Syracuse University
SURFACE

Dissertations - ALL SURFACE

12-2013

Putting the Ghost Back in the Machine: A Defense of Common Sense Dualism

Matthew Andrew Skene

Follow this and additional works at: http://surface.syr.edu/etd

 Part of the Philosophy Commons

Recommended Citation
Skene, Matthew Andrew, "Putting the Ghost Back in the Machine: A Defense of Common Sense Dualism" (2013). *Dissertations - ALL*. Paper 25.

This Dissertation is brought to you for free and open access by the SURFACE at SURFACE. It has been accepted for inclusion in Dissertations - ALL by an authorized administrator of SURFACE. For more information, please contact surface@syr.edu.

Dissertation: *Putting the Ghost Back in the Machine: A Defense of Common Sense Dualism*

Abstract: Gilbert Ryle once ridiculed substance dualism, describing it as the view that we are a "ghost in the machine." Since that time, substance dualism has found few defenders, and a presumption toward naturalism has dominated philosophical inquiry. Here, I offer an unapologetic defense this unfashionable view of the self. To do so, I first explain why philosophy should endorse a shift in method away from naturalism and toward common sense philosophy. I then show how, from within that approach, substance dualism is far better supported than its competitors. My defense of the common sense method rests heavily on an account of justification and of evidence defended by Michael Huemer, and results in the view that we should give strong presumptive weight to common sense. I proceed to argue that a commitment to active, immaterial minds is built into the conception of human nature used in everyday life, and therefore qualifies as a tenet of common sense. This fact gives a strong presumption to dualism. I finish by considering objections to dualism, and showing that they not only fail to overcome this strong presumption, but would fail to overcome even a moderate conservatism with regard to beliefs about the nature of the mind.

Putting the Ghost Back in the Machine: A Defense of Common Sense Dualism

by

Matthew Skene

Degrees Earned
B.A., University of Colorado, Boulder, 2002
Ph.D., Syracuse University, 2013

Dissertation
Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Philosophy.

Syracuse University
December, 2013

Copyright © Matthew Skene, 2013
Chapter 2 of this work is a slightly extended and modified version of a paper originally appearing in Philosophical Studies in March, 2013 vol. 163, number 2. Reused with permission.
All Rights Reserved

Acknowledgments: I would like to thank my committee members, Robert van Gulick, Andre Gallois, Michael Huemer, Kris McDaniel, and Mark Heller for all of their work in helping me complete the dissertation. I would also like to thank Jordon Dodd, Deke Gould, Andrew Corsa, and those who attended a pair of ABD talks at Syracuse University for comments on earlier versions of chapters of this dissertation. I would like to thank all of those people who talked through these ideas with me and helped me these issues. While there are numerous such people, including those already mentioned, I would particularly like to thank Saikat Guha for continually beating me in arguments and forcing me to try to come up with ones that could stand up during our discussions. You are missed, and I only wish you were here to see the final result. Finally, I would like to thank my wife, Andi Skene, for all of her patience while I completed this work and for all of her love and support.

Analytical Contents

Section I: In Defense of Common Sense Philosophy

Chapter 1: Philosophical Error and the Economics of Belief Formation

1.1: I begin by discussing some recent work in medical meta-research and in cognitive psychology about the effectiveness of academic research. John Ioannidis has argued to the satisfaction of most of his colleagues that the majority of positive findings published in peer reviewed academic medical journals are false. He claims that a large number of influences, including financial influences facing researchers and journals, and cognitive biases influencing research demonstrate that the majority of positive findings in medical research are actually false. In addition, work in cognitive psychology has shown that academics tend to be both less accurate than they believe and far more confident about their beliefs than the evidence warrants. The incentives faced by researchers, and the subconscious cognitive influences held by researchers both suggest that success rates of published work in potentially every intellectual discipline are likely to be far lower than generally believed.

1.2 I continue by discussing some recent work in behavioral economics by Bryan Caplan. Caplan discusses the issue of political error, and contends that the primary cause of poor policies in democratic societies derives from voter irrationality. Caplan's account of this phenomena rests in a general account of irrational belief formation which he calls "rational irrationality." I discuss this work, and its apparent consequence that incentive structures we have over our beliefs likely play an important role in explaining the presence of widespread error in any number of areas, including academic areas with low costs of error.

1.3: I examine the consequences of the previous sections for academic philosophy, and find that the results are not good. Academic philosophers face incentives that are worse than the average voter, and that incentivize certain types of error in the field. In addition, the discipline has fewer checks against error than other disciplines, and therefore is more likely to accept methods of belief formation that are subject to the problems that face other academic disciplines.

1.4: I provide an account of the mechanism by which non-epistemic incentives can infect academic research in philosophy. Specifically, I claim that the most likely cause of error in the field is the presence of under-scrutinized presuppositions of method endorsed by various philosophical fields. I explain how, given the presence of the incentives discussed in the previous section, presuppositions of method that permit the acceptance of beliefs that fulfill these incentives will lead to profligate but inaccurate literatures, and will appear to be sound to those who engage in them at the time. This account gives rise to a need among philosophers to critically evaluate the methodological assumptions we use in philosophical practice.

1.5: I begin that investigation by looking at the nature of arguments, and explaining why it is likely that philosophers dramatically over-estimate the evidential relevance of philosophical arguments as sources of evidence. This over-estimation of the relevance of arguments is a wide-spread philosophical error built into nearly every philosophical method. In addition, it is the type of error that fits the incentive structure described above. I conclude by suggesting that this

example gives us more reason to carefully reflect on even the most widely held presuppositions of philosophy. I also suggest that the right place to begin that investigation is with a good account of the nature of evidence.

Chapter 2: Evidence and Epistemic Justification

2.1: I begin to defend the account of evidence I endorse by examining Michael Huemer's Principle of Phenomenal Conservatism (PC). Huemer's principle states "If it seems to S that p, then S thereby has at least some reason to believe that p." Huemer takes this principle to be self-evident. I explain the principle, and evaluate this claim.

2.2: I discuss Huemer's self-defeat argument for (PC). Huemer believes that (PC) is a presupposition of rational thought, and that it is therefore impossible to argue against (PC) without assuming its truth. I consider two different versions of this argument, and explain why Huemer needs to accept a version that takes (PC) to be a necessary truth about the nature of justification.

2.3: I begin to explain why we should think that seemings are necessarily tied to epistemic justification. I provide an account of seemings that grows out of Huemer's work. Specifically, I claim that seemings are propositional attitudes that serve the function of allowing us to add to our existing beliefs in a way that is guided by our epistemic aims. I explain why this shows that the possibility of epistemic rationality depends on the exitence of seemings, and why this suggests that we should think that, in the absence of defeaters, seemings must confer at least some degree of justification.

2.4: I consider a series of objections to Huemer's position. First, I consider objections that suggest that only a limited subset of seemings confer justification, even in the absence of defeat. I argue that the proposed limitations to (PC) all stem from some misunderstanding of the principle, and that, in the absence of examples that fail to depend on such a misunderstanding, we should think that (PC) captures a necessary truth. Next, I consider objections that attack the broader epistemic framework developed in the paper. Specifically, I consider objections from John DePoe about the need for seemings as sources of evidence, an argument from Michael Bergmann objecting to the necessity of (PC), and an argument from William Alston against the connection between (PC) and deontological, voluntaristic accounts of justification.

Chapter 3: In Defense of Common Sense Philosophy

3.1: I begin with an explanation of the nature of common sense. Taking my cue from Thomas Reid, I argue that common sense beliefs are presuppositions of ordinary thought and discourse.

3.2: I proceed to explain in more detail what common sense is, and how to identify common sense beliefs. Reid suggests that one can identify common sense by (1) examining the structure of languages for commonalities, (2) identifying unstated assumptions of every day life that are necessary to make sense of ordinary human action and interaction, and (3) identifying widely

shared beliefs that seem to stem from surface-level, ordinary exercises of our intellectual powers, such as observation or intuition.

3.3: Having explained what common sense is, I proceed to discuss the elements of common sense philosophy. Common sense philosophy typically adopts the following methodological presuppositions: (1) A rejection of contextualist accounts of both justification and proof which suggest that philosophical contexts should accept higher or distinct standards from those used in ordinary life, (2) A broad-based foundationalism that takes all of our natural intellectual abilities as legitimate sources of knowledge and as the basis for our initial data in philosophy, (3) A negative view of the use of speculation as a source of evidence, (4) An affinity for at least some aspects of ordinary language philosophy, and an acceptance of the idea that ordinary speech is ontologically committing, and (5) a privileging of the authority of ordinary human judgment over theoretical or philosophical reasoning in cases where the two conflict.

3.4: I examine a theoretical defense of common sense that can be found in the works of Reid, G. E. Moore, and Michael Huemer. According to this defense, there are limits on the extent of rational belief revision connected to the nature of evidence and to the structure of justification. According to this view, the dependency of theoretical positions on non-theoretical starting points, and the strong presumptive weight given to widespread, heavily integrated aspects of our common sense worldivew grant them a near-immunity from revision on the basis of theoretical considerations. In addition to discussing these arguments, I offer my own version of this defense, drawing on earlier sections of the work.

3.5 I discuss the question of why we should think that common sense beliefs are correct. I explain how Reid's providentialist account of why we should trust our beliefs is unlikely to be accepted at present, and some of the consequences of this. I then provide my own account of why we should accept common sense. On my view, the reason to privilege common sense judgments isn't because they must be correct, it is because whether or not they are correct, they are the best we can hope to achieve given both human nature and the nature of evidence.

3.6: I evaluate common sense philosophy with respect to its ability to solve the methodological problem discussed in chapter 1. I explain how the elements of common sense philosophy discussed in section 3.3 include valuable and important limitations on philosophical positions that can curtail the influence of perverse incentives to a high degree. I finish with a discussion of the comparative value of the common sense approach and of naturalism, explaining why naturalism should be viewed as subject to the same basic flaws as most philosophical methods, given the results of the first chapter.

Section II: The Common Sense View of the Mind

Chapter 4: Mind and the Physicalist's World

4.1: I begin with a preliminary discussion of whether or not dualism is the common sense view of the mind. I discuss some previous claims made about this issue and conclude that while it appears to be the common sense view, we need a more in depth examination of the nature of the mind in order to be certain.

4.2: I discuss the various features the mind is typically thought to have. These include the fact that minds include thoughts and beliefs, that conscious experiences occur there, that these experiences include sensations and other qualitative states, that the mind is the locus of our choices about how to act, and that they are the essential parts of us.

4.3: I discuss the nature of physicalism, arguing that physicalism is committed to the view that mental states are realized by brain states. I provide an account of realization in order to explain the minimal commitments of physicalists in their efforts to account for the mind.

4.4: I divide the physicalist's projects into efforts to explain the apparent components of our mental states, such as qualia and content, and the apparent functional or causal role of our mental states. I provide a brief explanation of why our intuitions of distinctness in philosophy of mind provide a serious challenge for physicalists to overcome.

Chapter 5: Compositional Components of the Mind

5.1: I begin chapter 5 with a discussion of the role of intuition and introspection as sources of evidence in philosophy of mind. I explain why an advocate of the common sense method should trust introspection as a source of foundational judgment. I also explain why the best way to think of intuitions in philosophy of mind is in terms of thought experiments. I claim that thought experiments serve the role of getting us to attend more directly and carefully to specific aspects of our minds, while simultaneously helping us develop inter-subjective data to build our theories on.

5.2: I reiterate the basic case for dualism from the appeal of thought experiments in favor of dualism that we find in the literature. I claim that people naturally accept the conclusion of these thought experiments that thoughts and feelings aren't physical.

5.3: I examine specific thought experiments and intuitions discussed in the literature, including Leibniz' mill, Jackson's cases of Mary and Fred, Searle's Chinese room argument, and various issues discussed by Alvin Plantinga. I claim that the entirety of the appeal of these examples lies in getting us to focus on specific features of our mental states. They are, in a sense, ostensive efforts to get us to see which features the dualist thinks must be non-physical. Their appeal rests in our automatic understanding of these components of our minds as non-physical states.

5.4: I discuss the question of how to view the conclusions of these thought experiments. I claim that the thought experiments justify the claim that mental states are non-physical, not just the claim that mental states don't appear to be physical. I also discuss related objections to these conclusions that claim that the mind is mysterious in itself, and that dualism fails to alleviate this mystery. I reply to this objection by explaining why mental substances are uniquely suited to explain the existence of these features.

5.5: I provide an epistemological objection to functionalism. Specifically, I argue that functionalism cannot provide an account of how we learn the functional roles of mental states. Functionalists claim that our mental states are exhausted by their functional nature. However,

our knowledge of mental functions, especially purely internal ones, must derive from an introspective awareness of specific mental states fulfilling those functions. This awareness requires that those states have identifiable intrinsic components that we can track and recognize in order to learn those functions. Given this, functionalism appears unable to account for our knowledge of those aspects of our mental lives which they claim exhaust our mental lives.

5.6: I conclude the chapter by discussing some of the history of physicalism. Specifically, I claim that the explanatory gap used to be seen to establish dualism as the default view, and physicalists recognized a burden of proving that this gap could be overcome. However, despite failing to actually overcome it, physicalism somehow managed to become the default position in the field. I claim that this change was unjustified, and that the best explanation for the inability of physicalists to overcome the explanatory gap remains the initial dualist hypothesis that it is conceptually impossible that the intrinsic components of our mental lives could be physical.

Chapter 6: The Causal Argument against Physicalism

6.1: I begin chapter 6 with a discussion of the commitments of physicalism with respect to causation. I explain why physicalists are limited to event causal accounts of efficient causation given their view about the nature of reality. I further explain how the options available for event-causal explanations appear to be incompatible with the nature of free will and moral responsibility.

6.2: I provide a detailed account of the nature of agent causation in order to further elucidate the conflict between physicalism and agency. I describe agent causation as singular, non-derivative, and intentional exercise of the active power inherent in agents. This positive account provides content to the nature of agent causation and helps us see how this type of causation is regularly appealed to in human history.

6.3: I discuss some experimental data related to issues of free will and agency. First, I explain why Paul Bloom contends that people are natural dualists and I discuss how notions of agency and intentionality are central to this account. Next, I look at some results from experimental philosophy relevant to this issue. X-phi results have not been uniform when it comes to questions about our intuitions on free will and moral responsibility. I explain why the results supporting libertarianism are more reliable than the ones suggesting a more compatibilist account of moral responsibility, and why many of the results of x-phi experiments are likely to be questionable in this area.

6.4: I tie together the results of this chapter and the previous one to explain why they combine to produce a serious and perhaps insurmountable problem for physicalists. Realism in philosophy of mind requires sufficient fidelity to common sense concepts about the mind for the project to count as a reductive or realization account of the mind rather than an error theory about the mind. Given the conceptual inadequacies of the compositional and the functional or causal aspects of the physicalist's view of the mind, they don't appear to have sufficient fidelity to any aspect of the pre-theoretical understanding of the mind to grab a foothold to explain the rest of the features in terms of. Given this, it is likely that the only versions of physicalism with any hope of success will have to be anti-realist views such as error theories or eliminativism.

Chapter 7: Dualism as Common Sense

7.1: In chapter 7 I seek to develop an overall account of the mind that unifies the results of the previous chapters with some additional common intuitions about the nature of the mind we find in the literature. I begin with a brief description of the intuition that minds are, in principle, separable from bodies. I discuss how this intuition can be seen both in the history of philosophy and in the history of thought in general.

7.2: I continue by explaining how the intuitions about disembodiment connect to the role of the mind as a source of original causal power. I explain why substance dualism is best suited to explain the existence of a loci of a new type of efficient causation given its introduction of a fundamentally different player into the causal network.

7.3: I discuss yet another feature of the mind often thought to support dualism: subjectivity. I explain why a substance capable of subjective experiences is necessary for rational action, and how rational action is directly related to the existence of the compositional features of the mind discussed in chapter 5.

7.4: I explain how the results of the previous sections combine to form a comprehensive view of the mind and its place in human behavior. I explain how the possibility of rational choice created by the combination of subjectivity and the compositional features of the mind combine with the combination of an immaterial mind as a source of new causal power to provide a compelling picture of human beings as rational agents. I explain why a substance dualist view of the mind should be preferred to other forms of dualism that lose plausibility from their lack of comprehensiveness.

7.5: I tie the results of this chapter back to the issues discussed in chapter 1 by discussing whether or not we should be suspicious of dualism as an interesting philosophical view. I claim that dualism isn't really interesting. It is instead the combination of a few basic intuitions and introspective observations. It appears interesting only because philosophers have rejected it. I further explain why I think this should never have happened, and how strong the case against physicalism appears to me to be.

7.6: I sum up the section, and explain how it connects to section I and sets the stage for section III.

Section III: Objections and Replies

Chapter 8: Causal Objections to Dualism

8.1: I begin by providing an overview of the objections to be considered in section III. I point out that if someone has been fully convinced by the work of the first two sections, then they will see the task of the physicalist in overcoming the presumption in favor of dualism as nearly insurmountable. However, given that I can't reasonably expect everyone to have fully embraced

those conclusions, I set myself the higher task of showing that these objections fail to overcome even a far milder presumption in favor of dualism.

8.2: The first objection I consider rests on the intuition many people share that there is something odd or impossible about immaterial causation. I discuss a few examples of this intuition, and argue that they fail to provide a good objection to dualism. First, I point out that these intuitions are far from universal, and run counter to examples of causal stories people have often had no problem understanding or accepting. I also point out that the puzzlement seems to depend on insisting that the explanation fit the model of event causation used in other areas, and therefore begs the question. Finally, I point out that even if one struggled to provide an explanation of immaterial causation, this wouldn't mean we should ignore the apparent evidence of its existence, as uncertainty about how something happens is a weak basis for claiming it doesn't.

8.3: Jaegwon Kim has produced a far more developed version of the objection claiming that immaterial causation is problematic, which he calls the "pairing problem." This problem revolves around the issue of how we can pair up mental causes with physical effects in the absence of spatial relations. Kim's pairing problem can interpreted as either an epistemological or as a metaphysical worry. I explain why the epistemological worry fails to qualify as a genuine philosophical mystery, and then move on to discuss the metaphysical worry. Kim asks for a description of the relationship between the mind and the body that explains how they can interact without spatial relations. I describe the relationship between the mind and body as the combination of experiential awareness and of agency, or direct control of the body. I explain why this account can fulfill Kim's request in a way that isn't question-begging or empty.

8.4: I continue by moving on to objections grounded in empirical discoveries. I start with the discovery of neural correlates of mental activity we have found in the brain. I claim that these discoveries are fully compatible with dualism, and that the presence of regular, continuous correlations of this sort are exactly what we should expect once we think more deeply about the intimate and continuous nature of mind/body interaction. I also explain why the expertise of neuroscientists in their ability to reach these discoveries fails to give them any authority in questions about the metaphysics of the mind.

8.5: I consider several objections to agent causation. In 8.5.1, I consider van Inwagen's rollback argument, which claims that agents would be redundant as causes due to the presence of a statistical explanation of our behavior in terms of the probabilities of each of our actions occurring. I explain why the presence of statistical outcomes from a rollback scenario fail to provide an adequate account of our behaviors, and how this shows that van Inwagen's argument fails to provide a problem for the view. In section 8.5.2, I discuss the role of motives in human behavior and consider the objection that indeterminism negates the possibility of comprehensible action. I explain why this objection not only fails to show a problem for libertarian views, but also why libertarianism does a much better job capturing the data we have about free choice than does a view according to which motives are sufficient causes of our behavior. In section 8.5.3, I discuss the studies done by Liber and others on the timing of choices and influences of the will. Several studies claim to show that people are systematically mistaken about when they make choices, typically believing they have made a choice only after the relevant brain activity leading up to the action has begun. I explain why these studies can be better understood as reflecting the

time of resolution, and why we should think that physical activities would begin prior to resolving on a specific decision, and thus should result in data like we actually find.

8.6: I discuss the relevance of nature and nurture on human behavior and its relevance to the existence of free will. I claim that worries that our knowledge of the influences of these factors on our actions are has left no room for free will are dramatically overstated. Although there is a significant influence of genetics on human behavior and a varying but far less significant influence of upbringing and identifiable aspects of shared environment on human behavior, the data is nowhere close to establishing a deterministic view of human action. There is plenty of room left for free will to exert influence over our actions.

8.7: I discuss the claim that the conservation of energy is inconsistent with dualism, and explain why it is false. The law is compatible with interactive dualism for two reasons. First, the law discusses the overall quantity of energy in a system over time, and is therefore compatible with changes in distribution that don't produce changes in quantity. More importantly, however, the law is only intended to apply to closed systems. If interactive dualism is true, then physical bodies aren't closed physical systems; they are systems that are being influenced from the outside by an immaterial mind. As a result, the law of conservation of energy simply doesn't apply to human behavior if dualism is true.

8.8: I discuss the idea that the universe is causally closed, and the related objection to dualism outlined by Papineau. I explain why such a principle might be used as a guiding principle of science, but why this role it plays fails to have relevance in philosophical contexts, and fails to provide grounds for questioning dualism.

8.9: I finish the chapter by discussing the inductive case for physicalism grounded in the past successes of science. I claim that the inductive case is far weaker than is traditionally believed. First, the past successes of science have primarily served to overcome speculative views of cultures, not substantive data of common sense. Second, the current data we have in relevant areas such as neuroscience fails to justify confidence in its eventual ability to account for everything we do. Most neuroscientists interpret and describe their findings under physicalist assumptions. Those who look at the data without this guiding assumption often interpret it differently, sometimes even claiming a need for immaterial causes. These facts suggest that physicalism isn't a conclusion being derived from the data, but rather an interpretive constraint on how many researchers describe the data in the first place. Metaphysical assumptions of scientists, however, don't provide good guidance in philosophy, while the newness of the field together with facts about scientific practice suggest that even if there are serious problems with this assumption, we should not have expected these problem to cause a shift in this presupposition of those researchers.

Chapter 9: Non-Causal Objections to Dualism

9.1: I begin this section with a discussion of parsimony and its application to philosophy of mind. I examine some of the accounts of parsimony and argue that they either fail to offer support on either side of the issue or else offer support for dualism. In particular, I focus on a defense of parsimony grounded in anti-speculative philosophy, and argue that in philosophy of

mind it is the physicalist and not the dualist who is engaged in speculative ontology, and therefore physicalism is less parsimonious than dualism.

9.2: I discuss the relevance of introspective illusions and of thought experiments in philosophy of mind. I claim that the illusions we find in introspection can be easily accommodated by a reasonably fallibilist epistemology and don't justify skepticism about introspection in general. In addition to introspective illusions, Daniel Dennett has attacked the use of thought experiments in philosophy of mind. I provide a quick, general defense of the use of intuitions in philosophy and then explain why intuitions derived from thought experiments should be used as data points in philosophy of mind.

9.3: I consider Paul Churchland's claim that dualists are asking for a special epistemic status for introspection that we don't offer in other areas, namely that our experience fully reveals the underlying nature of what we are aware of. I explain why introspective data favoring dualism provides such strong evidence about the nature of our minds. First, I claim that in our experiences of all types we have a good sense of the categories into which the objects of awareness fall, even when we lack a full grasp of their underlying nature. Second, I claim that unlike in the case of perception, in introspection there doesn't appear to be an additional underlying nature beyond the qualitative aspects of the experience themselves. Since they appear to be irreducible, the suggestion that they have an underlying nature runs counter to experience.

9.4: I discuss some apparent materialist intuitions about persons and explain how they can be incorporated into a dualist metaphysic. I sketch a general ontology of persons according to which we are comprised of both a physical body and an immaterial mind in virtue of the causal interconnectedness of the two substances. This account can make sense of the idea that I have a certain weight and height without sacrificing the view that I am essentially a mind or a soul.

9.5: I conclude with a brief summary of the work and with some optimistic claims about the viability of common sense dualism. I explain why I have sought to establish such strong conclusions, and how much weaker versions of those conclusions would still suffice to justify substance dualism. I conclude by offering some reason to think that the establishment of common sense philosophy as a viable approach to intellectual inquiry should result in a return to dualism as the default view in the field.

Section I: In Defense of Common Sense Philosophy

Introduction:

The vast majority of human knowledge is obtainable in simple and direct ways. We know about the world around us because we can see it, we can touch it, and we can experience it directly. We know about our thoughts and our feelings, our abilities, our hopes and our fears because we can sense them within ourselves. We know about what happened in the course of our lives because we can remember it. And we know much about what is true and false in a variety of other matters because we can think, and we can understand and evaluate what we hear and read and consider when we do. Most of human knowledge is comprised of these surface-level facts available to all of us through the everyday use of our intellectual abilities. Sometimes, we can expand our knowledge beyond these surface level, easily accessible facts to know more than what is readily apparent. We can examine and study the world, we can think and reason about it, and we can engage in debate and seek to understand more than we previously knew. We can grow. Sometimes the discoveries, or apparent discoveries we come to in these endeavors are surprising. Sometimes, these efforts lead to results or views about the nature of things that conflict, or at least are in tension with, those things we thought we knew from the exercise of our everyday abilities. When such conflicts arise, what are we to do?

Over the past few centuries, the scientific approach to understanding the world has drastically increased the scope of human knowledge. The successes of science have motivated many to adopt a worldview grounded in this approach. This view of the world, known as 'physicalism,' takes the world to be entirely composed out of physical particles, and takes the successes of science to be grounded in the fact that science has the methods best suited for understanding the nature of our entirely physical world. This picture of reality comes into

conflict with a number of claims that people have made in the course of history. It conflicts, obviously, with the traditional religious views of the world, which claim that there are non-physical beings such as angels and demons, and of course, an all-powerful immaterial being: God. More than this, however, it at least apparently conflicts with the view that most people have of themselves; a view that seems to be grounded in those simple, everyday evaluations of our nature that we engage in. Traditionally, people have believed that they have souls. They believe that they are composed not merely of the matter arranged in their bodies, but also of a mind, or soul, that is believed to be immaterial, and thus not to fall under the scope of scientific investigation. This view of persons, according to which we are composed of both a physical body and an immaterial mind or soul is known as "substance dualism." Paul Bloom recently said

> I would put those who reject dualism in the same category as those who, through scientific reasoning or philosophical deliberations, come to believe that there is no external world, just sensory impressions (as did Bishop Berkeley), or that thoughts and feelings do not exist (as some radical behaviorists assert), or that there is no such thing as morality, or truth, or pain. These scientists and philosophers may be perfectly sincere in these beliefs. But such views are held at an airy intellectual level, slapped on top of our foundational appreciation that the world contains objects, minds, morals, truth, and experience.[1]

Bloom views physicalism as an *error theory* in philosophy of mind; a view that seeks to replace our ordinary understanding of something with a view that is thought, by its proponents, to better represent how things really are. Although most people acknowledge that dualism is the common

[1] Bloom, 2004 p. 191

sense view of the self, most people in contemporary philosophy of mind are willing to reject common sense on this point, and to adopt the physicalist view of persons; one according to which our minds are simply another part of the physical world. They think, in this case, that the view of the world suggested and endorsed by those who engage in scientific pursuits should trump our everyday understanding of ourselves, and that we should come to adopt a revisionist picture of human nature. This is understandable. It is also a mistake. The common sense view of the self is essentially correct, and it is being given up far too easily, and at far too great a cost.

To establish this, I will begin with a discussion of philosophical method. I will argue that in cases of dispute between the outputs of our more speculative or theoretical endeavors and those facts available to us in every day experience, the common sense view should win out. This is true both because theoretical conclusions are less reliable than it is commonly believed and because the nature of evidence supports the view that we should trust the reliability of our foundational judgments. The common sense approach to philosophy takes our pre-theoretical view of reality as the foundation from which to examine philosophical questions. As Reid puts it,

> In this unequal contest betwixt Common Sense and Philosophy, the latter will always come off both with dishounor and loss; nor can she ever thrive till this rivalship is dropt, these encroachments given up, and a cordial friendship restored: for, in reality, Common Sense holds nothing of Philosophy, nor needs her aid. But, on the other hand, Philosophy... has no other root but the principles of Common Sense; it grows out of them, and draws its nourishment from them. Severed from this root, its honours wither, its sap is dried up, it dies and rots.[2]

[2] Reid, 1997, p. 19

Underwriting the common sense approach is a position recently developed by Michael Huemer[3] known as "phenomenal conservatism." According to this view, we should believe that the world is the way it seems to be unless we have compelling reasons to change our minds. Initial appearances are granted foundational, defeasible justification, and they establish the presumptive positions in philosophy. The degree of justification an appearance holds depends upon its strength, and thus those appearances that are constitutive of our common sense view of reality are elevated, on such a view, to a degree of justification that makes them virtually immune to revision. These facts about the nature of justification legitimize the common sense approach to philosophy, including defenses of it offered by Reid and by G. E. Moore. In addition, common sense philosophy offers the best hope of overcoming some of the difficulties that commonly face theoretical reasoning and academic research.

After defending the common sense approach, I apply it to the philosophy of mind. Here, I argue that substance dualism best captures our initial understanding of the nature of the mind, and that it should therefore be the default position in the field. The current endorsement of physicalism in philosophy of mind is primarily motivated by an endorsement of the methods of science as our best chance to understand the nature of reality, and a related view of philosophy as subordinate to the work done in scientific fields. Philosophers, seeking to aid science in its pursuit of the truth, have sought to understand their own subject matter in ways that are compatible with and motivated by the appeal of the modern scientific worldview. This approach, known as 'naturalism,' is incredibly popular, and has influenced a great deal of the philosophical work over the last century. A shift from a naturalist to a common sense approach dramatically changes the evaluation of the various arguments and pieces of evidence we have available. The

[3] See Huemer, 2001, 2007

common sense approach gives no *a priori* advantage to any general theory about the underlying nature of the world, but rather sets constraints on what such a picture can look like which are grounded in our ordinary grasp of the nature of reality. The common sense approach invites us to take reality as we find it, full of whatever clearly seems to exist, and to build our metaphysics in a way that respects our ordinary understanding of the world. Freed from the methodological constraints of naturalism, we can see that the common sense view of the world includes an ontology of immaterial minds as well as physical bodies, and a view of human nature that is fundamentally at odds with the modern scientific worldview.

Given the initial defense of the common sense approach, this conflict suggests that we should embrace substance dualism. However, there have been a number of attempts to show that this view must be mistaken. Although it's status as common sense makes dualism highly resistant to revision, the position is still defeasible, and if a very compelling case can be made to justify changing our minds, then we should. The remainder of the work examines the case against dualism, and explains why it fails to justify the revision of common sense that physicalism requires. The objections fail both because physicalism is far more revisionist than people recognize, and because the arguments against dualism are much weaker than people take them to be. In fact, I argue that even a mild conservatism could not be overridden by the current objections to dualism, and thus even on more neutral grounds of evaluation than those I defend, the physicalist's efforts to overturn our ordinary view of the self fail.

Before moving on to this defense, it is worth explaining why this is important. Much of the support for the modern scientific worldview comes from the fact that whenever it has faced a conflict with some long-standing view of the world, it has triumphed. It has overcome superstition and dogma, and vastly expanded our knowledge of the world in the process.

Standing in its path therefore strikes many as a fool's errand. On its march toward explaining the world, this is what is next. I expect the fight over the soul to shape the atmosphere of the next century in much the way that the fight over evolution shaped the last one. Here, however, the fight is not between faith and reason or superstition and evidence. The fight is between our ingrained sense of who and what we are and a view of human nature that fits more neatly into the picture that has been developing of what the world as a whole is like. In the past, these debates led much of society to embrace irrationalism, anti-intellectualism, and a mistrust of science. Left without a rational avenue of defense, but unable to give up on a central aspect of what they believe, many turn away from rational debate. In a matter so central to our sense of self, such responses will only intensify, further marring the great achievements science, and leaving us all worse off.

There is a picture of ourselves integrated into our thoughts, our actions, and our sense of worth, freedom, and responsibility that, for most of us, cannot be abandoned, except at that "airy, intellectual level" Bloom spoke of. There is another picture of the world that stands, perhaps undefeated in its efforts to overcome the challenges it has faced. In the conflict between them, it is essential that our sense of self not be hastily abandoned out of a fear of a hefty opponent; that we fight for what we know and value about our nature, and that we evaluate carefully and honestly the evidence at hand. There is an alternative to the naturalist approach to evaluating the evidence, one that lies in a respectable tradition that can be well defended. From within this tradition, our sense of self remains whole, and remains rationally justified. This offers an opportunity for a fair fight, and for one that does not require us to ignore or mistrust the great achievements of science, achievements that are surely among the crowning glories of the human

race. It is a debate that can be had honestly, and openly, and rationally, and it is one that I believe will, at the end, vindicate the understanding of human nature we have always had.

Chapter 1: Philosophical Error and the Economics of Belief Formation

When I was young and naïve, I believed that important people took positions based on careful consideration of the options. Now I know better. Much of what Serious People believe rests on prejudices, not analysis. And these prejudices are subject to fads and fashions[4]*. –Paul Krugman*

Overview of Section I

This work calls for a rather dramatic shift in how people working in philosophy of mind should approach philosophical questions, and for a return to a now unpopular position about the nature of the mind. To begin, I will discuss why we should change our approach to addressing questions in this area, and in philosophy more generally. This effort begins with an explanation of some problems that face academic research, and some difficulties we all face in making rational judgments. I combine these issues to provide an account of one important cause of philosophical error. This account requires an explanation of the mechanisms that produce such error. I argue that the source of most error in philosophy comes from the uncritical acceptance of certain aspects of philosophical method that motivate the misapplication of philosophical reasoning to problems it is not best suited to solve. This should motivate us to examine foundational questions in epistemology about the nature of evidence. In chapter 2, I provide an account of the nature of evidence while defending Michael Huemer's "Principle of Phenomenal Conservatism." In the third chapter, I discuss the common sense approach to philosophy, and defend this approach on the grounds that it best solves the problem discussed in Chapter 1, and that it is well supported by the account of evidence offered in Chapter 2. This will create a basis for the shift in importance in the evaluation of the philosophical and empirical data relevant to

[4] Retrieved July 3, 2010 from http://www.nytimes.com/2010/07/02/opinion/02krugman.html?_r=2&ref=opinion

deciding issues in philosophy of mind. The implications of this approach are then pursued through the remaining chapters.

Overview of Chapter 1

Philosophers have often recognized shortcomings in our field. People are fond of saying that there is no idea so absurd some philosopher has not held it,[5] and recently, Eric Dietrich has gone so far as to claim that "there is no progress in philosophy."[6] While these are overstatements, difficulties facing the profession are hard to ignore. However, these problems have not inspired the level of self-reflection one might have hoped for. In this chapter, I seek to account for at least some of the causes of philosophical error. Here, I detail some problems facing academic research in other fields, as well as recent work in economics that depicts a general problem facing epistemic rationality, and show how the results of these studies can be brought to bear on understanding the difficulty philosophy has had in making steady progress and in solving longstanding problems. I conclude that much philosophical error arises from too quickly embracing methodological assumptions that create opportunities to defend false but interesting positions.

1.1 Challenges to Successful Academic Research

Recently, work has been done in several fields that calls into question the reliability of academic research. One of the most detailed accounts of this problem is in the area of medical research. By combining statistical modeling and some moderate assumptions about possible sources of bias, John Ioannidis has mathematically demonstrated that most positive findings in medical research are incorrect.[7] Specifically, he showed how starting with appropriate base rates, and taking into account the difficulties that face many actual studies, the probability that a

[5] This statement seems to have originated with Cicero in Book II of *De Divinatione*.
[6] Dietrich, 2011
[7] Ioannidis, 2005

given positive finding is accurate is less than 50%. Given the low initial probability of a positive correlation, the small size of typical positive correlations discovered, and the room to adjust parameters or otherwise allow bias to influence results, the majority of instances where researchers believe they have found a positive discovery turn out to be either statistical noise or errors resulting from either bias or methodological inadequacy in the study. This problem even affects the most often cited research published in high quality journals. According to Ioannidis, only 44% of highly cited studies are confirmed by future studies.[8] Although these finding are initially surprising and troubling, perhaps the best reason to accept them is that they are widely embraced by people in the field.[9]

Two of the most interesting conclusions Ioannidis reaches are that financial and professional considerations as well as levels of interest in a research area can both negatively influence reliability. The fact that financial considerations can affect the reliability of studies isn't very surprising, but financial incentives can take a number of forms. Some examples from Ioannidis include the observations that "Many otherwise seemingly independent, university-based studies may be conducted for no other reason than to give physicians and researchers qualifications for promotion or tenure,"[10] and that "With many teams working on the same field and with massive experimental data being produced, timing is of the essence in beating competition. Thus, each team may prioritize on pursuing and disseminating its most impressive 'positive' results."[11] These academic influences on what we decide to research should be familiar. The advice that one should pursue a line of thought because no one has done it yet, or it sounds interesting, or that you should try to carve out your own new position on an issue are

[8] Friedman, 2010
[9] Friedman, 2010
[10] Ioannidis, 2005, p. 0968
[11] Ioannidis, 2005, p. 0968

often given to graduate students trying to decide on what to do with their careers. The advice that one should only publish what they know to be true would likely be seriously professionally detrimental. But if academic disciplines know that pursuing the truth is a destructive professional goal, this suggests some awareness of the fact we face additional incentives in deciding what to research and publish.

Although practical incentives seem to influence the reliability of research, this isn't the only or even necessarily the primary source of the problem. Ioannidis also points out that "scientists may be prejudiced purely because of their belief in a scientific theory or commitment to their own findings."[12] He claims that emotional commitments to theories and to past assertions can lead both researchers and those who work with them to hold on to and continue to publish erroneous results. The attachment people feel to their views, especially once they have published them, has also been documented in other research. In one startling instance, a study involving over 284 experts on political and economic issues asked them to predict the likelihood of various future events. The results, as described by Daniel Kahneman, were that "people who spend their time, and earn their living, studying a particular topic produce poorer predictions than dart-throwing monkeys who would have distributed their choices evenly over the options."[13] Despite learning of their lack of success after the fact, "experts resisted admitting that they had been wrong, and when they were compelled to admit error, they had a large collection of excuses."[14]

This emotional commitment to one's own views isn't the only source of problems, however. Often one's success similarly hinges on accepting the views of colleagues. Work that diverges too far from prevailing assumptions often face greater difficulties in their efforts to

[12] Ioannidis, 2005, p. 0968
[13] Kahneman, 2011, p. 217
[14] Kahneman, 2011, p. 217-18

receive attention. This is one reason Ioannidis claims that often "the extent that observed findings deviate from what is expected by chance alone would be simply a pure measure of the prevailing bias."[15] If an erroneous assumption, such as a high level of confidence in a positive correlation between two things takes hold, then data is often collected in a way that will conform to this assumption, and inconsistent data is often ignored or explained away. In this way, errors can spread quickly through a field and corrupt the reliability of the results in that field.[16]

One might have hoped that these problems would be solved by peer review. After all, academic journals are supposed to prevent inaccurate findings from seeing the light of day. One reason to be skeptical of such hope is that it is unclear how reliable peer review is as an evaluation of quality. Research into the inter-rater reliability of peer review has consistently shown only a small positive correlation between the opinions of reviewers. For example, a study led by Richard L. Kravitz concluded that at the *Journal of General Internal Medicine* "reviewers…agreed on the disposition of manuscripts at a rate barely exceeding what would be expected by chance."[17] Indeed, the study found a higher correlation between a single reviewer's evaluation of two separate works than between the ratings of two reviewers on a single work, suggesting that the personality of a reviewer is a more important determinant to publication than the quality of the work itself. A meta-analysis of the research on inter-reviewer reliability reached the overall result that across studies it was shown that "reviewers agreed in their evaluations for 17% more of the manuscripts than would have been predicted on the basis of chance alone."[18] Given this, it is unlikely peer review can adequately correct for these problems. In fact, journals also face almost the same incentives as researchers when deciding what to

[15] Ioannidis, 2005, p. 0970
[16] This result fits very well with the account of error-producing mechanisms in philosophy I develop later.
[17] Kravitz, et al, 2010
[18] Bornman, et al, 2010

publish. Journals want prestige, and they obtain it in part by being the first to publish novel and interesting findings. They also typically reject papers at very high rates in an effort to maintain this appearance of prestige. This selectivity ends up weeding out methodologically sound works that obtain boring, negative findings or that confirm existing findings in favor of papers that have something interesting to say, even when they are less likely to be correct. This problem of unnecessarily high rejection rates in medical research eventually led to the establishment of a special online journal, the Public Library of Science, with lower rejection rates, which only considers the quality of the work and ignores the interestingness or the originality of its findings in its decisions on what to publish. The need for this is a clear recognition by the field that there is a problem with the incentives faced by the majority of journals in deciding what to publish.

Even in a context such as medical research with potentially high practical costs of error and rather well-established methods for ascertaining the truth, intelligent, qualified people more often than not find themselves publishing false things. Given that most academic fields share these features, it would be surprising if most academic disciplines didn't have similarly high rates of error in their publications. If it happens in fields that possess genuine, measurable progress and important consequences for people in the real world, then it likely infects all intellectual pursuits. Impractical areas of belief formation with little progress are certainly among the most suspect, as are ones with weaker constraints on method or with fewer external measures of accuracy to test against.

1.2: Preferences over Beliefs

The problems in academic research discussed above appear to fit a pattern of belief formation recently developed by economist Bryan Caplan. In *The Myth of the Rational Voter*, Caplan argues that most voters are irrational in the formulation of their political beliefs. Caplan

focuses on the relevance of what he calls "preferences over beliefs." Caplan treats beliefs as a consumable good. We value beliefs for a variety of reasons, and it is easy to think of them as providing a benefit to the holder, and of being obtained as a means of procuring that benefit. In most cases, the value of a belief is a product of its truth. If my belief about how far away a cliff's edge is from me is false, it has virtually no value, and if it is off by much, it is likely to have a high disvalue. Generally speaking, we cannot afford to be very wrong in our beliefs, and our primary aim in forming them is to believe what is true, since only then is the belief likely to be valuable. However, truth isn't the only thing we care about. As Caplan points out "on many topics, one position is more comforting, flattering, or exciting, raising the danger that our judgment will be corrupted."[19]

When we form beliefs in order to meet the goal of believing what is true and avoiding error, we are epistemically rational. But since beliefs can fulfill other goals, it is possible for there to be "beliefs that are irrational from the standpoint of truth-seeking [but that] are rational from the standpoint of individual utility maximization."[20] This creates the opportunity for what Caplan calls "rational irrationality" in the formation of our beliefs. Given the fact that beliefs can be formed in order to satisfy non-epistemic motives, it is important to know which sorts of beliefs we are most likely to be irrational about. Here, a straightforward cost/benefit analysis can be developed. Beliefs with low practical costs of error typically won't change the value of the believer's life no matter what opinion they hold. Therefore, if there are non-epistemic motives one could satisfy by forming a belief, and these motives outweigh one's desire to believe what is true about the issue, it would be practically rational to follow those motives. One could then

[19] Caplan, 2007, p. 155-16
[20] Caplan, 2007, p. 141

obtain the benefit of holding the belief she wants without having to pay a cost for getting the belief wrong.

Caplan contends that in the case of political beliefs the cost of holding false beliefs is virtually negated. Since your opinion will only be put into effect if enough people share it, your own political opinions have effectively zero impact on which policies get enacted.[21] Even if the effect of the decision being implemented is high, even for the voter herself, she is unlikely to cast the deciding vote, and so the truth is pragmatically irrelevant to the individual voter's belief. As Caplan puts it

> Since delusional political beliefs are free, the voter consumes until he reaches his 'satiation point,' believing whatever makes him feel best. When a person puts on his voting hat, he does not have to give up practical efficacy in exchange for self-image, *because he has no practical efficacy to give up in the first place.*[22]

Freed from the individual costs of error, voters can instead consider the value of these beliefs in terms of their cohesion with the beliefs of their friends, the sort of self-image such beliefs produce, and, generally, how happy or smart or clever or kind it makes us feel to hold the belief in question.[23] It is worth pointing out that deliberate self-deception is rarely involved. Rather, we seem to shift our methods of evaluating evidence to better suit our new aims.[24]

A similar effect can be seen in the research problems discussed earlier. The set of incentives faced by researchers and by journals converge to create a situation where accuracy

---

[21] See p. 17-18 of Caplan, 2007 for a more detailed explanation.
[22] Caplan, 2007, p. 132, emphasis original
[23] The voter, of course, doesn't do all of this self-consciously, but there are a variety of ways to rationalize one's own views after the fact. Michael Huemer has an interesting treatment of the irrationality of political beliefs that deals with such mechanisms and with Caplan's earlier work in an unpublished paper "Why People are Irrational About Politics" available on his web site at http://home.sprynet.com/~owl1/irrationality.htm.
[24] For a good discussion of the mechanisms involved in rationalization, see Kornblith, (1999). For Caplan's approach, see pp. 125-131

must compete with non-epistemic motives. The fact that practically important areas also see this problem may initially appear as a challenge to the view that practical costs affects the rationality of beliefs. However, it is worth remembering who is negatively affected by false beliefs about things like medicine. As Caplan points out, "errors with drastic real-world repercussions can be cheap for the individual who makes them. How? When most or all of the cost of the mistake falls upon strangers."[25] Typically, errors in research don't cause harm to reserachers. Since most diseases are rare, and since findings rarely claim dramatic effects of drugs or behaviors on disease, the results rarely motivate behavioral changes for those who understand them (although public reaction can be inappropriately dramatic). None of this shows that the research is bound to be corrupted. But it does help explain how researchers can be insulated from the cost of error even about issues that have great practical importance for other people. Like for the average voter, the cost of error for the individual researcher is quite low. Since philosophical beliefs aren't known for their practical consequences to their holders, it is reasonable to wonder if non-rational preferences over our philosophical beliefs affect philosophical practice, and if so, what we can do about it.

1.3: The Problem in Philosophy

The existence of regular error in philosophy is often recognized by people working in the field. I believe the problems discussed in the previous sections can help us account for such error. To see if this is true, we should look at the incentives philosophers face in acquiring their beliefs. One way to look at this is to think of what philosophers are selling. Philosophers make a living by getting people to want to know what we think about a particular issue. To get people to want to listen to us, we need to make claims that they will find interesting. In order to get people to listen to what *you* think, instead of what other people think, you need to be able to say

[25] Caplan, 2007, p. 121

things that are original. To make a career of this, we need people to keep talking about us, and the easiest way to do that is to say something controversial, in hopes of creating a lively debate. Therefore, if one wishes to be published and widely read, it is a good idea to publish views that are interesting, original, and controversial. Unfortunately, there are only so many truths like this to be discovered. The set of false claims has a far greater variety of interesting and original positions to draw from, and since it is rare for everyone to embrace the same error, they'll almost always be controversial. So, there is a practical reason for philosophers to find ways to endorse false views. Given the absence of the practical costs of error in philosophy, this suggests that there is likely to be a problem in philosophy analogous to the ones we find in other areas. The problem can be captured by the following argument:

(1) Principle of Epistemic Rationality: (PER)- It is not epistemically rational to believe something just because it is interesting, original, or controversial.

(2) Publishability Fact: (PF)- One is more likely to be published in philosophy if she is able to say things that are interesting, original, and controversial.

(3) Perverse Incentives: (PI)-The practice of philosophy incentivizes non-rational motives of belief formation; specifically, it incentivizes believing things that are interesting, original, and controversial, whether or not they are true.[26]

(4) Sad Truth: (ST)- Almost all claims that are interesting, original, and controversial are false.

(5) Unfortunate Conclusion: (UC)- The practice of philosophy encourages philosophers to say and to publish things that are false.

[26] This claim is less substantial than it may appear at first glance. The claim is purely about the possibility of satisfying a set of relevant aims for publication in the absence of a concern for truth, it does not say that philosophers actually fail to care about the truth.

I think the first premise of the argument is self-evident, and the second is an obvious empirical observation. (3) follows from (1) and (2) plus the uncontroversial assumption that philosophers have an incentive to publish papers, and the hopeful assumption that philosophers believe what they say.[27] (4) is the least obvious of the premises, but I think some reflection will bear it out. For a true view to be both interesting and original, the discovery must involve a case where someone sees a valuable truth that has eluded the attention of her epistemic peers. Such discoveries do happen, but for the them to happen regularly enough to allow thousands of researchers in a discipline to each find their own interesting and original truths on a somewhat regular basis would require a widespread blindness to important but discoverable truths on the part of those peers. It is unlikely that human intelligence is sufficiently diverse to create prolific discoveries of new truths in longstanding fields. In addition, controversial views must either go against what most people believe, or else take a stance on an issue where there are already a significant number of opinions. Given the variety of opinions in philosophy, it appears to be a mere statistical point that the majority of those views must be inaccurate. As Richard Fumerton points out, "each positive philosophical view is usually such that most philosophers think that it is false, and there is typically nothing approaching a consensus on the correct alternative. I can, therefore, infer that most philosophical views are false" (Fumerton, 2010, p. 109). It seems, therefore, that most things people would find interesting, original, and controversial are things that are likely to be false. Since (5) follows from (3) and (4), the argument appears to be sound.

Considered merely from the perspective of the incentive structure, the position of the average philosopher is actually worse than that of the average voter. The things that influence

[27] I tend to think this is correct. Philosophers don't often strike me as insincere, and it is usually difficult for people to defend and fight for views they don't actually accept. The suggestion that philosophers subconsciously fall prey to problems commonly faced by other researchers and by humans in general strikes me as far more charitable than the suggestion that philosophers are academically dishonest.

the average voter tend to be irrelevant to the truth of the claims they hold. Voting typically fulfills our aim of expressing something about who we are. In philosophy, our ideas aren't simply expressions of ourselves, they are intended for an audience, and therefore have to satisfy the interests of that audience. There are likely many things that a philosopher's audience values, hopefully including the truth. However, insofar as they are interested in features such as the ones we have discussed, there will be an incentive to try to meet these demands, and work that is able to meet them will be more successful. Captivating or enchanting ideas require specific characteristics that aren't required for the views of the average voter. These characteristics, however, are often positive indications that the belief is false. Given these incentives, there is a serious danger of philosophy struggling to make progress when work is evaluated in a way that emphasizes these features. If we find reason to believe that this problem is more than theoretical, then any philosophical method that hopes to produce genuine progress or even consensus on these issues must find ways to restrict such incentives.

The above argument shows that there is potential for philosophers to be motivated to adopt false beliefs. However, the fact that there are incentives to adopt false views doesn't, by itself, show that anyone will actually follow those incentives. Even in a situation where there is a benefit to be obtained from following motives that tend to lead to error, we will only consume as much error as we can stomach. Since philosophers are supposed to love the truth, there is at least some hope that we won't be as prone to error as we otherwise could be. In addition to this, there is some progress made in philosophy.[28] This is particularly true in ethics. Revolutionary ideas such as the contention that we shouldn't oppress women and minorities, or that average people should have a fair degree of control over their own lives have won out and become stable

[28] For a dissenting position on this issue, see Dietrich (2011)

components of our modern worldview.[29] However, these successes exist against a backdrop of fairly extensive and routine error. Such widespread error among many of the greatest thinkers in human history calls out for some explanation.

One reason to believe that philosophy is being influenced by incentives to believe false things is the things that philosophers have actually said. The peculiarity of philosophical beliefs has not escaped the attention of philosophers in the past, and examining their complaints should help us see that there is a problem in the field that requires an explanation. David Stove, for example, was so vexed by philosophical error that he once called for a nosology of its causes.[30] Others, meanwhile, have bemoaned the effect of some of the more extravagant reaches of philosophical error on the reputation of the field. Thomas Reid, for example, frequently speaks of the responses of "the vulgar" to the theory of ideas in all its forms as one of deserved ridicule. Richard Fumerton recently complained about the political views of philosophers, claiming that "people who clearly are in a position to know better use arguments that they would never endorse but for a desire to reach a certain conclusion," and goes on to say that "If there is one thing I can be virtually certain of, it is that most philosophers are *not* reliable when it comes to arriving at interesting philosophical truth. And it does not help much to turn to "brilliant" philosophers."[31] This idea is also present in Descartes motivation for engaging in philosophy. He perhaps comes closest to expressing the theme of this chapter when he says

> For it seemed to me that I might find much more truth in the reasonings which each one makes in matters that affect him closely, the results of which must be detrimental to him if his judgment is faulty, than from the speculations of a man of letters in his study which produce no concrete effect and which are of no other

[29] Huemer provides a nice account of this point in his (2012), p. 223-4, 232-33.
[30] Stove, 1991, p. 187
[31] Fumerton, 2010, p. 107

> consequence to him except perhaps that the further away they are from common sense, the more vanity he will derive from them, for he will have had to use that much more skill and subtlety in order to try and make them seem dialectically probable.[32]

In addition to seeing this problem in the views of others, philosophers have also occasionally recognized these problems in themselves or in the community as a whole. Socrates, for example, admits in the *Phaedo* that

> …at last I concluded myself to be utterly and absolutely incapable of these enquiries, as I will satisfactorily prove to you. For I was fascinated by them to such a degree that my eyes grew blind to things which I had seemed to myself, and also to others, to know quite well; I forgot what I had before thought self-evident truths…[33]

Here, Socrates acknowledges the effect that fascinating ideas had on his ability to reason accurately about certain issues that interest him. Thomas Reid similarly claimed that philosophers find more cautious approaches that limit imaginative speculation difficult, saying

> To a man of genius, self-denial is a difficult lesson in philosophy as well as in religion. To bring his fine imaginations and most ingenious conjectures to the fiery trial of experiment and induction, by which the greater part, if not the whole, will be found to be dross, is a humiliating task. This is to condemn him to dig in a mine, when he would fly with the wings of an eagle.[34]

[32] Descartes, 1968, p. 33
[33] Plato, *Phaedo*, 95 b-c
[34] (Reid, 2002, p. 535

Reid says elsewhere that "It is genius, and not the want of it, that adulterates philosophy, and fills it with error and false theory."[35] In both these instances, Reid acknowledges that our desire to prove our creativity and brilliance drives us to endorse any number of views and to accept methods that will allow us to do so, even when they lead to error. This recognition by philosophers of certain propensities toward error in the field also seems introspectively accurate. Honest self-assessment frequently brings with it an awareness of certain shortcomings in how we think about issues.

While it is true that philosophy has made some progress over time, this progress has not been anywhere near the level or rate of progress one would have expected from some of the most brilliant minds in history dedicating their efforts to the same areas of study for thousands of years. In addition to this, what progress we have made has been regularly accompanied by large areas of study where not only is progress absent, but the sorts of beliefs philosophers advocate include ideas that are among the most peculiar and implausible claims anyone has ever made. An occasional error grounded in confusion or mistakes in reasoning is understandable, but one would not expect drastic, frequent, and persistent error in a field seeking to get at the truth. The fact that there are incentives to believe interesting ideas, however, does a good job of explaining how so many brilliant minds could consistently endorse a rather wide array of errors. Given this, we have some reason to think that philosophy has often succumbed to such incentives.

1.4: Error Producing Mechanisms

One problem this view faces is that we know our colleagues and ourselves, and we often find it hard to believe that we are being dishonest or unscrupulous in forming our beliefs. Any account that suggests that professional philosophers don't care about the truth or about good evidence is going to face strong and well-justified resistance. Before this claim can be accepted,

[35] Reid, 1997, p. 15

we need some account of the error-producing mechanism that does not require us to reject the evidence of our own experience in dealing with our colleagues and thinking about our own beliefs. One reason to recognize the possibility that such explanations exist is that are found so often in other areas. People often have inaccurate beliefs about the scope of their expertise or the rationality of their judgments. Kahneman reports on several instances where experts mistakenly think that their views in an area are well supported when they are not. For example,[36] real estate agents think that they are not subject to priming effects when evaluating the value of a home, stock brokers believe their knowledge of how stocks work will provide them with a large advantage in the stock market, and psychologists believe their expertise in understanding people's current mental state will allow them to predict long-term recovery rates accurately. In each case, these self-evaluations have been disproven by empirical study. Kahneman reaches the conclusion that "the unrecognized limits of professional skill help explain why experts are often overconfident."[37] A reasonable amount of humility should prevent us from thinking that we aren't equally likely to be blind to certain shortcomings in our thought. As for the source of those shortcomings, when Caplan sought to identify errors in economic beliefs of the average voter, he looked for identifiable biases built in to how we think about economics[38]. Given the fact that methodological presuppositions are an effective way to skew our evaluation of data without our realizing it, we should look first for the source of philosophical problems in the various methods that philosophers endorse.

Biases that remain part of our means of evaluating evidence rather than exerting themselves as overt influences over belief formation do occur, and are typically unrecognized by

[36] Kahneman discusses these cases in his (2011) in section, parts 2 and 3.
[37] Kahneman, 2011, p. 240
[38] Caplan identifies four key biases the average person has about economics: anti-market bias, anti-foreign bias, make-work bias, and pessimistic bias. For a discussion of these see pp. 30-48

those who have them.[39] In addition, it is in background assumptions about how to form our beliefs where biases tend to be most effective. Given the incentive structure we have seen above, errors of method that justify thinking in ways that allow us to pursue these incentives will provide ample opportunity to write interesting papers, at least until and unless the illicitness of the underlying assumptions are brought to light. Since excessive, deliberate, self-aware irrationality isn't plausibly attributed on a large scale, this avenue provides a much better means of identifying how people find ways to justify evaluating evidence in a way likely to permit widespread irrationality. Philosophical methods that allow us to believe we are earnestly and properly pursuing the truth while simultaneously allowing us to pursue implausible and original avenues of thought will lead to well-intentioned errors. Given the incentives in the field, these methodological errors will lead to widely published and potentially very influential work. Since an apparently productive presupposition is likely to earn greater support from the fact that work is frequently published by those who assume it, this could quickly and easily lead to large portions of the philosophical community suddenly embracing such presuppositions. So, if an illicit presupposition can find its way into our thinking, we will likely not only fail to see it, but we will come to view it as a valuable guiding assumption that allows us to produce publishable work. This seems to provide a legitimate account of a mechanism for satisfying these incentives that needn't appeal to implausible or nefarious motivations of individual philosophers.

On the account we will be examining, it is thought that philosophers are subject to unconscious biases grounded in an unquestioned acceptance of some aspects of philosophical method. If the practice of philosophy encourages methods of belief formation that are influenced by these motives, then well-intentioned people will come to think they are pursuing the truth

[39] For an insightful discussion of the role of implicit bias in the profession, particularly with respect to things like hiring decisions, see Jennifer Saul (forthcoming)

when they follow methods that are used to justify error. Indeed, methods that inculcate error at their core are far more likely to create successful philosophers, since they will legitimize approaches that allow one to publish regularly. The incentive structure therefore does not to have to work on individual philosophers in a self-conscious way. Instead, the discipline is set up to both encourage errors grounded in philosophical method and to legitimize those methods in the field by rewarding approaches that permit fascinating errors by producing well-respected publications and the illusion of progress. This realization helps us move from *ad hominem* attacks on particular people to a more charitable and plausible claim that methods inculcate irrational evaluation procedures by masking them as virtues or at least as harmless heuristics.

Kahneman has noted the willingness of theorists to blindly accept such errors, and has coined the phrase "theory-induced blindness" to describe it. He defines this phenomenon as follows:

> once you have accepted a theory and used it as a tool in your thinking, it is extraordinarily difficult to notice its flaws. If you come upon an observation that does not seem to fit the model, you assume that there must be a perfectly good explanation that you are somehow missing. You give the theory the benefit of the doubt, trusting the community of experts who have accepted it.[40]

Even in otherwise ideal theoretical contexts, such blindness can create problems for some time. In ones where the errors create personal success and the illusion of fruitfulness, their grip will likely persist even longer, and people will feel even less inclined to question them. This problem makes it even more important to engage in a constant evaluation of methodological assumptions in modern academic life.

[40] Kahneman, 2011, p. 273

Given what we have seen, methods of evaluating evidence that open up areas of interesting, original, and controversial beliefs as legitimate candidates for consideration should be suspect. Such methods establish mechanisms for creating richer arrays of publishable ideas, and can be adopted as part of a philosophical approach while bypassing excessive scrutiny. In addition, methodological assumptions that instill biases that permit a greater capacity to satisfy our incentives will lead to prolific fields (if not successful ones). Therefore, it is in such background methodological assumptions that we should seek out the mechanisms for allowing us to seriously entertain views that are likely to be false. There are no doubt a number of these sources of error, but for now I want to focus on one nearly universal bias that seems pervasive in philosophy in order to illustrate the problem. If a nearly universal presupposition of philosophical method can be shown to fit this pattern, this should provide us with good reason to think we need to look more carefully at our philosophical methods, no matter who we are. This pervasive assumption is the widespread and effusive praise of arguments as sources of evidence for our views.

1.5: Arguments and Rational Irrationality

Teaching students that arguments are the holy grail of evidence is standard philosophical practice. Jerry Fodor and Ernest Lepore, for example, have said that "what there is no argument for, there is no reason to believe" (Fodor and Lepore, 1992, p. xiii) and go on to refer to this assertion as a "chilly methodological truism." However, this claim is highly implausible. It entails, among other things, that virtually everything that every person in the history of the world has ever believed, they have believed for no reason at all. And yet it is merely a clear expression of an attitude that most philosophers have to claims that aren't supported by an argument. In nearly every approach to philosophy, very high value is placed on rational arguments as sources

of evidence. Given the nature of what an argument is, however, I can't see how this level of enthusiasm for arguments can be well founded. Since a disproportional appreciation of arguments as sources of evidence would be precisely the sort of background assumption that we are looking for, its near universality in philosophy would be good evidence that philosophy needs to carefully re-evaluate its methodological presuppositions, if in fact such appreciation is misplaced.

To see why we should think that the evidential relevance of arguments is overstated in philosophy, we should once again look at a more common problem with human reasoning and then consider the usefulness of arguments in light of this problem. One of the more common sources of error in contexts where people are reaching new conclusions based on evaluations of new considerations is a tendency to ignore the importance of base rates, or prior probabilities to the overall results. A good example of this phenomenon was discussed by Eliezer Yudkowski during a rather successful attempt to provide an intuitive explanation of Bayes' theorem.[41] There, Yudkowski discusses a question asked to groups of doctors where they are asked to assess the probability that a patient has breast cancer given that the patient has a positive result on a mammogram. Here, the patient has no other signs of cancer, and is in a demographic with a 1% base rate of cancer. When a patient has cancer, the test indicates this fact 80% of the time. However, it also delivers false positives 9.6% of the time. When asked what the probability is that a patient with a positive result has cancer, only about 15% of doctors get the correct answer, with the majority of respondents believing the answer to be between 70% and 80%.[42] The correct answer, however, is less than 8%. The example shows that for a proposition with an initial probability of 1%, if there is a test that is usually 80% reliable at indicating the truth about

[41] This example can be found at http://yudkowsky.net/rational/bayes/. The page clearly explains the math behind the correct answer, if the answer sounds wrong to you.
[42] Yudkowski cites several studies replicating these findings, including (Casscells, et al 1978).

the question, but has about a 10% chance of delivering a false positive, then the probability of the proposition after a positive result is still less than 8%. The problem the doctors faced here was in appreciating that one's conclusions in the presence of new data needs to be tempered by the facts one had before examining that data. If the base rate or initial probability is very low, then it takes a lot more than some data, even reasonably reliable data, indicating a different answer in order to become justified in changing one's mind. This problem of ignoring base rates has been detected in many other fields[43] as well, and appears to capture a basic human tendency to engage in a reactionary bias to new considerations.

There is no uncontroversial way to establish initial probabilities in philosophy, but there are certainly pre-theoretical evaluations of how likely a view is to be true. To evaluate how much evidence we can obtain from an argument, it is worth remembering what an argument is. An argument is a structured set of propositions where we are encouraged to believe one of them on the basis of the others. The conclusion of an argument is viewed, by the advocate, as in need of evidential support, which it is supposed to obtain from its connection to other claims. In order to draw support from them, however, those other claims must be better supported than the conclusion. Since the conclusion of an argument is dependent on the credibility of its support, conclusions typically must be less well supported than the premises they depend on. Arguments can, of course, give us *good* reason to believe, but only if we can already have *better* reason to believe the premises on which they depend. This fact about the evidential dependence of conclusions on premises shows that there is an asymmetry between philosophical conclusions and the evidence they rest on. Without this, there could be no basis for allowing some propositions to serve as premises while believing that others require evidential support.

[43] For more examples, see Kahneman, (2011), esp. chapters 14 and 23.

Given the dependence of conclusions on their premises for support, there is good reason to think that philosophical reasoning depends on the idea that propositions come to us with some level of initial probability. Given this, there is reason to think that philosophers might, at times, fall prey to the sort of irrational rejection of base rates present in other fields. I believe that some of the praise of arguments present in philosophy is due to this sort of problem. Consider the idea that arguments can be used as useful tests of accuracy. Mirroring the earlier example, an appearance of soundness can be seen as a positive result, and illusions of soundness as false positives. Arguments usually have superficially plausible premises and appear to be valid most of the time. Suppose that the likelihood of a superficially plausible argument being sound in general is 80%, but that people are also subject to a cognitive illusion of soundness 10% of the time. Then, for a proposition whose initial probability is 1%, the degree of confidence we have in the proposition after hearing an apparently sound argument for it should not rise above 8%, and so should not be believed, even in light of the argument. Even if one thinks that the rate is something more like 90% to 5% accurate findings to false positives, the probability of a proposition with a 1% prior probability is still only 15.4%, which far too low to justify belief. However, even the initial numbers respectively seem too high and too low. Philosophical arguments usually don't have premises that, in conjunction, exceed an 80% rate of apparent truth, and given the number of times subtle logical errors have been found in arguments, the rate of illusions of apparent soundness is likely much higher than 10%.

It has been said that there is no position so absurd that some philosopher has not believed it. I doubt this is true, but it is certainly true that the literature is full of examples of beliefs that, prior to philosophical inquiry, have probabilities around 1% or lower. Prior to philosophical inquiry, the idea that tables are mental images, that people don't actually have beliefs or

experience sensations, or that feeling hungry is causally irrelevant to eating is so remote and bizarre that the priors for such positions must be very low. In general, skeptical or highly counter-intuitive philosophical conclusions often have initial probabilities of 1% or less. Having a good argument for such conclusions therefore fails to confer anywhere near adequate justification for belief, given facts about the reliability of appearances of soundness for philosophical arguments. This would seem to show that having a good argument for a position is insufficient grounds for belief. To be adequately supported by an argument, a position must first have a high enough prior probability. Similarly, for a belief to be overturned by an argument, it must have sufficiently low prior probability. If the prior for a belief is 99%, it can't be hurt too greatly by an argument; certainly not enough for rejection, although possibly enough to withhold judgment. This would seem to show that there is an important limit to the support an argument can offer for a philosophical position, and a limit to the role that arguments should play in certain philosophical discussions. Not all conclusions are fair game, for instance, just because you can create an argument for them. Reid puts the point as follows

> The conclusions of reason are all built upon first principles, and can have no other foundation. Most justly, therefore, do such principles disdain to be tried by reason, and laugh at all the artillery of the logician, when it is directed against them.[44]

The scope of this mistake is further demonstrated by the preference many philosophers have for complex arguments over simple ones. The probability of an illusory connection between premises and conclusion increases whenever the argument is more complex, and the probability of soundness should diminish as the number of premises increases. Thus, the best arguments in philosophy would have to be the simplest ones, since they would diminish the risk of error on

[44] Reid, 1969, p. 172

both of those scores. And yet it is complex or creative arguments that usually receive the most praise and attention. However much we should appreciate the skill involved in crafting such arguments, their complexity alone should give us good reason to be concerned about the accuracy of the conclusion, particularly if the conclusion seems unlikely on its own.

Forming accurate beliefs in speculative and complicated fields is never a sure bet. We are at our worst in forming judgments when we are asked to move beyond the data we have available to us in everyday life and easily confirmable empirical investigation and into theoretical areas. Typically, the rational thing to do in such a situation is either to form hesitant, cautious judgments or to refrain from belief entirely. When forced to take a side as part of your life, as we are asked to do in democracies and highly motivated to do in philosophy, we are pushed to evaluate the evidence incorrectly. On evidential grounds, the philosopher's love of arguments seems unfounded. It can, however, be explained given the incentive structure discussed earlier. Philosophers are good at arguing, and people are impressed by interesting arguments. Instilling the belief that any view you have a good argument for is a view worthy of serious consideration allows anyone who is sufficiently creative in argumentation to satisfy their incentives to hold false views while simultaneously getting them to believe that they are virtuously pursuing the truth, since they are following "sound" philosophical practice of developing interesting arguments for their views. If even such a basic methodological assumption as the idea that rational belief formation critically hinges on philosophical argumentation is found to be an illicit methodological presupposition, then less pervasive standards that similarly open up areas of belief formation that initially appear quite unlikely are also likely to be suspect.

Concluding Remarks

Problems in philosophy often seem to last forever, changing form, but rarely producing results. Often philosophers I have spoken with casually express the view that this is okay because philosophical problems aren't meant to be solved, they are only meant to be discussed. I find this attitude toward philosophy both baffling and deeply troubling. Disciplines have no reason to continue to exist if they don't produce anything of value. The track record of philosophy in terms of production are not good. There seem to be two main possibilities as to why philosophy hasn't been able to produce many stable results. The first possibility that is that the questions that philosophers ask are simply too difficult to answer; that people are incapable of discovering philosophical truths, or of recognizing them when they do, and so philosophy as a discipline is doomed to remain free of results, or at least of any serious consensus. The second is that there is something wrong with how philosophers engage in the practice that keeps them from recognizing the truth when they find it, or of pursuing it in ways that are likely to be fruitful. I have argued that the second answer is the correct one. This fact may be somewhat uncomfortable to admit for a discipline filled with such great minds, but overall I think it should give us hope. If the lack of productivity in philosophy has a cause that rests in our practices rather than in the nature of the problems it faces, then by correcting these problems we can eventually be more successful in establishing lasting and valuable results. Given the importance of so many of the issues addressed by philosophers, such accomplishments should be well worth the effort.

The next two chapters take up the challenge of articulating a well-grounded method that can overcome these difficulties. Eventually, I will argue that common sense philosophy does the best job of fighting the incentives that often lead to error in the discipline. Before I get there, however, it would be good to have an account of the nature of evidence to work from. With such

an account, one would be in a better position to see whether or not various methodological assumptions are ones that are focused on our proper epistemic motives, and which ones appear to instead serve the function of allowing us to accept interesting but improbable conclusions. The account of evidence I endorse arises from Michael Huemer's Phenomenal Conservatism, according to which evidence is entirely a matter of how things seem. In chapter 2, I explain and defend this account before using it in chapter 3 to evaluate and defend common sense philosophy.

Chapter 2: Seemings and the Nature of Evidence[45]

Overview of Chapter 2

Having examined a practical problem facing philosophical methods, we should consider approaches to philosophy that might do well at solving this problem. In chapter 3, I will discuss common sense philosophy and explain why it does the best job of fighting those incentives discussed in chapter 1. Before examining the comparative virtues of this approach, however, it would be good to have a clear sense of how to adjudicate between these positions. In order to do so it would be helpful to have a clear sense of the nature of evidence we have when we are deciding what to believe. If we have a clear sense of what evidence consists in, then we should be in a better position to evaluate the merits of various methods in terms of whether or not their views accord with or conflict with a proper understanding of the nature of evidence. Here, I defend an account of evidence grounded in a position developed by Michael Huemer known as "Phenomenal Conservatism."

Michael Huemer's principle of phenomenal conservatism states

> (PC) If it seems to *S* that *P*, then, in the absence of defeaters, *S* thereby has at least some degree of justification for believing that *P*.[46]

While Huemer's principle is true, it is not obvious why seemings are centrally tied to justification. Since Huemer takes (PC) to be the sole fundamental epistemological principle, it is important to understand this connection. The chapter proceeds as follows. Fist, I explain Huemer's account of (PC), and evaluate his claim that (PC) is self-evident. Next, I discuss Huemer's self-defeat argument, offering two formulations of it, and arguing for the stronger version. I then examine the nature of seemings to show how the right account of seemings

[45] A slightly different version of this chapter appeared in Philosophical Studies, Vol. 163-Number 2, March, 2013. Reused with permission.
[46] Huemer, 2007

explains their connection to justification, and motivates the stronger version of the self-defeat argument. Finally, I examine several objections to (PC), showing how the understanding of seemings I offer helps us to respond to them, and further supports the position.

2.1: Explaining (PC)

Two important things should be noticed about Huemer's principle before evaluating it. First, Huemer's principle concerns an internalist account of justification[47]. An internalist account of justification is one according to which justification supervenes entirely on facts about our mental states. What a person should believe is determined by what the person can do to satisfy her epistemic goals given her epistemic position; that is, it is the question of how to be epistemically rational in forming our beliefs[48]. Huemer's answer to this question is that if we believe what seems true, then, in the absence of defeaters, we believe in an epistemically rational way. As he puts it "if my goal is to have true beliefs and avoid having false ones, and if *P* seems to me to be true, while I have no evidence against *P*, then from my own point of view, it would make sense to accept *P*."[49] This internalist notion of justification is crucial to Huemer's account, and is important to keep in mind when evaluating (PC) and the arguments in this paper.

The next thing to notice is that (PC) provides a perfectly general account of justification. Although developed in defense of direct realism about perception, the view covers all types of appearances. As Huemer points out elsewhere

> …we can say 'This line seems longer than that one', 'I seem to recall reading something about that', 'It seems to me that I have a headache', and 'It seems that

[47] While some people are interested in (PC) as a principle of externalist justification as well, Huemer states that the principle deals with internalist justification in his (2001), p. 104.

[48] Not all versions of internalism focus on our epistemic goals in this manner. Since Huemer's version clearly does, however (see, e. g. his 2001, p. 115), I will be focusing on such versions of internalism here.

[49] Huemer, 2001, p. 104

> any two points can be joined by a single straight line'. All of those statements make sense, using the same sense of 'seems.'[50]

On Huemer's view, seemings are present in the exercise of every faculty of judgment. Their presence is what makes it rational for us to adopt the beliefs we do. Whenever we make a judgment, we are evaluating how a proposition seems with respect to its accuracy. When a proposition seems true, whether perceptually, mnemonically, introspectively, or intellectually (or in another way, if there are others) we thereby have some degree of justification for believing it. Since all faculties of judgment are uniform with respect to how we use them to make epistemic evaluations, Huemer's position provides a unified account of the source of all justified beliefs.

Having articulated (PC), we can consider the claim that it is self-evident. According to Huemer, we are justified in believing something if we form our beliefs in a way that is reasonable given our evidence. The evidence we have available depends on how things seem. So, Huemer's claim essentially says that when something seems true, and we don't have good reason to think it isn't, it's reasonable for us to believe that it is. Seen in this light, the principle seems self-evident to me. However, many people have denied (PC), or have denied that they find it self-evident. I believe that this results from a failure to appreciate or accept some aspect of Huemer's broader epistemological framework in which he places (PC). By articulating this framework more fully, especially by explaining the nature of seemings, I hope to make the truth of (PC) more readily apparent. In addition, Huemer offers arguments to help motivate (PC). We will turn to one now.

2.2: The Self-Defeat Argument

One argument Huemer offers for (PC) is the "self-defeat argument." Whether one can justifiably deny phenomenal conservatism depends on whether one can rationally form the belief

[50] Huemer, 2005, p.99

that (PC) is false without depending upon (PC)'s truth. Huemer takes rationality and internalist justification to be equivalent. He says "another word for what is justified, or *should* be done or believed, from the first person perspective is 'rational.'"[51] Huemer thinks other positions are self-defeating because when forming beliefs "there is no (rational) escape from how things strike *you*."[52] That is, he thinks we cannot rationally form beliefs unless phenomenal conservatism is true, because we cannot rationally form beliefs except by forming them on the basis of how things seem to us. This reasoning suggests the following formulation of the self-defeat argument:

> (1) For it to be possible to be justified in any belief, it must be possible to rationally form beliefs.
>
> (2) If phenomenal conservatism is false, then it's not possible to rationally form beliefs.
>
> (3) Therefore, if phenomenal conservatism is false, then it is not possible to be justified in any belief.
>
> (4) Therefore, if phenomenal conservatism is false, it is not possible to be justified in believing that phenomenal conservatism is false. That is, any denial of phenomenal conservatism is self-defeating.

Given an internalist sense of justification, (1) must be right. (3) clearly follows from (1) and (2), and (4) clearly follows from (3). The crucial premise, therefore, is (2). There are two ways to read this premise given different senses of 'possible'. A weaker reading treats it as a contingent fact about creatures like us that we can rationally form beliefs only in this way. A defense of this position requires an empirical account of our psychology when forming beliefs with our epistemic goals in mind. A stronger reading would take the premise to apply to all possible

[51] Huemer, 2001 p.22
[52] Huemer, 2001 p.105

epistemic agents, and to claim that it is a necessary truth about justification that no being could rationally form beliefs unless (PC) is true. A defense of this stronger claim requires an *a priori* argument showing that the nature of justification precludes the possibility of rationally forming beliefs in any way other than by believing what seems true.

So which version should Huemer adopt? Given the role (PC) plays in Huemer's epistemology, he needs to defend the stronger version. According to Huemer "the correct fundamental epistemological theory will be a necessary truth that, when combined with the contingent circumstances of the world, yields appropriate verdicts on which beliefs in that world are justified."[53] Given that the necessary truth in question is supposed to answer all questions of justification when combined with the contingent features of *any* world, Huemer needs an account of why seemings confer justification that doesn't depend on contingent facts about our psychology, but instead captures something about the very nature of justification. Since (PC) serves the role of a fundamental epistemological principle, we need an *a priori* defense of premise (2). This defense must show that justification necessarily depends on how things seem to us. At first glance, it is not obvious why we should believe that (2) captures a necessary truth rather than a contingent one, but I believe this is because we have not examined seemings closely enough. I will seek to rectify this by giving a detailed account of seemings which explains their connection to justification.

2.3: The Nature of Seemings and their Role in Justification

We have seen that on Huemer's view, all beliefs are justified on the basis of seemings. Given this, we need a clear understanding of what seemings are. Huemer does not spend a great deal of time explaining the nature of seemings. Part of the reason for this is that he expects us to recognize them introspectively, at least when prompted with examples. We are sufficiently

[53] Huemer, 2007, p. 37

familiar with seemings to do this, and I think we can do so well enough to see that his principle is true. However, recognizing these states is different from having a clear understanding of their nature. I want to address this by providing an explanation of the nature of seemings in line with Huemer's which allows us to see why seemings are necessarily connected to justification.

To begin, we should compare and contrast seemings and beliefs. Seemings and beliefs are both propositional attitudes, but the nature of the attitude is slightly different. We can capture the difference by examining the sentences used when talking about these states. It makes sense to say "I believe that p" and "it seems to me that p" but it does not make sense to say "I seem that p" or "it believes to me that p." Beliefs are *affirmed* mental representations, and affirmation is a mental activity we take ourselves to be responsible for. This suggests that a similar but passive attitude attends seemings. Seemings play an instructive role in our mental lives, and like verbal instructions, we receive them passively. The instructions they provide are typically followed and therefore lead to beliefs. Seemings, then, are states that tell us what we *should* believe. They are mental representations of what is *to be affirmed* rather than what is *in fact affirmed*. Seen in this light, (PC) is the claim that following the instructions seemings offer is the source of all justification.

Since seemings always occur in cases of awareness, which Huemer offers an account of, we should continue our examination of their nature by studying this account. It goes as follows:

> S is aware of x if and only if:
>
> i. S has an assertive mental representation (an apprehension),
>
> ii. x exists and at least roughly satisfies the content of that representation, and

> iii. it is not accidental (not due to chance) that the content of the representation is satisfied.[54]

Since seemings have to do with internalist justification, they cannot be a matter of either (ii) or (iii). Seemings, then, must be a matter of the 'apprehension' or assertive mental representation. So what does Huemer say about apprehensions? Huemer explains that "Apprehensions are *mental* representations… [that] represent their contents *as actualized*."[55] Seemings are a species of apprehensions just as beliefs are. Both states represent their content as actualized. Huemer explains this feature of "representing as actualized" as follows:

> This feature I refer to as their 'assertiveness,' based on the following analogy to sentences. Consider the sentences, 'Please close the door,' 'Will you close the door?' and 'You will close the door.' These sentences have common content—they are all about closing the door. The first sentence requests or commands that the state be actualized, while the second asks whether it will be actualized. Only the third is assertive: it represents that state of affairs 'as actual,' that is, as what is actually going to happen.[56]

Huemer proceeds to give a functional definition of apprehensions. According to Huemer, apprehensions are mental representations that fulfill "the function of representing the world as it actually is."[57]

Given this functional account of apprehensions, the distinction between different species of apprehensions must rest in a difference in how this function is fulfilled. We entertain content in mental representations for a variety of purposes. Sometimes, such as in cases of creative

[54] Huemer, 2001, p.55
[55] Huemer, 2001, p.22
[56] Huemer, 2001, p.53
[57] Huemer, 2001, p.54

imagination, we are not concerned with the truth of the content. In these cases, the representation is not an apprehension. Given the central importance of truth to our success in planning and in negotiating our way through the world, however, it is hardly surprising that some representations fulfill the role of representing the world as it is. We want to acquire information about the world around us and to use the information we have in helping us to live our lives. Beliefs serve the role of letting us know what we have already affirmed about the world so that we can use the information they contain in our planning. Before we form beliefs, however, we need some means of evaluating what is true. Seemings, rather than letting us know what we already affirm, serve the function of representing what is true for the purpose of adding to our knowledge. To be able to form beliefs in a way that is guided by the aims of believing what is true and avoiding error is what it is to be an epistemic agent. Therefore, if it can be shown that the ability to form beliefs based on seemings is required in order to be guided by this aim, then it can be shown that one cannot be an epistemic agent unless one can reasonably form beliefs based on seemings. This is what I will attempt to show now.

The need for seemings is clearest when considering what it is to make an epistemic judgment. Epistemic judgments are evaluations of propositions with respect to their accuracy. Such evaluations can only occur if it is possible to discriminate between propositions with respect to truth and falsity. Since the content of a proposition cannot, by itself, permit us to make this discrimination, the only way such judgments could occur is if there is something in addition to the content that provides us with an indication of the content's accuracy, and thereby gives us epistemic guidance (that is, instructions or at least indications about what we should believe). Here, the analogy between apprehensions and assertions becomes clear. Both are norm-governed in that they aim at representing the truth, and both have features that permit their

content to be presented in a way that indicates this aim. Just as speech acts could not allow us to make judgments unless the utterances whose content was to be believed could be differentiated from those whose content was intended to serve another role (as in the case of questions or commands), mental representations could not allow us to make judgments unless they allowed us to differentiate their content from mental representations whose function was not to indicate what was true. Without this, there would be no way to discriminate between what is true and what is false, and therefore it would not be possible to make any epistemic judgments. In other words, if entertaining content gave us no idea what to do with that content, we could never have a reason, from our own point of view, to do one thing rather than another with it. And if it gave us no indication of how to go about fulfilling our epistemic aims, we could never have an epistemic reason to form or not to form a belief with corresponding content; that is, we could never make an epistemically rational judgment. So, the possibility of epistemic rationality is dependent on the presence of apprehensions.

However, not all apprehensions are suitable candidates for indicating what we should believe. In the case of agentive apprehensions (beliefs) the content is represented as true because we have affirmed it. Engaging in this affirmation, though, is typically what making a judgment is[58], and so it cannot justify the judgment itself. The apprehension permitting the discrimination involved in judgment must be passive rather than active. But as we have seen, to have a passive apprehension that provides us with epistemic guidance is to have a seeming. Therefore, if the fact that something seems true does not give us justification for our beliefs, then there is simply no way that we could ever rationally form a belief. This is true, not because of our psychology, but because of the nature of justification itself. Judgments cannot occur without seemings, and

[58] I say "typically" because sometime we make affirmations without making judgments, such as in cases of wishful thinking or in cases of arbitrary belief formation.

one cannot be epistemically justified in a belief if one cannot make judgments about the truth of the content of that belief. Therefore, it is simply not possible to be an epistemic agent unless we can use seemings as a source of justification in forming our beliefs.

To give an overview of the preceding reasoning, here is the argument for the view that justification necessarily depends on seemings:

(5) One can rationally form beliefs only if one can make epistemic judgments (evaluations of their truth) about entertained content.

(6) One can make epistemic judgments about entertained contents only if there are passive apprehensions that allow us to make those judgments.

(7) Passive apprehensions that permit us to make judgments are seemings.

(8) So, one can rationally form beliefs only if they can form them based on seemings.

Given this argument, we are in a better position to see why (2) would be true, and thus why any position that denies phenomenal conservatism must be self-defeating. The question of justification in the internalist sense simply cannot arise if there either are no seemings or if seemings are incapable of providing some degree of justification. Given this, to deny the relevance of seemings to justification would be to deny that we are capable of having epistemic justification in the first place. This includes the possibility of having epistemic justification for forming the belief that (PC) is false.

One of the primary reasons Huemer offered for accepting (PC) was the claim that only seemings were capable of making our beliefs rational[59]. I have explained why we should accept this claim. This gives us some reason to think that (PC) captures a necessary truth about justification. However, while (8) shows that we can *only* be rational by forming beliefs based on

[59] This seems to be what Huemer has in mind when he claims "If phenomenal conservatism is false, so that the way things seem to oneself is irrelevant to epistemic justification, then all judgment must be irrational" (Huemer, 2001, p. 107).

seemings, (PC), and premise (2) of the self-defeat argument, claim that any time we form beliefs based on seemings in the absence of defeaters, we act rationally. Given that only seemings are capable of providing us with justification, we should certainly hope that this is true. After all, if, even in the absence of countervailing reasons, we still aren't justified in accepting the evidence provided by seemings, then it is unclear how rational belief formation could proceed. However, if only certain seemings provide justification, it could turn out that sometimes when we form beliefs based on seemings, even in the absence of defeaters, we nonetheless believe irrationally.[60] So, for (8) to provide adequate support for (2), it would need to be supplemented with a reason to believe that we should not place such limits on the scope of (PC). As we turn to considering objections to (PC), we begin with objections that take issue with this claim. After closing this gap in the overall argument, we will examine objections that take broader issue with the epistemic position as a whole.

2.4: Objections and Replies

2.4.1 Restrictions to (PC) and Related Concerns

Restricted versions of phenomenal conservatism could grant that seemings are necessary for justification without maintainng that all seemings are suitable sources of justification, even in the absence of defeaters. Laurence BonJour, for example, holds that only the seemings of introspection and intuition confer justification.[61] Such a view could avoid the self-defeat argument so long as the remaining seemings justify the denial of (PC) in its perfectly general form. In "Compassionate Phenomenal Conservatism", Huemer responds to BonJour and similar views by challenging advocates of less inclusive, discriminatory positions to meet the burden "of identifying an epistemically relevant difference between those appearances that do and those

[60] Given the above argument, if this is the case, then only those remaining seemings could ever provide justification for our beliefs.
[61] BonJour, 2004

appearances that do not confer justification on their contents."[62] Huemer then looks at the reasons offered for the proposed restrictions and explains why he finds them lacking. I think Huemer's replies are successful against the objections he considers. However, if (PC) captures a necessary truth, there should be a general strategy of response to these objections. In light of the account of seemings provided here, I believe we are in position to develop such a strategy.

Before looking at the details of the response, it will be helpful to look more closely at when seemings occur in the process of belief formation. This will help us partition the proposed restrictions on the scope of viable seemings. As we have already seen, seemings differ from beliefs. Deciding whether or not to form a belief follows the occurrence of the seeming, and is distinct from it. In addition, before we entertain propositional content, the content is shaped or conceptualized. The raw data of experience is not the same as the propositional content that is entertained or evaluated. An example of this comes from Thomas Reid. Reid pointed out that when looking at a sphere of uniform color "the eye originally could only perceive two dimensions, and a gradual variation of color on the different sides of the object,"[63] but we quickly acquire the ability to take this data and conceive of it very differently. As "experience...teaches [us] that the variation of colour is an effect of spherical convexity, and the distribution of light and shade…the acquired perception in a manner effaces the original one; for the sphere is seen to be of one uniform colour, though originally there would have appeared a gradual variation of colour."[64] In this case the seeming attends an experience of a three-dimensional object of uniform color. Since the seeming attends the content as it is experienced rather than the raw data, it is clear that prior to the occurrence of the seeming, the content was shaped or conceptualized in order to allow us to evaluate it properly. So, there is a time before

[62] Huemer, 2007, p. 32
[63] Reid, 2002, p.236
[64] Reid, 2002, p. 237

the presence of the seeming where the content is shaped, an occurrence of entertaining the content where the seeming occurs, and a time after its occurrence where the content is processed. If (PC) is necessarily true, then the apparent problems with certain seemings must be caused by a confusion related to one of these areas. If we can uncover these errors, we should be able to develop general strategies of response. While this won't provide a demonstrative proof of (PC)'s necessity, it should provide us sound reason to accept it in the absence of objections that circumvent these replies.

We begin with objections focusing on what happens after a seeming occurs. These objections claim that despite the presence of a seeming, additional considerations show that the belief is unjustified. Matthias Steup, for example, raises a case of echolocation.[65] He asks us to suppose that James, the first patient to be given echolocation, has been told that it might not work, and that if it doesn't, then he will have hallucinatory experiences. After the surgery, James is in a dark room and seems to have an experience of a table. According to Steup, James should not believe the table is there. Instead, he should withhold judgment about its existence. Steup claims that Huemer's position would entail that James should believe the table is there because it seems to be and James has "no evidence for believing that sense experiences of that kind are unreliable."[66]

Steup is right that withholding judgment in this case is the correct approach. However, his objection requires an uncharitably restricted view of the nature and extent of defeaters. Defeaters are not limited to reasons to believe that a kind of experience is unreliable. Huemer's view is that the proper doxastic response is entirely determined by seemings. Since he also asserts that defeaters are relevant to justification, it is clear that he cannot believe that seemings

[65] Steup, 2004, p. 413
[66] Steup, 2004, p. 416

only occur in the initial evaluation of content. Seemings occur not only immediately, but also in processes of reflection and evaluation that follow the initial experience. Any subsequent seeming could override the support the initial seeming offers. In Steup's case, there is something James believes, namely that there is a chance the experiment did not work, that, when taken into consideration, changes how things seem overall. The fact that the experience may be hallucinatory is a reason to withhold judgment precisely because it makes it seem like there may not be a table there after all[67]. There are any number of things we might be aware of that could change how things seem to us. When they do, Huemer's position requires us to modify our doxastic response accordingly. If a problem in processing the content of a seeming is relevant to internalist justification, it will be within the scope of S's awareness. If the problem would change how things seem to S, as these objections suggest, then it would change what should be done when processing what initially seemed true. This is the purpose of the defeaters clause in (PC). Insufficient clarity about the ease with which defeaters can be obtained seems to motivate this type of objection. The first strategy of response, then, is to clarify the extent of defeat to demonstrate how proposed counter-examples simply provide examples of defeaters.

To examine the next two categories, it will be helpful to look at an example from Peter Markie. Markie offers the following proposed counterexample to (PC):

> I…see a beautiful sunset, and, having learned that I gain a feeling of spiritual peace from believing that the angel Gabriel exists on the basis of such experience, I so believe. Early in my learning, I was conscious of what I was doing as I responded to the experience with my belief and gained my feeling of peace. I

[67] This claim should not be confused with the claim that it would make the initial seeming stop occurring. The same content can appear to be true when considered in some ways and not seem to be true when considered in others. Consider, for example, the apparent length of the lines in a Muller-Lyer appear when you look at them compared to their apparent lengths when you measure them.

> now make the connection without thinking about it. In the course of forming my belief, it also seems to me as if Gabriel exists.[68]

Markie claims that although this belief will seem true just like most of our ordinary experientially based beliefs do, this belief is improper since "it has been formed relative to the epistemically improper end of gaining spiritual peace."[69] Markie claims that Huemer's position cannot explain what is improper about the Gabriel belief, since all justification is a matter of seemings on Huemer's view, and this belief seems true.

The right way to respond to this example depends on how we understand the case. The three important claims here are that (A) the belief seems true, (B) the belief is formed in order to provide spiritual peace, and (C) forming beliefs for that reason is epistemically improper. Although (A) stipulates that the belief seems true, and (B) and (C) together show is that the belief is not justified, by themselves these claims are fully compatible with (PC). One can have justification for a belief and still not be justified in holding it if the belief is formed for a reason other than the one that provided justification. So, even if a claim's seeming to be true provides one with justification, one's belief can still be unjustified if they don't believe it *because* it seems true. Therefore, to complete his counter-example, Markie must have been assuming something in addition to these three claims.

One way to complete the counter-example would be to add the assumption that in this case the proposition's seeming to be true *is the same as* its satisfying the goal of obtaining spiritual peace.[70] On this interpretation, the reason he actually believes and the thing that supposedly gives him justification would be the same thing, and the above reply would not be available. However, given the account of seemings on offer, we can see that this assumption is

[68] Markie, 2006
[69] Markie, 2006
[70] The relata of this identity statement are motivations, or reasons for belief.

incorrect. Notice that this account of seemings is purely functional. Mental representations count as seemings in virtue of fulfilling the role of providing us with doxastic instructions *so as to fulfill our epistemic goals of believing what is true and avoiding error*. Therefore, if a representation serves the aim of instructing us on what to believe for non-epistemic reasons, then it does not count as a seeming by definition. This is not a mere semantic point; it is central to the overall cognitive framework. We have a variety of motives, or reasons for and against the adoption of our beliefs. However, *epistemic* agency requires us to have specific aims directed toward the truth. Fulfilling such aims requires a type of motive for the adoption of beliefs that is guided by this goal. Given the presence of a motive that has this characteristic function, we will be able to pursue our epistemic goals by following the indications they provide. Seemings are these motives for belief characterized by this function. Although we certainly have other motives for adopting beliefs, including the sense of peace they may provide, following such motives rather than the evidence provided by how things seem is the essence of epistemically irrational behavior. When we believe in accordance with how things seem, however, we are rationally guided toward the truth, and so following such motives is constitutive of epistemic rationality. This conflation of seemings with non-epistemic motives in the production of our beliefs seems to be the origin of the objections that focus on improper motivations for belief formation. Carefully individuating such states should suffice for dealing with objections grounded in the influence of such motives.

There is, however, another way to construe Markie's objection. In the previous version, I assumed Markie conflated seemings with another motivational state, and I claimed that he therefore assumed that the belief seemed true when it did not. Given the details in the case about how these experiences have changed over time, however, it is possible that Markie is suggesting

that the belief about Gabriel now seems true, but that the seeming was caused by epistemically improper influences prior to entertaining the content, and that these causes corrupted the reliability of the seeming. Since the subject no longer recalls the influence of the prior motivation, he now seems able to trust this seeming, even though it is corrupt. This seems problematic to many people. Surprisingly, Huemer is (or at least was) one of those people. Huemer offers a nearly parallel case in "The Problem of Memory Knowledge." There, Huemer considers the following problem for foundationalist views of memory:

> Suppose that I initially accepted the existence of life after death by wishful thinking. I now no longer remember where I got that belief, but I just seem to remember that that's something I know. On the other hand, my brother Pete adopted the same belief in exactly the same way. However, his memory is better than mine, so he also remembers how he got the belief. As a result, my belief system is rational and his is not. That seems wrong.[71]

Huemer proceeds to defend what he calls a "dualistic" theory of justification for memory; one that does not fit well with his later work, or with (PC).

To the best of my knowledge, Huemer has never explicitly responded to the position he advocated in his earlier paper. I would like to produce a reconciliation by offering a general response to these worries. To begin, remember that we are dealing with an internalist account of justification. Huemer recently explained his underlying reason for endorsing internalism, which he calls the "central intuition of internalism about justification," and I label (CII). By examining it, we should see how to respond to this sort of objection.

> (CII) there cannot be a pair of cases in which everything seems to the subject to be the same in all epistemically relevant respects, and yet the subject ought,

[71] Huemer, 1999, p. 348

> rationally, to take different doxastic attitudes in the two cases—for instance, in one case to affirm a proposition and in the other to withhold"[72]

Now, given the subject's epistemic state in the situations described, what response would be more reasonable for S than belief formation? For these particular beliefs, one may think there are available defeaters, but supposing, as the cases would have it, that these individuals have no accessible reason to doubt the legitimacy of their current seemings. Then on what grounds could they rationally reject their evidence? I can't think of an epistemic position that would constitute a more proper response to their current epistemic state than adopting the relevant beliefs. Asserting otherwise seems to amount to rejecting internalism, or at least (CII), since it requires S's beliefs be responsive to features of her epistemic state that she has no capacity to take into account.

While the above may be right, it does seem that we should be able to say *something* about why Pete is not a worse epistemic agent than his brother, and why Markie's belief about Gabriel is troubling. In response to the objection that Utilitarianism fails to appropriately account for the significance of motives, Mill once asserted that "these considerations are relevant, not to the estimation of actions, but of persons."[73] Echoing Mill, it seems to me that the right thing to say is that the above objections are relevant, not to the estimation of beliefs, but of persons. It is certainly true that there is something epistemically blameworthy and contemptible about someone who deliberately corrupts their ability to reason. Indeed, Reid contends that "a greater crime can hardly be supposed" than that of "a man [who] has entirely destroyed his rational faculties." And yet he also says that such a man could not "[incur] new guilt when he was not a

[72] Huemer, 2006, p.151 Huemer explains in the paper why each version of internalism should adopt such a principle. Versions of internalism that reject deontogical accounts of justification, however, may be resistant to such a principle, and this response would likely not be available to those who would reject such a conception of internalism. I will say more in defense of deontological views later in the paper.
[73] Mill, 2009, 337

moral agent."[74] I think a similar point follows here. While these believers certainly have done something wrong, forming the belief in question is not it. Judging individual actions, or individual beliefs, is not the same as judging character, and the epistemic character of an individual need not be a function of the justification of each individual belief. In this case, the agents have engaged in the act of making themselves inferior epistemic agents, and for this, we should lower our evaluation of their epistemic character. Having done so, however, they have made it possible for them to be blameless in particular beliefs they form, since these problems of character, in lying outside the range of available evidence, cannot provide a reason for the agent to engage in a different response. This is true even though, at least when the problem results from his own behavior, it provides us with good reason to diminish our epistemic evaluation of the individual. This division between reasons for diminishing our evaluation of an epistemic agent and reasons for denying the justification of particular beliefs of that agent allows us to deal with concerns about seemings that are objectionable in ways the agent is unaware of.

We have looked at various reasons people might want to restrict the scope of (PC). Some objections arise from failing to appreciate the extent of possible defeaters, and thus are not incompatible with (PC). Others arise from conflating seemings with non-epistemic motives, and so are not cases of seemings failing to provide justification. Finally, some objections conflate issues of justification, particularly internalist justification, with broader questions of the epistemic status of individuals. If (PC) is a necessary truth, then objections to it must ultimately be based on a misunderstanding of some aspect of what the claim says. I have examined the most common types of objections to (PC) and developed general strategies of reply depending on where the misunderstanding resides. The general point appears to be this: If a seeming produces a belief, there are two sorts of problems that can arise; problems that are among one's currently

[74] Reid, 1969, 318

available evidence and ones that are not. Problems that are available to the subject will provide defeaters, just as (PC) claims. Ones that are not available may effect our evaluation of the subject with respect to various "epistemic disiderata[75]," as Alston calls them, but they will not be relevant to justification as it is understood here. Given the argument in section III, some version of (PC) (restricted or otherwise) must be true if we have justification for any of our beliefs. However, if all that is required to defend (PC) against proposed restrictions is to clarify which states count as seemings, which problems count as defeaters, and exactly how justification is being understood, then (PC) would appear to be a necessary truth, since only a misunderstanding of it would lead to endorsing those restrictions. Unless an objection arises that fails to suffer from such a misunderstanding, (PC) seems to rest on a solid foundation, and the support (8) offers for (2) seems to stand. If these strategies work, then there would be no reason from within Huemer's epistemic framework to cast doubt on (PC). Some people, however, have raised concerns for the framework itself. We will turn to some of those objections now.

2.4.2 DePoe

John DePoe has recently argued that there are ways to justifiably form beliefs, even on internalist grounds, that don't depend on how things seem. Specifically, he claims that we can form justified beliefs on the basis of direct acquaintance even in the absence of seemings, and that, in fact, we should think that direct acquaintance, and not seemings, serve as the rational foundation of beliefs. To motivate this, he claims that

> it is not impossible to imagine intelligent creatures that due to odd mutations in their evolutionary ancestry (or some other reason) have either: (i) no seemings (in Huemer's sense) but still have direct acquaintance with some of their pains; or (ii)

[75] For a comprehensive discussion of the various types of epistemic evaluation we may be interested in, see Alston's (2005).

> who have seemings and direct acquaintance with some of their pains, but where the two have no nomological correlation with one another. Huemer would have us believe that creatures in circumstances like (i) are not justified in having any beliefs (since they have no appearances or seemings). But this is strongly counterintuitive. Even without any seeming that it is in pain, a creature in (i) who is directly acquainted its pain, the belief it's in pain, and the correspondence that holds between its belief and pain, would have a solid internalist basis for forming that belief.[76]

The details of this suggestion aren't very clear as it is described. Specifically, it's unclear exactly what an instance of direct acquaintance without a seeming would be like.[77] I think that when we come up with our best cases of examples of such states, however, the epistemic claims are exactly as Huemer would have them.

The best example I can think of where someone is directly acquainted with something but fails to have it seem to them that they are acquainted with it come from the Perkey experiments involving imagination and perception. In these experiments, people were asked to vividly imagine things, such as a banana, in a certain spot on a screen. An image of a banana was then gradually projected onto that spot on the screen, and the participants believed that the picture they were actually seeing was still the image they were vividly imagining. In these experiments, it seems that the subjects were directly acquainted with a picture of a banana, even though it did not seem to them that there was a picture of a banana in front of them. In this case, however, it seems quite clear that their actual acquaintance with the picture of a banana gave them no reason

[76] DePoe, 2009

[77] In fact, when it comes to pain, I'm not sure the view is even coherent, since seeming to be in pain may well be partially constitutive of the phenomenology of experiencing pain, but I won't press this point here.

whatsoever to believe that there was a picture of a banana in front of them. And the reason for this is quite straightforward: because it occurred without the accompaniment of a seeming.

The "Perkey effect," as it is called, shows it is possible for us to be like the aliens DePoe has in mind. But in such a scenario, we don't have justification, even though we do have direct acquaintance. A creature who had a state of direct acquaintance that was phenomenologically no different than a vivid imagining of the object they are acquainted with would not have good reason to believe that the thing they are acquainted with exists. Seemings are central to the way in which apparent objects of awareness are presented to the mind, and therefore can provide us with an indication of what to do with the content of our appearance states. Without the instructions provided by seemings, even genuine cases of direct acquaintance are in no more position to provide us with reasons to believe than vivid cases of imagination are. This shows that, contrary to DePoe's assertion, it is the seeming that is the relevant basis of justified beliefs, and not the state of direct acquaintance.

2.4.3- Against the Necessity of (PC)

Another potential objection concerns an apparent conflict between the necessity of (PC) and a principle many epistemologists would accept, which Michael Bergmann calls "contingency." Contingency is "the view that the fittingness or unfittingness of an automatic unlearned belief response B to experience E is a contingent feature of that belief response to that experience."[78] An apparent advocate of contingency is Thomas Reid, who claims

> We know, that when certain impressions are made upon our organs, nerves, and brain, certain corresponding sensations are felt, and certain objects are both conceived and believed to exist. But in this train of operations Nature works in

[78] Bergmann, ms

> the dark. We can neither discover the cause of any one of them, nor any necessary connection of one with another.[79]

Bergmann's own example is as follows:

> Although it's epistemically appropriate for me to believe, based on the tactile sensations I'm experiencing while touching an object, that the object is hard, this same noninferential belief might be an inappropriate response to that same sensory experience for an alien cognizer.[80]

Bergmann, Reid, and others endorse the claim that no particular sensations require the adoption of any particular belief, or indeed the formation of any particular concepts in response to them. Bergman thinks this is incompatible with treating (PC) as a necessary truth; after all, the alien cognizer and the human one should adopt different responses to the same tactile sensations. Given the plausibility of these examples, this would pose a serious problem for (PC).

While I believe Reid and Bergmann are on to something important here, I think the insight needs to be clarified. In his initial discussion of this issue, Bergmann assumed that our main evidence for a belief that we were holding a billiard ball was our having "tactile sensations of the type you experience when you grab a billiard ball."[81] And when defending contingency, advocates typically focus on sensations that occur during experience. Bergmann's own example given above, the ones he uses in describing Reid,[82] and the ones he quotes from Reid[83] all focus on the disconnect between particular sensations and subsequent beliefs. However, there is more to an experience than sensation. Also, as we saw with the example from the Perkey effect where the sensations were present but justification was not, it is implausible to think that having such

[79] Reid, 2002, 227
[80] Bergman, ms
[81] Bergman, 2006,, 115
[82] Bergmann, 2006, 119
[83] ibid

sensations is actually the reason we have for adopting the relevant beliefs. So, if the necessity of the connection between our experiences and the appropriateness or inappropriateness of the subsequent beliefs is not due to the sensations present in the experience but due to some other feature of the experience, then the primary motivation typically given for endorsing contingency could be retained without abandoning (PC).[84] Fortunately, given the account of seemings offered here, we can see that this is precisely what (PC) asserts.

On the present account, seemings are not identified by their phenomenology, but by the role they play in allowing us to fulfill our epistemic aims. Indeed, even the phenomenology of human seemings is not heterogeneous. Seemings of different sorts of experiences differ from one another, and even within a category, seemings vary in their phenomenology. Sometimes, when something seems true to the intellect, it feels as if we are drawn or tugged towards it. Sometimes, it seems illuminated or lit by an internal light. Sometimes, it produces what could be called an "Ah-ha!" sensation where the truth suddenly becomes clear. In addition to sensations, experiences contain modes of presentation, or ways the content appears during the experience. What unifies these instances isn't a particular sensation, but rather having a mode of presentation whose function is to allow us to fulfill our epistemic aims with respect to that content[85]. If (PC) treated the sensations that accompany experience as the source of evidence for our beliefs, or if it assumed there was some magical connection between a particular type of sensation and the presence of justification, it would implausibly contradict the motivation usually given for

[84] This reply may not be sufficient for those who advocate additional positions Bergmann endorses, including what he calls "objectivity," nor for those who may be motivated to endorse contingency for other reasons. Bergmann's case of the brain-damaged human (2006, 120; ms, section II), for example, may not be subject to this reply. However, that example requires not just contingency, but the view that some doxastic responses are objectively fitting or unfitting independently of one's available evidence. Since such a principle seems to be a straightforward denial of (CII), I have set it aside as well as related concerns, since I see no reason for an advocate of (PC) to endorse such a principle. For Bergmann's treatment of objectivity, see his 2006, esp. pp.114-16.

[85] This appeal to modes of presentation seems to be what Huemer was getting at when he referred to the differences between types of speech in describing the nature of apprehensions in the quote in Section III.

contingency. However, construing seemings in a way that connects them to facts about how content is presented to the mind, rather than to facts about phenomenology, allows one to endorse versions of contingency motivated by the most common examples without abandoning (PC).

2.4.4- Alston and Voluntarism

One concern that many people may have about the view articulated in this paper is that it seems committed to a voluntarist view of belief formation, and voluntarism is far from universally accepted. Although I think voluntarism is true, the account here only seems to require a view according to which we can be responsible for our beliefs, and therefore compatibilist views would also suffice. There is a wide variety of opinion on whether or not we have sufficient control over our beliefs to ground such responsibility. Here are two drastically different takes on the degree of control we have over our beliefs:

> "Can you, at this moment, start to believe that the Roman Empire is still in control of Western Europe, just by deciding to do so? If you find it incredible that you should be sufficiently motivated to even try to believe this, suppose that someone offers you $500 million to believe it, and that you are much more interested in money than in believing the truth... Can you switch propositional attitudes toward that proposition just by deciding to do so? It seems clear to me that I have no such power." –William Alston, *Beyond Justification*
>
> "If there's a steady paycheck in it, I'll believe anything you want."—Winston Zedman, Ghostbusters

We see here a disparity of opinion that mirrors a divergence in the views of philosophers and ordinary people. Most philosophers seem to defend versions of involuntarism, but people

regularly talk about belief formation as something we have control over. Judgment is typically viewed, in common speech, as an act that follows after an examination of the evidence, and involves a contribution of the individual's will. What is a judge, after all, but someone who is thought to be good at *deciding* what to believe after evaluating the evidence, and is therefore given the job of doing so? Holding judgment to be an involuntary matter surely strips such a profession of much of its merit, and runs counter to how we understand what they, and what we do when reaching our decisions about what to believe.

Alston grants that if the term "justification" in epistemology has an affiliation with common speech, it grows out of a sense of judgment as something we have control over. Alston claims that the use of the term 'justification' in epistemology likely came about as an extension of the term's application to voluntary action. He says that due to this "it is natural to think of believing, when taken to be subject to being justified or unjustified, as subject to requirement, prohibition, and permission", and, further, that "it seems clear that the terms of the deontological triad, *permitted*, *required*, and *forbidden*, apply to something only if it is under effective voluntary control."[86]

Despite this affiliation with common sense, Alston does not believe he actually has control over his beliefs, and thinks he can't be peculiar in lacking this power. At least for some of our beliefs, there is some appeal to this claim, and many people are drawn to it. Thomas Reid, for example, seems to have such a view about our perceptual beliefs, saying "Nature has doomed us to believe the testimony of our senses, whether we can give a good reason for doing so or not."[87] But for most of our beliefs, this claim of Alston's about his abilities strikes me, and others, as odd. I share Sharon Ryan's sentiment that "it seems clear to *me* that I *do* have this

[86] ALSTON, 2005, P. 60
[87] REID, 1969, P.462

power and I exercise it quite a bit. I am sure that I do not have special powers."[88] In a case of such radical difference of opinion, there is typically some confusion underlying at least one party's grasp of the issue at hand. A reasonable degree of conservatism would suggest that we should first look for confusion among those who are rejecting common sense. Fortunately, when we look for causes of confusion here, we are able to find them.

A common cause of confusion about the evidence for and against voluntarism stems from a failure to apply lessons we have learned in discussions of free will to the issue of belief formation. Since, as Mark Heller claims, "in the end, what we say about voluntarism depends upon what we say about free will in general,"[89] we should try to keep our views on these matters in line. For my own part, I find that very little can be said in favor of involuntarism as a theory of belief formation, just as little can be said for determinism as a theory of human behavior. The reason, in both cases, is that it runs contrary to obvious empirical facts. Thomas Reid puts the problems with these views as follows:

> Contrary motives may very properly be compared to advocates pleading the opposite sides of a cause at the bar. It would be very weak reasoning to say, that such an advocate is the most powerful pleader, because sentence was given on his side. The sentence is in the power of the judge, not the advocate. It is equally weak reasoning, in proof of necessity, to say, such a motive prevailed, therefore it is the strongest; since the defenders of liberty maintain that the determination was made by the man, and not by the motive.[90]

Evidence certainly has motivational force with respect to what we believe, but it isn't the only relevant consideration. People have the ability to hear and to become aware of evidence for

[88] Ryan, 2003, p. 65
[89] Heller, 2000, p. 132
[90] Reid, 1969, p. 287

views and *then decide* how to take that evidence into consideration in forming their beliefs. Such a view isn't speculative or surprising, it's simply a fact about people. We have the ability to ignore the force of evidence, to stubbornly hold on to views, to refuse to listen to reason, to weigh evidence differently than others, to consider whether or not it provides us with enough reason to believe, to choose to believe it anyway even if it doesn't. People do these things all the time. The entire language of belief formation is rife with the vocabulary of freedom, choice, and responsibility. It speaks of willfulness, irrationality, and sometimes even of epistemic triumph where we act rationally despite our stronger desires to believe what makes us happy. Those who generally accept the freedom of the will should recognize all the hallmarks of freedom in the area of belief formation, and should be inclined to a similar view here.

Sharon Ryan has done an excellent job detailing the various ways in which views on belief formation inappropriately fail to parallel our views about action. Seeing the comparisons should help us see belief formation in a more complex way, one that will help to diminish the appeal involuntarism draws from an oversimplified picture of voluntary control. Here are four examples from Ryan:

> (A) Suppose I am driving down the road and I see a mother and her child walking down the road. I see them and ask myself, "should I run them over?" I immediately and freely decide not to and then freely drive on. In fact, in a very important sense, I just can't do something like that. You could give me a million dollars and I could not do it. You could give me one hundred million dollars and I could not do it. That's just not something I could do. Does this show that I did not freely decide not to run over the mother and child? Of course not.[91]

---

[91] These examples are from Ryan, (2003), pgs. 63, 65, 66, and 67, respectively.

(B) It's true that I can't just decide, independently of how the evidence looks, to believe or disbelieve or suspend judgment on a proposition. My doxastic attitudes are guided by what seems to me to be good evidence (or whatever it is that motivates me). However, if I make a decision to switch doxastic attitudes, I can adjust my doxastic attitudes accordingly.

(C) Many of our beliefs come upon us quite quickly. Some perceptual beliefs are virtually spontaneous. The fact that a belief or action is virtually spontaneous is no reason at all to regard it as not free, involuntary, or out of our control. I spontaneously, intentionally, and voluntarily turn off my alarm clock when it rings every morning. I spontaneously, intentionally, and freely say "hello" when I see my friend walking into the coffee shop.

(D) Although I might not intend to move my right leg and then my left and then my right, each of these movements is done intentionally. As long as I'm the one doing the moving of my limbs and they are moving how I want them to move, I'm running freely. My running is under my control. As long as I believe what I meant to believe, I believe freely. One need not explicitly intend to do A to do A intentionally.

Ryan directs these considerations to anyone with compatibilist intuitions. I have no compatibilist intuitions, but the view that voluntary control extends far beyond cases where the influence of motives is weak, or where actions result from lengthy, reflective deliberation, is simply a fact about how we conceive of action. We take all sorts of actions, including unreflective ones, to be free. We take the automatic responses that reflect our higher-order intentions to be free. We take easy decisions to be free. We also don't believe that self-imposed limits on what we can do, grounded in the sort of person we are or want to be, limit our freedom or our responsibility.

Therefore, the failure to extend our sense of the scope of voluntary behavior to these parallel instances of belief formation seems to be an unjustified double standard.

I think the above considerations give us good reason to think that voluntary control extends much farther than Alston supposes. Despite this, there is still a problem for the view I'm defending here. On the account of seemings we have examined, they are a type of motive to adopt beliefs; one whose function is characterized by the aim of getting us to believe was it true and to avoid error. The justification for a belief lies in the presence of such a state. Alston's concern, in this context, is that the role of such motives in producing these states is an instance of compulsion or reflex, not an instance of voluntary behavior, and therefore talk of justification or responsibility is misguided. Even if we have voluntary control over some of our beliefs, even if it's over more of them than Alston claims, it seems problematic to grant that a broad area of our beliefs lack justification. However, the notion that we have control over our perceptual beliefs and many of the beliefs we form as children is far from obvious, even given the aforementioned considerations. As Alston puts it,

> Even if beliefs can be formed at will in these kinds of cases, there still remain vast stretches of our belief... where it seems obvious what is the case, where we have already seen believing at will not to be a possible move. And so it would still be true that believing permissibly would not be generally viable as an epistemic desideratum.[92]

Even supposing I am right, then, in thinking that the presumption of voluntary control over our beliefs should extend to a much larger set than Alston supposes, it may well still appear that far too many of our beliefs, including most of our perceptual beliefs, are not adequately accounted for on this picture of justification. One of the main focuses of epistemological endeavors has

[92] Alston, 2005, p. 64

been seeking to salvage the justification of our perceptual judgments from the talons of skeptical arguments. An epistemological theory, if it does nothing else, should at least assure us that seeing the table in front of gives us justification for believing it is there. But if these beliefs are the result of an involuntary response, and justification is deontological in nature, then justification is not adequate for the role it is supposed to play in epistemological debates. Although this concern sounds quite substantive on its face, I think there is a sound reply to it.

As we saw when responding to Markie, there is a difference between having justification for a belief and being justified in holding a particular belief. To be justified requires something more than merely having justification. To see this, consider a parallel in ethics. Take, for example, a slight modification of an example from Frankfurt.[93] Suppose that Jones is on his way to kill someone, and, instead of being threatened as in Frankfurt's case, he notices that his intended victim has his hand on the plunger about to blow up an orphanage. This would give Jones justification for what he is about to do, since it would save innocent lives. Suppose Jones still kills him because he wants to, rather than to save the orphans. Does his having justification for the murder mean he isn't responsible? It seems clear that it doesn't. The reason for this is that the responsibility tracks back to its cause, and the cause here was the initial motive, not the potential motive that would have given him justification. In the case of beliefs, the question of whether we are justified similarly goes back to the point of origin, since it is that action we typically take ourselves to be responsible for.

In cases where the belief was formed for a bad reason, the belief is unjustified. This fact wouldn't change if we were to acquire justification, any more than an immoral act would become moral merely because there is some justification for doing it. To obtain a justified belief in the presence of new evidence, the belief would have to be reevaluated on these current grounds. But

[93] see Frankfurt, 1969

what about compulsive beliefs? Here, the case is similar, but less problematic. It may very well be the case that compulsively formed beliefs, such as most beliefs had by young children and perhaps many perceptual beliefs, are formed in a way that prevents them from being justified, since we have no say in their formation. This, however, does not open them up to skeptical concerns or place undue limits on a theory of justification. Alston's fears can be allayed, because in spite of the fact that they aren't justified, we still have justification for such beliefs, and the beliefs weren't formed in a blameworthy manner. Such beliefs aren't irrational, epistemically valueless dictates of a "blind and powerful instinct of nature" like Hume claimed our natural perceptual judgments to be.[94] Instead, they are beliefs for which we have the highest degree of justification, even though they bypassed the possibility of being justified (at least until reflection raises the issue of our acceptance of them) due to the compulsive manner in which they were formed. We can see, then, that a deontological account of justification can be maintained even if we accept certain limits on our control. Some beliefs will be better described as blameless than justified on such an account, but the account needn't thereby be incomplete or subject to skeptical worries. We can therefore make use of the apparently commonsensical view of belief formation that takes it to be something we are responsible for, and to which questions of permissibility and rationality apply.

Concluding Remarks

I have defended Michael Huemer's principle of phenomenal conservatism by providing a more detailed account of the nature of seemings. This account allowed me to deal with some objections to Huemer's position and to defend a strong version of Huemer's self-defeat argument for (PC). Huemer claimed in his earlier work that

[94] Hume, 2002

> I think the principle of phenomenal conservatism underlies judgment in general. I think reflection will reveal that all judgment, whether inferential or not, is a process in which one accepts a proposition on the basis of how things seem to oneself. If phenomenal conservatism is false, so that the way things seem to oneself is irrelevant to epistemic justification, then all judgment must be irrational.[95]

I take myself to have engaged in the reflection that vindicates this claim. Huemer was right to suspect that the denial of (PC) entailed irrationality, and that seemings were therefore a precondition on the possibility of epistemic justification. To deny (PC) is to deny that we are epistemically rational. And this, as Huemer claims, "is something which no philosopher, not even a skeptic, can accept."[96]

This chapter has outlined an account of the nature of evidence. To have evidence for a belief, various propositions must appear to be true to us when we entertain them. Having developed a more complete account of the nature of evidence, we should be in a better position to evaluate the strengths and weaknesses of competing philosophical methods. In the next chapter, I will develop a full account of the common sense approach to philosophy and show how it is both motivated by this account of the nature of evidence and does a great job of addressing the concerns raised in chapter 1. These defenses of the approach should set it on a good foundation and thereby legitimize positions that are motivated by it. The rest of the work will then apply this approach to questions in the philosophy of mind.

[95] Huemer, 2001, p. 107
[96] Huemer, 2001, 107-8

Chapter 3: Defending the Common Sense Approach

Overview:

So far, we have looked at two large issues affecting philosophical method: a practical issue showing that methods are often subject to perverse incentives, and an account of the nature of justification and of evidence that should help shape the approach we take to philosophy. Here, I will argue that both of these issues strongly support endorsing common sense philosophy. Before we can see how this defense works, however, we need to understand the elements of that approach. We will start with an account of the nature of common sense, and then look at the different elements of the common sense method. Having done so, I will explain how the work of the previous two chapters motivates accepting this view.

3.1: What is Common Sense?

To begin, we should get clear on the nature of common sense. According to Thomas Reid, the biggest obstacle to understanding the nature of common sense is that philosophers have come to understand the word 'sense' differently than it is used in common speech. Philosophers now understand 'sense' as "the means of furnishing our mind with ideas, without any kind of judgment."[97] This is a departure from common usage since "in common language, sense always implies judgment. Good sense is good judgment. Nonsense is what is evidently contrary to right judgment. Common sense is that degree of judgment which is common to men with whom we can converse and transact business."[98] Common sense consists in the understanding of the world that is shared in common by all of us in our daily interactions. We can generally assume, in our conversations, that others can see what we do, that they have a mind, and can act upon their desires, that they will be responsive to basic principles of reasoning, remember what has just

[97] Reid, 2002, p.424
[98] Reid, 2002, p.424

transpired, and other things of this nature. These assumptions are grounded in a belief that some facts are easily and equally available to us all, are far too obvious to need to mention in everyday life, and are properly granted and assumed without need of permission in all ordinary discourse. Common sense philosophy takes these presuppositions of everyday life as the cornerstones of philosophical progress, and seeks to build philosophy upon them.

Common sense is comprised of the judgments that are available to us in the ordinary exercise of our various intellectual powers, especially those that are presupposed in our daily interactions with one another. Common sense philosophers typically accept an array of sources of foundational judgment. Reid, for example, considers all of our ordinary intellectual faculties as sources of rationally justified beliefs,[99] and other advocates of common sense, such as G. E. Moore with his acceptance of ethical intuitions,[100] demonstrate a greater than common willingness to countenance and embrace a large number of legitimate sources of knowledge. In addition, common sense philosophy takes the initial judgments obtained from these sources as rationally justified. This permits us to weigh such judgments against the results of theoretical efforts such as philosophy and scientific theorizing as independent pieces of evidence. So, common sense philosophy includes a broad-based foundationalism that treats our means of knowing as generally reliable, and the immediate deliverances of each source of knowledge as sufficiently well supported to bypass the need for independent justification. Although a broad-based foundationalism is part of the common sense method, not all foundational beliefs are elements of common sense. Among our foundational beliefs, some are so ingrained in human thought and action that they form a natural context for discourse and for understanding one another's intentions and conduct. These are the tenets of common sense. Although a helpful

[99] Reid includes at least perception, memory, introspection, and rational intuition as intellectual powers or faculties of man.
[100] For Moore's defense of ethical intuitions, see his (1903)

guideline, this account would benefit from more detail to help identify elements of common sense.

3.2 Identifying Common Sense Beliefs

While common judgment is an indication of common sense, not all widespread judgments qualify as tenets of common sense. Although he thinks it is impossible to draw a sharp line between them, Reid offers an account of the characteristic features that separate widespread judgments that are parts of common sense from those that are not. There are at least three ways of identifying these beliefs. First, one can discover elements of common sense by looking at the structure of languages, and identifying shared elements of how we talk about the world. Second, we can discover them by attending to apparent presuppositions of everyday life. Third, we can discover further elements of common sense by attending to those judgments of mankind that express self-evident propositions built in to how we conceive of the world, or that are sufficiently apparent that we needn't mention them in everyday life. We will look at each of these methods in turn.

According to Reid "there are...opinions that appear to be universal, from what is common in the structure of all languages, ancient and modern, polished and barbarous. Language is the express image and picture of human thoughts; and, from the picture, we may often draw very certain conclusions with regard to the original."[101] By drawing inferences from the structure of language to the beliefs that correspond to these structures, we can discover certain tenets of common sense. This is because

> There are certain common opinions of mankind, upon which the structure and grammar of all languages are founded. While these opinions are common to all men, there will be a great similarity in all languages that are to be found on the

[101] Reid, 2002, p.45

> face of the earth. Such a similarity there really is; for we find in all languages the same parts of speech, the distinction of nouns and verbs, the distinction of nouns into adjective in substantive, of verbs into active and passive… There are general rules of grammar, the same in all languages. This similarity of structure in all languages shews an uniformity among men in those opinions upon which the structure of language is founded.[102]

Reid suggests that close attention to the parts of speech and to the rules of syntax present in all languages will often reveal principles of common sense. Attending to the structure of language as a means of identifying common sense beliefs allows us to realize that common sense includes many tacit beliefs. Judgments that are commonly revealed in the structure of language are rarely evident to those who make use of them, and we rarely spend time attending to the truths that common discourse rests upon.

Another means of ascertaining principles of common sense that are often tacit is a close examination of the way that people conduct themselves in the ordinary affairs of life. "The actions of men are effects: Their sentiments, their passions, and their affections, are the causes of those effects; and we may, in many cases, form a judgment of the cause from the effect."[103] For example, we don't often attend to our beliefs in the general reliability of testimony, the existence of minds in those who provide testimony, the legitimacy of memory, or the possibility of movement, but when someone tells us our keys are on the coffee table we understand and trust her, remember the location of the table, and move to that location expecting to find the keys. A whole host of beliefs seem to be involved in this relatively simple process, and are borne out by

[102] Reid, 2002, p. 36. Reid proceeds to give the examples of our beliefs that attributes cannot exist without a subject, and that there is a difference between acting and being acted upon, as the sorts of things we can discover by attending to language.
[103] Reid, 2002, p.57

our actions, but it is very doubtful most people ever have cause to attend to them. Discovering such presuppositions of human action is a generally reliable means of ascertaining principles of common sense; for "when an opinion is so necessary in the conduct of life, that without the belief of it, a man must be led into a thousand absurdities in practice, such an opinion, when we can give no other reason for it, may safely be taken for a first principle."[104]

Although a number of principles of common sense may be tacit, we don't want to limit common sense to tacit judgments. Ascertaining the common sense of professed principles, however, is not significantly different from ascertaining those that are tacit since "the opinions of men may be considered as the effects of their intellectual powers, as their actions are the effects of their active principles."[105] We can therefore discover principles of common sense by examining the professed judgments of people. However, while tacit universal judgments presupposed in understanding common discourse, and in participating in the ordinary affairs of life are, possibly without exception, reasonable candidates for principles of common sense, the professed judgments of mankind are far less universal, far less clear, and far less apt for hitting the truth than these tacit judgments. Therefore, we need some means of separating those professed judgments of mankind that should be considered elements of common sense from those that are not.

Reid believes it is proper to count both tacit and express judgments of people as elements of common sense because "the same degree of understanding which makes a man capable of acting with common prudence in the conduct of life, makes him capable of discovering what is true and what is false in matters that are self-evident, and which he distinctly apprehends."[106] As we see from this quote, express judgments that qualify as common sense must be self-evident. In

[104] Reid, 2002, p.457
[105] Reid, 2002, p.57
[106] Reid, 2002, p. 426

addition to what nearly everyone would generally acknowledge to be self-evident, Reid believes that there are many things that would be self-evident were they to be considered. Thus he says

> It may be observed, that there are truths, both speculative and moral, which a man left to himself would never discover; yet, when they are fairly laid before him, he owns and adopts them, not barely upon the authority of his teacher, but upon their own intrinsic evidence, and perhaps wonders that he could be so blind as not to see them before.[107]

Not all such judgments are clear candidates for common sense, since they aren't always readily apparent, and needn't be part of our shared presuppositions of ordinary life. However, they do help bridge the gap between tacit and express judgments of common sense. Tacit elements of common sense are typically those most likely, when properly understood, to be among the claims that strike up as obvious when first "fairly laid before" us. In either case, because common sense beliefs are a subset of foundational beliefs, some beliefs clearly do not qualify. In particular, beliefs based upon conjecture or that deal with matters beyond what can be easily furnished by natural human faculties should not be considered elements of common sense. However, in cases "where there is such universal consent in things not deep nor intricate, but which lie, as it were, on the surface, there is the greatest presumption that can be, that it is the natural result of the human faculties; and it must have great authority with every sober mind that loves truth."[108] Common sense beliefs, then, include surface-level facts available through the routine use of our various intellectual powers as well as underlying presuppositions that shape our experience of the world and our means of communication, thereby creating a shared context for understanding and interacting with one another and with the world around us.

[107] Reid, 1969, p.249
[108] Reid, 2002, p. 45

3.3: The Common Sense Method:

Now that we have a better sense of which beliefs count as common sense, we need to look more closely at the elements of a philosophical method that gives those beliefs a central role. There are four elements of the common sense approach we will look at here. First, it gives our ordinary standards of justification authority in philosophical matters. Common sense philosophy denies the existence of special philosophical standards of justification, instead believing ordinary standards of justification and proof capture what is required for knowledge. Second, common sense philosophy takes a skeptical attitude toward speculation and correspondingly toward any arbitrary presuppositions in ontology or epistemology about the limits of human knowledge. This fact heavily influences the relationship between common sense and science in ways we shall examine. Third, common sense philosophy shares a strong affiliation with some features of ordinary language philosophy, including a firm insistence on using terms with their ordinary meanings and accepting the ontological commitments built into ordinary speech. Finally, common sense philosophy takes a stance on the comparative expertise of philosophers and non-philosophers on various matters. In particular, it holds that the judgment needed to make surface level observations is shared in common, and so the efforts of philosophers to overturn common sense requires rejecting the consensus of experts on matters where philosophers have no more ability to see the truth than those whose judgment they seek to overturn. We will look at these elements in the following sections.

3.3.1 Common Sense Philosophy and the Rejection of Shifting Standards

In our every day affairs, we take common sense judgments to be authoritative. To demonstrate this, Reid considers what would happen if the standards that philosophers had commonly accepted during his time were made use of in a court of law.

> ...if a sceptical counsel should plead against the testimony of the witnesses, that they had no other evidence for what they declared, but the testimony of their eyes and ears, and that we ought not to put so much faith in our senses, as to deprive men of their life or fortune upon their testimony; surely no upright judge would admit a plea of this kind...Can any stronger proof be given, that it is the universal judgment of mankind that the evidence of sense is a kind of evidence which we may securely rest upon in the most momentous concerns of mankind: That it is a kind of evidence against which we ought not to admit any reasoning; and therefore, that to reason either for or against it is an insult to common sense?[109]

The evidence of the senses is given immunity from philosophical reasoning in matters where the legitimacy of one's beliefs is of the highest importance. The reason for this is that the testimony of the senses is nothing more than the knowledge that any of us could obtain if we were in the same circumstance, and therefore needs no support from our other faculties, or from other people in order to be given strong weight as evidence. Nor is this limited to perception. Discussing our ethical intuitions, Reid says

> A man who is come to years of understanding, and who has exercised his faculties in judging of right and wrong, sees their truth as he sees day-light. Metaphysical arguments brought against them have the same effect as when brought against the evidence of sense: they may puzzle and confound, but they do not convince.[110]

Some matters are so clear that the judgment needed to see their truth is available to all, and efforts to overcome such claims by other means is viewed as a misguided effort on a common sense approach.

---

[109] Reid, 2002, p. 98
[110] Reid, 1969, p. 390

G. E. Moore raises a similar point about the nature of proof in his "Proof of an External World". Moore writes

> Suppose, for instance, it were a question whether there were as many as three misprints on a certain page in a certain book. A says there are, B is inclined to doubt it. How could A prove that he is right? Surely he *could* prove it by taking the book, turning to the page, and pointing to three separate places on it, saying "there's one misprint here, another here, and another here": surely that is a method by which it *might* be proved![111]

Here, and with his famous "hands" argument, Moore is reminding us that there are standards of proof that permit us to settle matters in a variety of ways, including demonstrating that something exists by showing it to people. Typically, philosophers start their inquiries with the idea that reasoning has authority in evaluating evidence. However, this does not seem to be how we think of evidence when we care about the truth in everyday life. Life is full of times when we want to know the truth, and standards of evidence and justification already pervade everyday life in matters where the truth is at issue. Shifting away from these standards requires a compelling reason to believe that they are somehow improper, or misguided. In the absence of a powerful defense of the idea that philosophers have access to the "real" standards of evidence or justification, this departure from ordinary practice seems ill-grounded.

This appeal to standards used in a courtroom, or to the standards of justification we appeal to when we care most about the truth, may seem at first to be a bare appeal to ordinary practices, but there is a deeper point here that connects to contextualist debates over how to treat skeptical concerns, and how to evaluate things in philosophical discussions. When Reid wrote, he was concerned with the Early Modern elevation of reasoning over perception that he thought

[111] Moore, 1939, p.

had led to so much error. As Reid describes the tradition of his day “”philosophers, pitying the credulity of the vulgar, resolve to have no faith but what is founded upon reason. They apply to philosophy to furnish them with reasons for the belief of those things which all mankind have believed, without being able to give any reason for it.”[112] The perceived need to justify all other beliefs by reason was so pervasive that any intellectual capacity not strictly resembling deductive reasoning was not thought to be, properly speaking, an intellectual power at all, but rather a source of vulgar prejudices that needed correction through philosophical argumentation. Reid believes that this is why “sense” became detached from “judgment” in early modern philosophy, and why perception and other forms of determining what the world was like were no longer labeled “intellectual” powers. What Reid intends to point out with the courtroom example is not merely that we don’t ordinarily accept such standards, but that

> (1) Our intellectual powers include all sources of initial judgment such as perception, memory, and rational intuition, and
>
> (2) The philosophical tradition had no grounds for limiting authority in epistemic matters to the sources of belief acquisition it prefers.

To use a phrase of William Alston’s, Early Modern Philosophy was guilty of a gross instance of *epistemic imperialism*, which Alston defines as “unwarrantedly taking the standards of one doxastic practice as normative for all.”[113] Knowledge, according to Reid, is not the province of any one of our intellectual powers; it is the province, in their respective areas of expertise, of each of them. When it is thought that philosophy, in following the methods of some preferred means of learning about the world, has found the true standards for justification and knowledge,

[112] Reid, 1997. p. 18
[113] Alston, 1991, p.

there seems to be little reason to accept such claims, especially when they fly in the face of how we evaluate evidence in matters of paramount importance in everyday life.

Reid describes the problems facing the philosophy of his day as follows:

> …the defects and blemishes in the received philosophy…which have most exposed it to the contempt and ridicule of sensible men, have chiefly been owing to this: That the votaries of this Philosophy, from a natural prejudice in her favor, have endeavoured to extend its jurisdiction beyond its just limits, and to call to her bar the dictates of common sense.[114]

At the heart of this complaint is an anti-contextualist view of justification, at least with regard to those standards that are relevant to philosophical discussions. Reid takes the fact that we all use a different set of standards, even in cases where the truth is of vital importance, to show that the elevation of reasoning involves a departure from the *genuine* standards of justification and knowledge: those we use in everyday life. The demand for philosophical proof in matters of common sense doesn't involve a move to a higher, or more authentic set of standards for justification; it involves a departure from considerations of justification in such matters toward questions of their *susceptibility* to philosophical proof. What remains is not justification, but mere accordance with a set of philosophical rules or practices. As interesting as they may be to follow, such rules do not determine our ability to know the things in question. This explains why perceptual judgments and other tenets of common sense are not correctable by philosophical argument. Standards of justification don't permit theoretical reasoning to override perceptual judgments, and attempts to elevate theoretical reasoning to a point where it is capable of having such authority removes one from the standards of justification required for knowledge.

3.3.2-Science and Common Sense

[114] Reid, 1997, p. 19

As we have just seen, common sense philosophy rejects the idea that we should try to fit all knowledge into a preferred method that may not be well suited to every area. As Reid puts it, "Even very ingenious men are apt to make a ridiculous figure, by drawing into the track, in which their thoughts have long run, things altogether foreign to it."[115] One may think at first glance that the application of this view to naturalism is straightforward. Naturalists tend to privilege a scientific means of understanding the world over all others in a way that is similar to the way Descartes elevated the methods of Euclidean Geometry as the paradigm of human knowledge. But common sense philosophy is certainly far from anti-scientific. Reid, for example, held science in very high esteem, going so far as to say "Wise men now agree, or ought to agree in this, that there is but one way to the knowledge of nature's works; the way of observation and experiment."[116] Though this quote may seem at odds with what has just been said, we need to place it, and its consequences, in context. Specifically, we must understand Reid's view of proper scientific reasoning, which he thought was exemplified by Newton.

Reid viewed Newton's success in science as grounded in laying down principles of common sense[117] as a foundation for pursing investigations in natural philosophy. He said of Newton's Rules of Philosophizing "though they have not the same kind of evidence that mathematical axioms have; yet have such evidence, that every man of common understanding readily assents to them, and finds it absolutely necessary to conduct his actions and opinions by them, in the ordinary affairs of life."[118] Reid viewed proper science as a discipline that followed general principles of common sense, and refused to speculate beyond what those, in addition to careful observation, offered. The reference to mathematical axioms draws attention to the fact

[115] Reid, 2002, p. 538
[116] Reid, 1997, p. 11
[117] See, e.g. Reid, 1997, p. 12, 2002, p. 40-41
[118] Reid, 2002, p. 40-41

that Reid thought this extended to all fields. He thinks we can use the clear deliverances of rational intuition to discover first principles of mathematics and, as we saw earlier, of ethics. In addition, Reid viewed the scope of "observation" in a broader way than it is frequently used at present. In addition to using common sense principles and empirical observation in studying the natural world, we should use introspection, and our common understanding of the mind in evaluating claims in epistemology and philosophy of mind. Indeed, Reid goes so far as to call introspection, or "reflection", "the only instrument by which we can discern the powers of the mind."[119] In every discipline, however, common sense desires careful attention to observational facts available to us as the cornerstones of the field, and demands a degree of humility and caution in positing additions to what we can learn through such observation. One helpful way to think of common sense philosophy is as combining the rationalist's willingness to embrace a wide array of sources of knowledge with an empiricist's method of only endorsing beliefs that follow more or less directly from the direct or observational data that we obtain from such sources. This point leads to another interesting aspect of the common sense view, the rejection of conjecture.

3.3.3: Common Sense and the Rejection of Conjecture:

Common sense philosophy takes a very critical stance of conjecture as a source of evidence. Reid takes his attitude toward conjecture from Newton, whom he believes to have demonstrated the proper way to proceed in examinations of the underlying nature of reality, and the illegitimacy of conjecture as such a means. According to Reid "we may learn something of the way in which Nature operates, from fact and observation; but if we conclude that it operates in such a manner, only because to our understanding, that appears to be the best and simplest

[119] Reid 1997, p. 15

manner, we shall always go wrong."[120] Nor is immunity from error a matter of intellectual ability according to Reid, who claims "when men attempt, by the force of genius, to discover the causes of the phenomena of Nature, they have only the chance of going wrong more ingeniously."[121] Reid views much of the history of philosophy as a continual replacement of old hypotheses grounded on mere conjecture with new ones that escape only briefly the absurdity that the past conjectures now bear on their face. Aristotle's speculations about what causes objects to fall seem rather silly to us, but they were far more sophisticated than what they replaced, and it was not through philosophical speculation that we came to learn of their illegitimacy, but through empirical observation. Reid takes the history of philosophy to bear out the view that no matter how ingenious the inventor, idle hypotheses about the nature of the world "may have a decent and plausible appearance to those who are not more knowing than the inventor; but when men come to be more enlightened, they will always appear ridiculous and childish."[122]

The error here is similar to that of using reason to justify the evidence of the senses: in both cases a tool of the intellect is misapplied and results in opinions that are deserving of ridicule. In both cases, philosophers have sought to use the tools they are most suited for; their powers of reasoning and of inventing and comprehending complex works of imaginative brilliance, and have demand that all knowledge fall under the purview of these skills. In Reid's view, the history of philosophy has shown, not that philosophy cannot succeed, but that its successes only occur when careful attention is paid to when philosophical reasoning needs to be applied, and is properly applied only in those cases. A similar point should hold for naturalistic assumptions. Assuming that science will one day answer questions that don't appear to be in

[120] Reid, 2002, p. 531
[121] Reid, 2002, p. 534
[122] Reid, 2002, p. 50

their purview is a pre-mature, and ultimately anti-scientific speculation. So is the assumption that reality must conform to a speculative metaphysic such as physicalism. Our overall view of the world should grow out of the data we have, and never presuppose future results that go beyond such data. Common sense philosophy takes the world as we find it, and doesn't demand that it fit any preconceived ideas about it nature.

3.3.4 Commitments of Ordinary Speech

According to Reid, "Common words…ought to be used in their common acceptation."[123] The reason for this is that "There is no greater impediment to the advancement of knowledge than the ambiguity of words."[124] Reid believed that common language, including the distinctions built into its structure, reflected common sense judgments of mankind. For example, he claims

> The distinction between substances, and the qualities belonging to them; between thought, and the being that thinks; between thought, and the objects of thought; is to be found in the structure of all languages: And therefore, systems of philosophy, which abolish those distinctions, wage war with the common sense of mankind.[125]

Reid cautions that the abuse of words is often a means of buttressing or legitimizing bad philosophy. There is a strong desire among philosophers to strip words of their apparent meanings. Reid criticizes these efforts because the corruption of language is often necessary to make a person's philosophy sound more palatable than it actually is, and in particular to mask conflicts with common sense.[126]

---

[123] Reid, 2002, p. 18
[124] Reid, 2002, p. 17
[125] Reid, 2002, p. 466
[126] Abuses also sometimes occur in an effort to make an obvious claim sound more interesting than it really is. Reid complains about this as well, since he believes it leads to fruitless debate and unneeded confusion, but since this is less likely to be a problem for an advocate of common sense, I don't plan to discuss it here.

According to Reid, "Philosophers ought not to escape censure when they corrupt a language, by using words in a way which the purity of the language will not admit."[127] As an example, Reid criticizes Hume for his misuse of the word 'impression.' Here is Hume's definition of 'impression':

> By the term *impression*, then, I mean all our more lively perceptions, when we hear, or see, or feel, or love, or hate, or desire, or will. Ideas are the less lively perceptions, of which we are conscious, when we reflect on any of those sensations or movements above mentioned.[128]

Reid criticizes this definition for three reasons. First, Hume's account of which things are perceptions is grossly promiscuous. Throughout his works, Hume "gives the name of perceptions to every operation of the mind. Love is a perception, hatred a perception. Desire is a perception, will is a perception…"[129] Reid laments that "this is an intolerable abuse of language, which no Philosopher has authority to introduce."[130] Second, Hume claims that a distinction in degree is sufficient for a distinction in kind, dividing every operation of the mind into categories on the basis of their vivacity. This runs counter to our natural understanding of proper categorization, since "to say…that two different classes, or species of perceptions, are distinguished by the degrees of their force and vivacity, is to confound a difference in degree with a difference in species, which every man of understanding knows how to distinguish."[131] Finally, Reid criticizes Hume's use of the term for being persistently ambiguous since "when I read all that he has written on this subject, I find this word impression sometimes used to signify

[127] Reid, 2002, p. 32
[128] Hume, 1975, quoted in Reid, 2002, p. 32
[129] Reid, 2002, p. 32
[130] Reid, 2002, p. 32
[131] Reid, 2002, p. 33

an operation of the mind, sometimes the object of the operation; but, for the most part, it is a vague and indetermined word that signifies both."[132]

When it comes to this final point, Hume's ambiguity seems borrowed from Locke and the rest of the tradition in their use of the word "idea", which is often misapplied to the object of the mind's attention, or to an operation of the mind, rather than to a feature of the mind. Hume's division causes new problems, since "common sense convinces every man, that a lively dream is no nearer to reality than a faint one; and that if a man should dream that he had all the wealth of Croesus, it would not put one farthing in his pocket."[133] However, using either term to refer to an operation of the mind or an object of the mind's awareness is wholly at odds with common sense. There is a use of the term 'impression' that allows it to apply to the mind, but "in the most extensive sense, an impression is a change produced in some passive subject by the operation of an external cause."[134] The various things Hume calls impressions, however, cannot fit this description. Uses that imply that operations of the mind are impressions are inconsistent with it because "in the structure of all languages, they are considered as acts or operations of the mind"[135] and not as passive receptions of the mind. The application of the term to the objects of awareness is an even worse offense, because "would any man be thought to speak with propriety, who should say that the sun is an impression, that the earth and the sea are impressions?"[136]

Reid focuses heavily on this example because of the importance a common sense approach places on maintaining clear, straightforward, and ordinary use of language. According to Reid, "if a man would persuade me that the moon which I see, and my seeing it, are not two

[132] Reid, 2002, p. 34
[133] Reid, 2002, p. 34
[134] Reid, 2002, p. 35
[135] Reid, 2002, p. 35
[136] Reid, 2002, p. 35+6

things, but one and the same thing, he will answer his purpose less by arguing the point in plain English, than by confounding the two under one name, such as that of an impression…."[137][138] He goes on to say "The structure of all languages is grounded upon common notions, which Mr Hume's philosophy opposes, and endeavours to overturn. This no doubt led him to warp the common language into a conformity with his principles…"[139] Since abusing language is frequently a means of masking conflicts with common sense, it is important that an advocate of common sense be careful to use words with their common meanings and avoid the trap of pretending there is an accord with common sense where none exists.

In addition to his concerns about the elicit misuse of words when they deviate from ordinary language, Reid believed that ordinary language carried with it ontological commitments. When speaking of a belief in the existence of a subject who has the various experiences we are aware of, he says "It is impossible to trace the origin of this opinion in history: for all languages have it interwoven in their original construction. All nations have always believed it."[140] Reid thinks that ordinary speech commits us to, and includes as a foundation, a belief in the existence of the things we talk about. He criticizes the theory of ideas for resulting in a conclusion that no such subjects of experience exist, lamenting that "We were always apt to imagine, that thought supposed a thinker, and love a lover, and treason a traitor: but this, it seems, was all a mistake; and it is found out, that there may be treason without a traitor, and love without a lover, laws without a legislator, and punishment without a sufferer…"[141] The claim that common sense, or common language, carries with it a set of ontological commitments runs against a current, popular view of analytic metaphysicians. However, it is not surprising

[137] Reid, 2002, p. 34
[138] David Stove has a hilarious and instructive discussion of a similar efforts by Berkeley in his 1991, p. 141-144
[139] Reid, 2002, p. 36
[140] Reid, 1997, p. 36
[141] Reid, 1997, p. 35

that common sense philosophy would reject such a claim. The central idea behind rejecting the commitments of ordinary speech seems to amount to the claim that people don't actually try to talk about the world, or that we don't have a means of conceiving of reality built in to ordinary thought. However, Reid further points out that we are naturally constituted to take all sorts of experiences as indicators of things that they are unlike. He gives the following examples:

> I think it appears…that there are natural suggestions; particularly, that sensation suggests the notion of present existence, and the belief that what we perceive or feel, does now exist; that memory suggest the notion of past existence, and the belief that what we remember did exist in time past; and that our sensations and thought do also suggest the notion of a mind, and the belief of its existence, and of its relation to our thoughts. By a like natural principle it is, that a beginning of existence, or any change in nature, suggests to us the notion of a cause, and compels our belief of its existence.[142]

We have, built into our ordinary means of conceiving of reality, and reflected in the structure of language, a common sense ontology that is presupposed in our understanding of one another, and of our experiences of the world. Commitments of ordinary speech are therefore commitments to an ontological picture of the world, and ones that should not be lightly dismissed.

3.3.5 An Appeal to Authority

We seek out authorities in various fields to provide us with information and to reinforce our conclusions. In general, we seek reassurance from others, particularly experts because "society in judgment, of those who are esteemed fair and competent judges, has effects very similar to those of civil society; it gives strength and courage to every individual; it removes that timidity which is as naturally the companion of solitary judgment, as of a solitary man in the

[142] Reid, 1997, p. 38

state of nature."[143] When it comes to matters of common sense, Reid thinks that we seek out the same reassurance. However, since "in matters of common sense, every man is no less a competent judge than a Mathematician is in a mathematical demonstration,"[144] the appeal to authority in these cases is an appeal to the good sense of those around us. When it comes to common sense beliefs, one does not need to develop any special skills not available to others. According to Reid "the power of judging in self-evident propositions, which are clearly understood may be compared to the power of swallowing our food...It requires ripeness of understanding and freedom from prejudice, but nothing else."[145]

The difference in abilities of reason and judgment between philosophers and ordinary believers is a matter of degree, not of kind. Philosophers have more refined and practiced judgment, and more training at conceptual analysis and argumentation, but they don't have new intellectual faculties. If our judgment in simple matters is so universally poor as to require serious revision, I see no reason to believe that higher degrees of skill will give us much hope in more complex ones. If simple matters can't be known with common judgment, I see no reason to think that complex ones can be with improvements of the same kind of judgment. It seems much more likely that there is a degree of judgment common to us all that can be used to understand much of the world around us, and that advancement of skill can extend the scope, but not overturn the foundations of human knowledge. When it comes to surface level facts, we had better be able to trust our common sense, or there is little hope we can ever go beyond it. Since it requires the rejection of the near universal agreement of those no less capable of ascertaining the truth in the matter than you are, rejecting common sense for theoretical reasons is a methodological mistake that we should insist on avoiding.

[143] Reid, 2002, p. 465
[144] Reid, 2002, p. 465
[145] Reid, 2002, p. 453

3.4: A Theoretical Defense of Common Sense

Over time, the idea that theoretical considerations cannot overturn common sense has been developed by proponents of the method into a general refutation of skeptical arguments. Reid claimed that "opinions which contradict first principles are distinguished from other errors by this: That they are not only false but absurd."[146] This suggests that skeptical arguments, which deny common sense beliefs, ought to be treated as *reductio ad absurdum* arguments. This point was later advanced by G. E. Moore, and further defended by Michael Huemer. Reid offers the beginnings of a theoretical defense of common sense, and of the claim that skepticism involves a reductio, when he says the following:

> But it is possible that, by setting out from false principles or by an error in reasoning, a man may be led to a conclusion that contradicts the decisions of common sense. In this case the conclusion is within the jurisdiction of common sense, though the reasoning on which it was grounded be not; and a man of common sense may fairly reject the conclusion without being able to show the error of the reasoning that led to it.[147]

In "Hume's Theory Examined," G. E. Moore takes up this suggestion, and gives a response to Hume's skepticism about our knowledge of the external world that mirrors it. As Moore sees it, the dialectic goes as follows[148]:

> If Hume's principles are true, then, I have admitted, I do *not* know *now* that this pencil—the material object—exists. If, therefore, I am to prove that I *do* know that this pencil exists, I must prove, somehow, that Hume's principles, one or both

[146] Reid, 2002, p. 368
[147] Reid, 2002, p. 340
[148] To hear a recent depiction of the argument, check out The 21st Century Monads' "The G. E. Moore Shift" at http://people.umass.edu/phil592w-klement/monads/TheGEMooreShift.mp3

> of them, are *not* true. In what sort of way, by what sort of argument, can I prove this?
>
> It seems to me that, in fact, there really is no stronger and better argument than the following. I *do* know that this pencil exists; but I could not know this, if Hume's principles were true; *therefore*, Hume's principles, one or both of them, are false. I think this argument really is as strong and good a one as any that could be used: and I think it really is conclusive.[149]

Michael Huemer has recently endorsed the point that Moore and Reid are making. In the processes of endorsing Moore's argument, Huemer formalizes it into a general argument pattern that ought to suffice for defending any common sense belief against skeptical arguments. Huemer asks us to consider the following three arguments where we "let K be the proposition denied by the skeptic, that we have a certain kind of knowledge."[150]

A	A	K
B	K	B
*~K	*~B	*~A

Huemer reminds us that "each of them is logically sound under exactly the same condition: namely, that A, B, and K are jointly incompatible."[151] In evaluating these arguments "the issue comes down to which of them has the more plausible premises," but "the nature of common sense beliefs, as such, is that they have the highest initial plausibility of all beliefs."[152] So, "it

[149] Moore, 1939, 606
[150] Huemer, 2001, p. 32
[151] Huemer, 2001, p. 32
[152] Huemer, 2001, p. 33

follows that a common sense belief could not be refuted by another, non-common sense belief."[153]

Each of these philosophers claim that positions that come into conflict with common sense are best understood as *reductio ad absurdums* of their opponent's premises. I provided a defense of a similar view in the first chapter of this work as well, when I claimed that not all beliefs are subject to correction by arguments. We should look at what has been said in defense of these arguments to better understand the motivation for endorsing the common sense approach. Reid has individual comments that are suggestive of some of the things we will see from later common sense philosophers, but his primary reason for thinking that we should not use reason to overturn common sense rested in his view about the comparative legitimacy of our various intellectual powers. Discussing perception and reason, Reid says

> Reason, says the sceptic, is the only judge of truth, and you ought to throw off every opinion and every belief that is not grounded on reason. Why, sir, should I believe the faculty of reason more than that of perception?—they came both out of the same shop, and were made by the same artist; and if he puts one piece of false ware into my hands, what should hinder him from putting another?[154]

This view is representative of Reid's view that since our abilities have the same source, arbitrarily privileging of one of those abilities over the others is a mistake. According to Reid, many of our beliefs are simply natural products of our constitution, and we can't be blamed for accepting them, nor should we feel justified in dismissing them.

For Moore, the defense of this approach rested primarily in facts about the legitimacy of arguments. Moore realized that his counter-argument to Hume wasn't likely to convince many

[153] Huemer, 2001, p. 35
[154] Reid, 1997, p. 168-9

people on its first reading, so he tried to explain why it should. First, he noted that if the Humean argument is valid, then so is the Moorean one. And since they share their first premise, we only need to examine the second one to see which argument is better. As Moore claims, what we want from an argument, in addition to validity, is "that we should know the premises to be true...And moreover the degree of certainty of the conclusion, in either case, supposing neither is quite certain, will be in proportion to the degree of certainty of the premise" (Moore, 607-608). Even though he did not state Hume's principles, Moore doesn't think it matters what they are, for his premise is "much more certain than any premise which could be used to prove that [it is] false; and also much more certain than any other premise which could be used to prove [it is] true."[155] So, for Moore, the certainty of the premise together with facts about the structure of arguments leads to the conclusion. While this sounds plausible enough, there is reason to be concerned about the term "certainty." Since this measures a degree of confidence, it seems to make rational evaluation subject to irrelevant differences in character.

When we turn to Huemer, however, we seem to have the most secure explanation for why we should endorse this line of reasoning. The account of evidence defended in the last chapter fits perfectly with the common sense view of how to evaluate skeptical arguments. Common sense beliefs seem most likely to be true. In addition, they serve as a basis for explaining why many other things seem to be true. This integration of common sense belief makes it difficult if not impossible to remove these beliefs from one's view of the world without severely changing one's overall understanding of reality. If rationality consists in following how things seem, then views that require one to overthrow central elements of our common sense worldview will almost certainly fail to provide a worldview with superior evidential weight. The nature of evidence creates certain limits on rational belief revision, and rules out the legitimacy

[155] Moore, 1939, 609

of views that take issue with heavily integrated, pre-theoretical assumptions about the nature of the world that shape how things seem to us in everyday life. Since such presuppositions of ordinary thought and action are what common sense consists in, it follows that one should not abandon key elements of common sense.

These points about the nature of common sense and the authority of foundational beliefs combine into an effective way of looking at the above considerations. First, consider some consequences of the fact that common sense beliefs are presuppositions of ordinary thought and discourse.

(1) -The fact that they are *presuppositions* of thought shows that they seem to be true to a sufficient degree that we don't feel the need to defend or state them. This shows they are among the strongest or most evident of seemings.

(2) -The fact that they are presuppositions of *ordinary* thought shows that they are heavily integrated into our belief structure. We use them often and in various arenas. As a result, rejecting one of them does a great deal of damage to our overall belief structure, and requires in most cases the rejection of a large disjunction of things that seem true.

(3) -The fact that they are presupposed *in discourse* shows that they are part of an inter-personal belief structure. That is, they are shared in common by large numbers of people with similar intellectual abilities. Since their presence is more or less unquestioned in others, it does not appear that their presence is likely to be something on which individuals have greater levels of expertise.

Now, suppose you believe that something is true. You then acquire some information and/or engage in a process of reasoning that seems to suggest it isn't. Then you have the following possibilities to consider:

(A) Your reasoning is sound and your information is accurate, so your belief was wrong, and you should change your mind.

(B) Your reasoning is sound but your information is inaccurate, so you shouldn't change your mind.

(C) Your information is accurate, but your reasoning is unsound, so you shouldn't change your mind.

In the presence of a conflict, you have a choice about which of these options to believe. Given the facts above, when it comes to a common sense belief, choosing (A) requires some degree of irrationality. (1) shows that in changing your mind, you are likely dismissing something you have stronger evidence for on the basis of things you have weaker evidence for. (2) shows that you are likely to have to subsequently reject a long disjunction of beliefs in favor of the conjunction of premises that currently seem true. Since a long disjunction becomes increasingly more likely and while conjunctions become increasingly less likely, it is probable that this preference is myopic. (3) shows that, with regard to beliefs other have an equal expertise about, you are none-the-less proclaiming yourself capable of judging the matter independently and of overturning the shared and nearly unanimous judgments of your epistemic peers. This requires a unpardonable level of arrogance. When your reasoning comes into conflict with common sense beliefs, options (B) and (C) are almost certainly both more probable than option (A). Accepting (A) seems to require a misevaluation of the comparative strength of seemings, a myopic focus on a small portion of the total data set at the expense of a larger one, or an inexcusable level of arrogance with respect to your own ability to reason.

3.5- The Providentialist Problem

Much of what has been said so far seems to many people to have a great deal of *prima facie* plausibility. People know lots of things, and have a certain natural way of understanding the world built in to how they think, reason, and act. We have various means of knowing, and when we use those means in their proper spheres, some things are just obvious, and we should believe them. When reasoning about what else to believe, we shouldn't trade weaker evidence for stronger evidence, and part of that is that we shouldn't give up on the obvious for the less than obvious. Most of this sounds rather sensible. It also strikes many people as dissatisfying. Why, after all, should we think that the obvious things are the right things to believe? Why should we think that the surface level facts aren't betrayed at deeper levels that we can discover in other ways? What assurances do we have that we really are making a mistake when we reject common sense?

Here, the tradition's foremost historical proponents have had less than fully compelling reasons to offer. They have pointed out that such practice is an aspect of how we do evaluate truth in matters of importance. Moore also offers a version of the regress argument for foundationalism,[156] plausibly asserting that common sense beliefs should be among the foundations. But Moore, to the best of my knowledge, has no well-developed epistemological account of the underpinnings of common sense. Reid, in contrast, does have one, but it is one that is less than likely to garner universal support. For Reid, much of his certainty that we can rationally follow common sense rests in what Derek Brookes calls his "Providentialist Epistemology."[157] Reid views the beliefs that naturally result from our constitution and from the unreflective use of our innate faculties as the inevitable effect of those sources of belief God has built into us. He believes these faculties are designed to represent what is true, and therefore

---

[156] see Moore, 1939

[157] See Brookes in Reid, 1997, pp.xi-xiii.

accepts them as dependable sources of evidence. Whatever merit such a view may have,[158] it is unlikely to become a secure contemporary foundation for a defense of common sense.

However, given the connection between (PC) and the common sense approach, and the argument offered against overvaluing arguments in Chapter 1, both put the issue in a clearer light. The reason to accept these views isn't because they are guaranteed to be correct; the reason to accept them is because views that accord with common sense are the most rational views for us to accept. If the truth is fundamentally incompatible with our basic ways of thinking about the world, then using our faculties of reason to understand the world around us simply won't allow us to accurately understand the world we live in. Going beyond the evidence in random ways is very unlikely to hit on the truth, and following the evidence in rational ways can't lead to a justified picture of the world that is too different from our common sense conception of the world, because it would have to conflict with our strongest and most heavily integrated seemings. How optimistic one should be about our prospects for understanding more of the world than is built in to common sense is likely affected by how reliable we should think our natural faculties are, but however reliable they are, they appear to set the limits on what sorts of worldviews one can rationally accept. Common sense philosophy isn't necessarily a view that says that human beings are well-designed to figure out the world, but it is a view that says that if we aren't well-suited to do so, there's not a lot we can do to make it better. Because we're all going to suffer from the limitations we face as a species, the best we can do is trust in the general reliability of our faculties of judgment and build the most compelling view of reality we can using the intellectual tools we have available. Since there is no rational escape from believing what seems true, and since the overall plausibility of worldviews can't be rationally revised in

[158] For a recent defense of the view that trusting our natural faculties is best defended by the view that God is responsible for them, see Plantinga, (2011), Part VI.

ways that dramatically conflict with common sense, the common sense approach is the optimal approach to understanding the world.

3.6 Common Sense as a Solution to Perverse Incentives

Another reason to endorse common sense philosophy is that it does a great job solving the problem presented in the first chapter of this work. As we saw from discussing the theoretical defense of common sense, this approach is one of the few methods for doing philosophical work that doesn't over-emphasize the value of philosophical arguments or allow sophistical reasoning to blind us to simple truths. A view that requires us to take highly counter-intuitive conclusions as evidence of our own philosophical errors rather than evidence of widespread errors in others will also help fight the desire to embrace shocking conclusions. In addition to this, there are three main areas where common sense helps fight these incentives. First, common sense philosophy shares many advantages with contemporary analytic philosophy. Second, its affiliation with ordinary language and judgment helps it fight certain contemporary methodological assumptions often endorsed in analytic philosophy. Finally, its methods don't place unwarranted ontological limits on results that one must fit their worldview into, and therefore it has a distinct advantage over naturalism. We will look at these in turn.

3.6.1:Praise for Analytic Philosophy

Since Socrates' condemnation of the Sophists, philosophers have recognized the dangers of using superficially persuasive reasoning to legitimize error, and the ability to use subtle mistakes of reasoning to convince others of pretty much any view is perhaps the longest-standing difficulty for philosophy. Analytic philosophy's focus on logical rigor fights the ability of philosophers to get away with such errors. At the very least, it requires philosophers to engage in far more clever and complex efforts in order to get away with mistakes in reasoning, since

philosophers are now well trained to uncover most gaps or errors in one another's arguments. This strength is enhanced by common sense philosophy, which sets us on guard against complex arguments in general by teaching us why they typically fail to provide good evidence, and by setting clear limits on which philosophical conclusions are deserving of serious philosophical consideration.

In addition to being unclear or sophistical in one's reasoning, unclarity about exactly what one is saying is another way for philosophers to appear to say deep or interesting things without others being able to easily identify the problems with their views. It was largely in response to this sort of problem that analytic philosophers began to endorse views like positivism. Rudolph Carnap and other positivists lamented the state of metaphysics to such a degree that they claimed it all to be meaningless.[159] While this went much too far, this approach did help curb the types of errors one could legitimize by couching it in complex and ambiguous language. By encouraging careful attention to both the reasoning one engages in and the precise meaning of the terms one uses, two of the most common methods for obscuring error or for defending confused but interesting ideas are heavily discouraged in analytic philosophy. Common sense philosophy endorses these aspects of the analytic approach, and therefore has similar advantages in terms of solving the incentive problem discussed in chapter one. In addition, however the aspects of common sense philosophy that set it apart do even more work in this regard.

Common sense philosophers should endorse these aspects of analytical philosophy. First, it should embrace careful conceptual analysis as part of its method. Presuppositions of thought are central to common sense philosophy. Carelessness in examining the concepts involved in such presuppositions is therefore likely to be even more problematic in common sense

[159] See, e.g., Carnap or Ayer's papers, both entitled "On the Elimination of Metaphysics"

philosophy than it would be in other fields. In addition, common sense philosophy emphasizes the importance of careful, gradual development of theories from initial data points. Such care must certainly include, in the case of deductive reasoning, an insistence on a rigorous and detailed examination of any process of reasoning. Therefore, common sense philosophy should include these aspects of the analytic approach that have helped it fight incentives to use sophistical reasoning to buttress philosophical errors.

3.6.2: The Authority of Ordinary Language and Judgment:

Philosophers often think that one of the main purposes of philosophy is to challenge all preconceptions and to use reasoning as a basis for all of our judgments about the world. This is believed to fight against the influence of unjustified prejudice and erroneous pre-conceptions. However, for reasons discussed earlier, there is good reason to think that this goes too far. For hundreds of years while trying and failing to find reasons to trust perception, philosophers asked the average person what they were going to believe: the philosophical reasoning of a great thinker or their own lying eyes. *And philosophers somehow sincerely expected people to choose the reasoning of the philosopher*. This dismissiveness of the authority of common judgment is mind-boggling when considered by itself, but it makes sense given the incentives in the field. Contempt for common belief is often widespread among those who wish to be more unusual in their views, because without such contempt maintaining those views is nearly impossible. If you don't think that your expertise in an area is vastly greater than those you are supposed to be educating, then you can't proclaim your own views sincerely. The idea that philosophical reasoning is the true source of authority in every area of belief creates such a vast disparity between philosophers and others that virtually every issue suddenly becomes one where we have the authority to use our reason and place it above the judgment of nearly everyone in the world,

no matter how implausible our conclusions are. Common sense philosophy does not afford philosophers this authority in most matters, and requires such careful reasoning in efforts to go beyond the data that it severely limits the range of issues that are properly philosophical. Because of this, common sense philosophy is far less likely to fall into error in these areas. It may occasionally find itself too willing to accept certain prejudices masquerading as common sense, but given the history of philosophical belief, I think this trade off is more than worth it.

This view about the authority of common sense extends to the view that ordinary language carries ontological commitments. Philosophers who claim that ordinary language doesn't commit us to anything about the nature of reality also claim that we therefore must do philosophy to understand the true nature of reality. The idea that words don't carry ontological commitments broadens the area of the proper application of philosophical reasoning by raising the issue what's "really" out there. Here, some of the lessons of Reid's criticism of the theory of ideas seem to apply. The effort to open up to philosophical debate ontological questions that are properly settled by common sense is a mistake. This is just as true when it is by pretending ordinary words don't try to refer to actual things, or that distinctions in language don't reflect judgments about the nature of reality, as it is in the effort to pretend that seeing things doesn't have bearing on justification, and therefore we must engage in speculative philosophy to find some grounds for knowing that what we see is really there. In both cases, there is an effort to legitimize the misuse of philosophical reasoning to applications for which it is both ill suited and unnecessary.[160] As Reid describes it

[160] There also seems to be much similarity in the defenses of such efforts to expand the scope of philosophy. I rarely find much said in defense of opening up ontological questions beyond the fact that some words seem not to carry any such commitment. But such an induction seems extremely hasty, since the examples appear rather like peculiar exceptions than like paradigm cases. Like the presence of illusion in perception, they stand out because they are surprising, and because they counter our clear expectations, in this case that terms will be used to refer to actual things. Also like the illusions, the exceptions don't seriously challenge the legitimacy of the ordinary case.

> The vulgar have undoubted right to give names to things which they are daily conversant about… [and] If it is a good rule, to think with philosophers, and to speak with the vulgar, it must be right to speak with the vulgar when we think with them.

Or, as David Stove more curtly puts it, "People who erect a metaphysics on the basis of saying that common English words mean something which everyone knows they do not mean, deserve all the neglect we can give them."[161]

Ordinary language philosophy is another aspect of 20th century analytic philosophy that occasionally went too far in its willingness to read directly off ordinary language, but it is still vastly superior to the alternative approach of pretending that language doesn't carry such commitments. Much as optical illusions give us reason to sometimes question vision, illusions of ontological commitment might give us reason to sometimes question the wisdom of reading off of ordinary speech. But, much like trusting vision, reading ontology off of ordinary speech is a much safer practice than ignoring the apparent content of either perception or ordinary language and trusting the issue to philosophical reasoning alone. The idea that we should ordinarily read commitments off of language is based in the fairly banal observation that people usually talk about the world around them. While this may have exceptions, general practices that ignore this fact will open up philosophy to widespread error and leave us with far less to work with in our efforts to understand the world.

3.6.3: Common Sense vs. Naturalism

Probably the main competitor to common sense philosophy as a means of fighting these incentives is naturalism. One great thing about naturalism is that it limits the scope of

[161] Stove, 1991, p. 143

philosophical inquiry in certain ways that it should. With the rise of naturalism, massive concerns about perception and many vague and useless arguments in speculative metaphysics faded into the background and lost much of their sway in philosophy. This limitation on the scope of philosophical authority was a benefit to philosophy, at least in part because it helped solve some of the problems discussed earlier. However, naturalism appears to go too far in the other direction. We saw from the tradition begun by Descartes how the idea that all of philosophy should try to make itself fit a particular patterns used in another respected field left us with inadequate resources to account for very basic knowledge in other areas. Although its focus is on science rather than math, naturalism appears to make a similar error by requiring all knowledge to fit the evidential pattern used in the natural sciences and ignoring the legitimacy of other types of evidence we frequently and appropriately use in other areas.

What common sense philosophy would oppose about the scientific worldview isn't the 'scientific' part, but the 'worldview' part. Common sense philosophy would take the premature, speculative acceptance of any worldview that wasn't grounded in and limited to those facts that clearly flowed from careful observation (in the broad sense) and just reasoning from such observation, to be unjustified. The opposition to naturalism rests in the fact that it is simply bad science because it places an artificial limitation on what we can come to know of the world that isn't grounded in a proper evaluation of the things we observe in the world. The fundamental difference between the scientific approach to philosophy and the common sense one is really a matter of the scope of the initial data set. Common sense philosophy takes the relevant data that needs to be explained to include the surface-level facts available to us through the use of all of our intellectual powers, and to include the aspects of our ordinary grasp of the nature of reality. The naturalist seeks to limit the initial data set to those results that come from empirical

observation of the external world. On the common sense approach, however, there is no privileged ontology that has any place in philosophy, except the one that rests in the objects of common knowledge.

One reason to prefer this broad approach of common sense philosophy is that a common pattern seems to be present here. If the account of evidence defended in the second chapter is correct, then all of our evidence is of the same basic type: it consists in seemings that occur when we evaluate various types of propositional content. Since seemings of great clarity and forcefulness exist in the exercise of each of our means of knowing about the world, our metaphysical picture of the world would have to account for the data provided by the seemings of each of these faculties of judgment in its development. This sets significant limits on the range of acceptable metaphysical views, and thereby significantly limits the range of interesting or novel pictures of reality a philosopher can legitimately endorse. Given the problem described in chapter one, this fact should lead us to suspect that most methods will encourage us to ignore or severely downgrade the importance of at least some of these sources of evidence. Getting rid of all of them would lead one too far astray from the appearance of rationally pursuing an understanding of the world. But getting rid of one or two opens up new metaphysical possibilities while maintaining enough rationality to seem to be admirably pursuing the truth. Ignoring the data of perception was necessary before philosophers could seriously endorse all those wild varieties of idealism. Ignoring the commitments of ordinary speech has been necessary for the promulgation of many contemporary metaphysical pictures of various parts of reality seen today. Ignoring the legitimacy of rational intuition was necessary to get most of 20th century meta-ethics off the ground. And ignoring the data of introspection has been necessary for physicalism to reign today.

Each case follows the same basic pattern: first, people have pretended that the faculty in question is supposed to be somehow infallible. Next, they have provided examples of ways in which the faculties sometimes go wrong. Finally, they conclude that this justifies a general skepticism about those faculties, rather than an acknowledgment of basic human fallibility that pervades all judgment, but occurs within a backdrop of sufficiently regular accuracy to have made those occurrences so surprising in the first place. Naturalism's privileging of empirical data and its dismissiveness of the significance of introspection and rational intuition fits the basic pattern that has been used throughout philosophy to open up new and interesting metaphysical possibilities. It requires us to ignore many clear and forceful seemings in order to maintain the legitimacy of a certain ranges of worldviews that otherwise would not appear to be well motivated in the absence of these efforts. Common sense philosophy includes no such presuppositions about the nature of the world, and can therefore let the data, in all its forms, speak for itself.

This fact gives common sense philosophy an advantage over naturalism because it fails to include unsupported methodological assumptions that open up wide areas of speculation. For example, if one simply accepts our apparent moral knowledge as evidence that we can learn evaluative facts, one needn't try to fit such facts uncomfortably into one's ontology or else ignore obvious moral truths. And, as I shall be arguing in the rest of the work, if one is willing to accept the introspective data about the apparent nature of our mental states, one doesn't have to try similar measures to account for those. Meta-ethics and philosophy of mind both appear at first glance, from the standpoint of common sense, to be fields motivated by the rejection of the apparent data offered by two legitimate sources of evidence that subsequently results in voluminous fields of philosophical inquiry. One can then either try desperately to account for

fairly obvious truths within an ontological and methodological framework not equipped to account for them, or else simply become hostile toward the data itself. Either way, they appear to be instances where a methodological presupposition leads to the immediate rejection of common sense and the subsequent willingness to accept one of a wide number of possible ways of going wrong.

Seen from this perspective, one cannot help but draw parallels between contemporary work in these fields and the early-modern obsession with finding a plausible variety of idealism. The direct and properly decisive evidence in favor of the common sense view is systematically trivialized by philosophers, the field opens up and creates wide arrays of attempted answers, but none of these answers really change the way things seem, how people act, or what they say in the areas in question. I believe Michael Huemer has done a great job illustrating this problem in meta-ethics in his book *Ethical Intuitionism*. I see the remainder of this work as a similar effort in philosophy of mind. When Reid opposed the foremost figures of his day by defending common sense realism in the presence of widespread rejection of this view about perception, he predicted of future views about Berkeley and Hume that

> Some ages hence, it will perhaps be looked upon as a curious anecdote, that two Philosophers of the 18th century, of very distinguished rank, were led by a philosophical hypothesis; one, to disbelieve in the existence of matter; and the other, to disbelieve the existence both of matter and of mind. Such an anecdote may not be uninstructive, if it prove a warning to Philosophers to beware of hypotheses, especially when they lead to conclusions which contradict the principles, upon which all men of common sense must act in common life.[162]

[162] Reid, 2002. P. 162-3

I hope that today we can learn from this example just how wrong things can go when a philosophical hypothesis leads to conflict with common sense, and not continue to repeat the practices of the past that waste so much genius in pursuit of views that conflict so deeply with what we know about ourselves and the world we live in. Given the pattern contemporary philosophy appears to be repeating, and the reasons we have seen for thinking this is reflective of a broader phenomena within the field, there is good reason to think that if a viable ontology can be obtained from common sense, we will one day look back with similar puzzlement at the materialism of today as we do when we look at the long-standing prominence of idealism in early modern philosophy. By unearthing this ontology, common sense philosophy, at the very least, will present a viable alternative to naturalism's materialist worldview as a legitimate metaphysical picture of the world.

SECTION II- The Common Sense View of the Mind

Chapter 4: Mind and the Physicalist's World

An Overview of Section II

So far I have defended a general approach to philosophical problem solving that takes our common sense intuitions as the cornerstones of philosophical progress. I have argued that a proper understanding of the nature of evidence underwrites a broad-based foundationalism that treats our various intellectual powers on a par in terms of how they furnish us with an understanding of the world. This epistemological view underwrites the common sense approach to philosophy defended by Moore and Reid. Due to the perfect generality of the account of evidence, the method thereby supported can be applied to an array of issues. In the remainder of this work, I will focus on the method's application to philosophy of mind.

The application of this approach leads me to deal with questions in philosophy of mind in a more traditional manner than is common in contemporary works. The widespread acceptance of naturalism in philosophy of mind has led to more of an empirically informed approach to the field, and has shaped the debate in terms of a view's ability to make sense of the mind within a physicalist framework. Although I will discuss empirical data where appropriate, this is a more classical work in philosophy of mind. My focus is the more general question of how a common sense approach to philosophy of mind shapes our understanding of the evidence. Since a common sense approach takes a much broader view of the relevant data points, I will focus extensively on unearthing and evaluating the implications of the data points of introspection and intuitions in the area. Therefore, many of the considerations I offer take the form of broad, general arguments in favor of various views. Certainly a deeper investigation into the minutia of the debates is important, and I think the influence of the more general approach carries over into

these areas. I hope to be able to address several of the areas I cover in more detail in future work, but here I focus on the big picture in order to a compelling framework for engaging in common sense philosophy of mind in the future.

I begin with a general discussion of the common sense view of the mind. I conclude that, although there is a fairly strong *prima facie* case for the claim that dualism is an aspect of common sense, a closer examination is needed to ensure that certain alternatives can be set aside. I then lay out what I take to be the common sense view of the mind and its characteristic features, avoiding for the moment the broader ontological questions about the significance of these features. This should provide us with a good sense of the nature of the ontological questions to be explored later by clarifying the nature of the subject matter of the debate. Before looking at the various features of the mind, I begin with an account of the physicalist's project, focusing on a minimal commitment of physicalism with regard to the nature of the mind. This account provides a background against which to evaluate the ability of physicalism to adequately account for the nature of the mind. In the next three chapters, I proceed through the various characteristics of the mind, showing, first one at a time, and then in conjunction, what it is about our grasp of the nature of the mind that is at odds with the physicalist world view, and how these intuitions of distinctness drive us toward a dualist metaphysics in order to capture an appropriate understanding of the nature of the mind.

4.1 The Common Sense View of the Mind

In ordinary language there is a natural tendency to differentiate between the physical and the mental. We speak of being physically well but psychologically troubled, of using mind over matter, of the need to feed the mind and the soul as well as the body. These sorts of statements suggest that there is a natural distinction built in to how we think of things that places our bodies

and the rest of the physical world into one category, and our minds and the rest of the psychological world into a different one. According to substance dualism, this tendency of speech mirrors a fact about the nature of persons. According to this view, people are composed of both a physical body and a non-physical mind. Most people take this to be the most natural and common sense way of thinking about persons. Thus, even as he hopes to "free us from its grip", David Papineau notes that "we find it almost impossible to free ourselves from the dualist thought that conscious feelings must be something additional to any material goings-on."[163] David Lund goes so far as to claim that all materialist views of persons run into several "head-on clashes with hard-core commonsense belief."[164]

Echoing these claims, Alvin Plantinga, citing others, says, "As Paul Churchland, Jaegwon Kim, and many others say, dualism is the natural, baseline belief of humankind, not an invention of Plato or Descartes."[165] However, while the view that dualism is common sense is certainly widespread, there are dissenters to this opinion. John Searle, after describing his own preferred version of materialism claims that it "is a kind of commonsense answer to the question, given what we know about how the world works."[166] Searle believes that much of the force of our apparent commitment to dualism rests in the fact that "we persist in talking about a twentieth century problem in an outmoded seventeenth-century vocabulary."[167] Confusions grounded in failures to adapt vocabulary as quickly as science progresses certainly aren't uncommon. As Reid points out

---

[163] Papineau, 2002. p. 3
[164] Lund, 2005. p. 20
[165] Plantinga, 2006, p. 120
[166] Searle. 1984. p. 28
[167] Ibid, p. 14

> ...when the Ptolemaic system of astronomy, which agrees with vulgar prejudice, and with vulgar language, has been universally rejected, by philosophers, they continue to use the phraseology that is grounded upon it, not only in speaking to the vulgar, but in speaking to one another. They say, the sun rises and sets, and moves annually through all the signs of the zodiac, while they believe that he never leaves his place.[168]

So it is possible for claims about the commitments of common sense to be affected by abuses of language. Searle seems to be suggesting that our progress in the sciences has made it more sensible to think of the mind as something realized in the brain. He thinks that, once we learn these things, we should even think of some version of materialism as the common sense view.

The history of philosophers discussing common sense has been full of peculiar claims. In perhaps the oddest instance of an appeal to common sense, Bishop Berkeley once claimed that idealism was the common sense view of the world, and that, in fact, the whole idea of matter was incoherent and no one ever really believed in it. However, on the other side of things, there is also a history of assuming that common sense says more than it actually does. People often claim that science has shown common sense to be mistaken about all sorts of things that in fact common sense was never really committed to. For example, it somehow came to be believed that the common sense meaning of 'solid' entailed that an object wasn't composed of any empty space. Clearly this isn't what the word means. 'Solidity', as it is used in common language, is a matter, first and foremost, of how penetrable something is, not a matter of its underlying constitution. Therefore, however scientists are using the terms, the fact that particles are mostly composed of empty space doesn't in any way refute the common sense view that objects are

---

[168] Reid, AP, 275

solid. I am quite sympathetic to the view that many of the counterintuitive, even bizarre-sounding "discoveries" of science are little more than abuses of terms, grounded in a rapid march of progress combined with a rigid maintenance of vocabulary no longer well suited to describe the observed phenomena. So the possibility of a misunderstanding of the actual commitments of our mental terms as a source of the apparent common sense commitment to dualism can't be dismissed out of hand. To figure out which sort of case Searle's claim is, we will need to look closely at our mental concepts and how we use them in understanding the world.

I have the natural reaction of most dualists that Searle is making a mistake akin to Berkeley's, but there seems to be less than a consensus on this point. Many of my colleagues insist that their intuitions of distinctness are far less clear and distinct than my own, or those of most people I discuss these issues with who aren't in the field. These apparently sincere claims of a lack of a robust sense of distinctness on the part of some members of the field further emphasizes that the case for holding dualism as the default, common sense view of mankind is less quick and straightforward than one might have hoped. To better evaluate whether physicalism is really the radically revisionist view I take it to be, or merely a clearer, if sparser account of the nature of the mind, we need to take a closer look at our mental concepts, and the resources available to the physicalist in accounting for them.

4.2 The Nature of the Mind

The ordinary understanding of the mind includes the following features. First, the mind is the thing that thinks, desires, and imagines. When engaged in any of these activities, it has states, such as beliefs and desires, that have content (that is, they are about something). This content represents various possible ways the world can be. Just as a statement or a written

sentence represents the world as being a certain way, certain states of our minds do as well. This feature of being about something else is typically called "intentionality" by philosophers.

Second, the mind has subjectively recognizable qualitative features. When we have experiences, there are aspects to them that are felt or sensed internally. These include things like the hurtfulness of pains, the sorrowfulness of loss, the tingly heart-throbbery of falling in love, and the characteristic sensations of tasting bacon, smelling bacon, and hearing the crackle of bacon when it cooks. These qualitative or "felt" aspects of our experiences are typically called "qualia" by philosophers, who take the expression from the word 'qualitative'.

Third, the mind has an active nature in the sense that it brings about events through acts of the will. The will is used in forming intentions to act. When we decide to eat cookies & cream ice cream instead of mint chocolate chip, we form an intention and then act on that intention by ordering the cookies & cream. We think that the exercise of the will in forming our intentions is what causes most of our behavior. We naturally think of the will as free in its activity. The will is free when nothing requires it to form the intention that it does. It was within our power to choose the mint chocolate chip instead of the cookies & cream, and if we had formed the other intention, then different behavior would have occurred. It is when we freely choose to act in this way that we take ourselves to be responsible for our behavior. So moral responsibility seems to be connected to activity that originates in the mind.

Finally, the mind is understood not as a mere amalgamation of these features, but as the part of us that has these features. We think of the mind as a distinct and (in principle) separable part of us in the same way that, say, our legs or our arms are distinct and separable parts of us. The possibility of the disembodiment of the mind is a feature of many views of the afterlife and is taken advantage of sometimes in storytelling where characters switch minds and thereby learn

what it's like to walk in someone else's shoes. Like losing our legs or arms, we think that when separate, the mind would maintain its intrinsic features such as our beliefs, desires, and memories, but of course lose its ability to affect and be influenced by the body to which it formerly belonged. Unlike legs and arms, however, we tend to think that the mind is the essential part of us. We go where our minds go, so when Donkey and Puss & Boots get switched in Shrek the 3rd, it is Donkey who gets transferred to Puss & Boots' body, not Puss & Boots who gets Donkey's mind.

The mind, then, is the distinctive substantial part of us that has thoughts and desires with intentional content, experiential states with qualitative aspects, and actively produces our behavior through the forming of intentions. Now that we have a fairly clear grasp on the nature of the mind, we can begin to ask questions to help us understand the mind's relation to our bodies and to the rest of the world. Later, I will attempt to show that the most natural and plausible account of the mind and the features it possesses commits us to substance dualism. Given the work of the earlier chapters, this accordance with common sense would grant substance dualism a default, presumptive status in philosophical debates.

4.3 What is Physicalism?

By defending dualism, I will be taking on the dominant position in the literature: physicalism. Arguments against physicalism face a difficulty of pinning down their target. Physicalists have offered a wide variety of views and taken on vastly different commitments in the literature. Here, I specify what I take to be a minimal commitment of physicalism. If it can be shown that, even on this construal, the mental is too closely tied to the physical to adequately account for our mental concepts, then it can be concluded that the stronger versions will also fail

in this respect. I take physicalism to be a metaphysical thesis.[169] Physicalism is a claim about what there is in the world. It is not a claim about which concepts apply to the world, or how we should approach philosophical issues, or what languages correctly describe the world. Given this, I treat issues in philosophy of mind as questions about the relationship between mental properties or states and physical properties or states, and not as a question of the relationship between, say, psychological language and physical language, or between mental concepts and physical ones. These connections can be illuminating, but they should not be confused with the fundamental issue, which is one of ontology.

Given our focus on ontology, the key notion to explain in order to understand the minimal commitment of physicalism is *realization*. Realization is clearly a metaphysical notion. In the broadest sense, realization is a matter of *making something real*, and this is certainly a matter of ontology, not of language or concepts. Since most of the recent debate in physicalism focuses on property instances, the account of realization I offer takes them as the relata.[170] Realization is captured by the following definition.

> For any two properties F and G, F realizes G in x iff Fx and Gx and x's having G consists in x's having F.

When two properties are both instantiated in something, one realizes the other just in case the existence of the realized property consists in the existence of the realizer. The "consisting in" relation is to be understood as establishing that the realized property is nothing over and above the realizer in this particular case. "Consisting in", then, is a stronger relation than

---

169 For a different approach to defining physicalism see Ney (2008)

170 I am remaining neutral here on exactly what property instances are. I think that these points can be translated into many different views on the nature of properties and how they are instantiated in particulars. Given the prevalence in the field of the type/token distinction and the popularity of views that endorse multiple realizability, if there is no sense to made of such talk then much of the literature is in trouble. Realization in other ontological categories may well be definable in terms of this relation as well (e.g. one object may be realized in another just in case each (non-modal) property instance of the realized object is realized by some property of the realizer).

supervenience or "in-virtue-of-ness." Not only do realized properties supervene on their realizers, but they have no existence independent of them (in any particular instance). This relation compares favorably to an account of how one fact can be constituted by another offered by John Foster. Foster gives the following explanation of this relationship:

> Let us say that a fact F is wholly constituted by a fact or set of facts F' if and only if the following two conditions hold: (i) F obtains in virtue of F' (in some way that makes the obtaining of F asymmetrically dependent on the obtaining of F'), and (II) the obtaining of F is nothing over and above the obtaining of F'.[171]

This account of the relationship between facts is roughly analogous to the account of realization of properties offered above. Some examples should help further clarify this idea.

Let's start by considering a quarter. One realized property of the quarter is its shininess. The shininess of any particular quarter is nothing over and above the surface reflectance properties of that quarter. One you have the surface properties you have the shininess. In addition to this (probably) intrinsic property, quarters also realize very complex relational properties like "being worth 25 cents." The monetary value of a quarter consists not just in its appearance and metallic composition but also, and more importantly, in the social agreements people have made about how quarters are to be used in transactions. Once we have the surface properties or the social agreements in place, that is all there is to shininess or monetary value of the quarter. We don't need to somehow add on these properties, nor do they magically appear by some causal process. What happens, rather, is that the existence of the realized property consists in the existence of the underlying realizer, and has no existence independent of it. The same thing holds in other cases of realization, such as being a university, a battalion, or a brake on a

[171] Foster, 1991, p. 5

vehicle. In each case, having the underlying realizer is all there is to the fact that the realized property is present.

The lack of independent existence of realized property instances is crucial to any discussion of realization. The debate in meta-ethics between naturalists and non-naturalists may help us keep this in mind. The non-naturalist like Moore is not going to deny that once the non-evaluative features have been fixed, the evaluative properties will supervene on them[172]. What they deny is that the natural property instances, such as say, instances of pain sensations, constitute a moral property like badness. According to the Moorean, goodness and badness are simple qualitative properties that are had in common by many different things and unify them under the same category. This simple property, though it supervenes upon and necessarily comes with the underlying base, does not consist in, and therefore cannot be reduced to, that base. Rather, it has a simple nature of its own, over-and-above the natural properties that underlie it. This distinction between the moral property dualist like Moore and the Cornell-style naturalists[173] in ethics is valuable to keep in mind as we use the notion of realization to help us understand the nature of physicalism in philosophy of mind. Since any weaker relations will be compatible with various versions of property dualism, the physicalist needs to be able to ensure that this minimum standard of constitution has been met to earn her physicalist credentials. Physicalism in philosophy of mind, then, is committed to the view that all mental property instances are identical to or constituted by some physical property instance.

4.4- Functional Features, Compositional Features, and the Physicalist's Project

Our concepts of properties typically include a functional or causal profile to their nature. This profile includes the sorts of causal relations that must hold for the property to be instanced,

[172] See Moore, 1903
[173] See, e.g., Boyd, 1988, Sturgeon, 1985

and the sorts of roles or functions these relations allow the thing that has the property to play. In addition to their functional profile, properties sometimes also include what I call a "compositional profile." This profile isn't always precise, but there is a general sense that certain properties only exist if the right sort of thing is fulfilling the profile in question. Thus, going back to the quarter, most of what it takes to be a quarter is to fulfill a functional role in our economy, but in addition to this, it has to be made out of the right sort of metal. Quarters can't be made of aluminum or of gold. Similarly, as we have learned more about water, it seems to have been accepted that it must be made out of H_2O to qualify. If it's not composed of the right type of stuff, then it's not water. This compositional component also seems to underlie the Moorean intuition in ethics. Those who have it believe the reductive project must fail because goodness has a simple compositional nature that isn't captured in any reductive analysis. The intuition of distinctness in philosophy of mind seems to suggest that our mental states have such a profile as well.

When we come to understand the application conditions of a certain concept, what we learn are the functional and compositional profiles of that concept. The functional profile tells us how something must behave in various situations in order to fall under the concept, while the compositional profile tells us what it must be composed of in order to fall under the concept. Many concepts require that both types of profiles be satisfied. Purely functional concepts, however, have no compositional profile. Functionalism in philosophy of mind is the view that all mental concepts are purely functional, in the sense that no particular type of thing is required to fulfill the role in question. Ideally, a realization account will allow us to understand how each of the constitutive elements of our concept of a property already consists in the underlying realizer. The reason that functionalism in philosophy of mind is so popular among physicalists is

that if all there is to our conception of a mental state such as a belief, or a state of pain, is that the right functional role is fulfilled, then it doesn't matter what satisfies the role, and we can let the neuroscientists figure that out for us. However, if, in addition to satisfying the role in question, beliefs and pain states need to have an appropriate composition for our concepts to apply to them, then the stuff the neuroscientist is looking at doesn't seem up to the task. Our mental concepts obviously do not have a compositional profile that includes being made out of synaptic activity in squishy gray stuff, or even of being made out of matter of any sort whatsoever. That can be seen in the fact that people understand the claim that God has thoughts, but they don't think that God is a big hunk of matter. When dualists claim that our mental states can't be realized by physical states, they do so for two reasons. First, dualists believe that mental states have a compositional nature. Part of the essence of a state of pain is the felt aspect of the experience of pain. The sensation is a component of the experience itself. Second, they think this is a simple, basic component, fully present, and fully recognized in introspection. That simple, basic component of our experience is thought to be missing in physicalist accounts, and what is offered as the compositional substitute for it is thought to be too different from what was recognized in introspection to qualify as the same compositional feature. So, if our concepts of mental properties have compositional profiles, then it is clear that the relevant constitutive bases would need to be immaterial, and physicalism would seem to be a nonstarter.

Due to the strength of these dualist intuitions, physicalists have had to tell a story explaining how these apparently distinct states could actually be one and the same. The initial, Rylean move was to tell the following sort of story:

> Oh, sure, they seem different. But that's no problem because mental "states" aren't really states of the world at all; there is no particular thing in the world which is, say, your

> belief that 2+2=4, there's just the fact that you're disposed to say 'four' when prompted with the utterance 'what does 2+2 equal?'[174]

This seemed wrong because we seem to be aware of such states within us, and it seems clear that we are referring to internal states, and not just asserting that some dispositional fact holds of us, when we make mental attributions. In light of this, the physicalist then wanted to say

> Well, indeed mental state ascriptions aren't just asserting some general dispositional fact. They do refer to a state of the world after all. But they only refer to it as whatever internal state happens to be the thing that occupies the right role, and this empty, functional account of the mental doesn't say anything about what the mental is composed of, so now we can trust the neuroscientists when they tell us it's constituted out of brain activity.[175]

However, arguments that seek to strengthen our intuitions of distinctness suggest that such an account still leaves out introspectively accessible features of the mind. In spite of their efforts, the intuition of distinctness still remains quite strong. However, physicalism maintains wide acceptance in philosophy, and its defenders are confident that no matter how difficult these problems are, in the end they can be overcome. Their confidence rests in the idea that neuroscience will show that the causal or functional profile of the mind is already being fulfilled by the brain, and since it would be absurd to think that two different things were doing the same work, we must conclude that they are actually one and the same.[176] To evaluate whether or not the physicalist can capture our intuitive grasp of the nature of the mind, we need to examine the strength of the case for the view that there are compositional features of our mental lives that we

---

[174] See, e.g., Ryle (1949). For a more contemporary effort to defend this sort of view on a more limited scope, see Dennett (1988)

[175] See, e.g., Armstrong (1980)

[176] For a clear formulation of this difficulty see Papineau (2002), esp. pp. 17-18

think of as immaterial, and to examine the physicalist's ability to capture our understanding of the causal or functional nature of the mind as well. Having done so, we will be able to see whether or not, and to what extent, the physicalist can ground our common sense view of the mind, and whether or not their reasons for continuing to try give us grounds for thinking the shortcomings they currently face can eventually be solved.

Chapter 5: Compositional Components of the Mind

5.1: Introspection and Intuition as Sources of Evidence

When attempting to uncover the common sense view of the mind, the first place to turn is introspection. Introspection gives us the most direct access to what is going on in our own minds. The data of introspection, however, clearly has limits. Since we don't have direct access to other minds, it is necessary for us to find some way to communicate what we find within ourselves to one another. In addition, since careful introspection often proves difficult, we need a way to isolate separate elements of our mental lives so that we can attend to them more carefully. The key means we have developed to do this is the creation of thought experiments. Thought experiments seek to extend the limits of introspection by getting others to attend to the features we are interested in. By drawing attention to specific features of our minds through isolating those elements from others, thought experiments can get others to attend to the features we are interested, and thereby create a shared context of experience for discussing the nature of the mind. From a common sense perspective, these aspects of our awareness in introspection serve as a foundation for our work in philosophy of mind. Moreover, since thought experiments work by getting us to share a better grasp of our own minds, these "intuition pumps," as they are sometimes called, enhance our ability to attend carefully to specific elements of our minds and of our ordinary conception of their nature, and they are therefore an invaluable means of aiding us in understanding these aspects clearly.

From within the common sense approach, there is nothing closer to raw data in the field than the objects of awareness in introspection and the clear intuitions about our mental concepts that can be brought out through careful attention to the mind. However, whether or not one endorses the common sense approach, everyone should attend closely to these pieces of

evidence. With the exception of eliminativists and self-avowed error theorists about such states, the staring point for any legitimate reductive effort has to lie with our ordinary conception of the mind. One cannot establish that we should believe in an ontological reduction unless one can show that the elements that constitute our initial grasp of the realized property are compatible with its consisting in some underlying state. One cannot show such a thing unless one already has a clear conception of what is to be reduced, and one can't seriously expect people to accept any such effort unless the conception is actually of the thing we are trying to understand. One cannot, for example, claim to have established a reduction between dogs and fish by claiming that we are mistaken in thinking that dogs are furry things with four legs that walk around and breathe air. The aspect of the relata to be reduced must stay roughly the same throughout the process, or we aren't really *reducing* it, we're *replacing* it, and are thereby committed to either an eliminativist or error theorist account of its nature.

For this reason, the results of the following chapters will go beyond what I would need to establish to justify dualism given the epistemic background in place. In the process of showing that common sense commits us to dualism, we will see that our ordinary conception of both the compositional and functional features of the mind is deeply wedded to a dualist metaphysic and that physicalism, in all its most popular forms, is conceptually deficient in its understanding of the mind. If I can establish that the common sense commits us to dualism, I can show that those who embrace the epistemology of the earlier chapters are strongly committed to accepting it. If I can further show that the commitment shapes and constitutes our very conception of the mental, I can show that anyone who wishes to be a mental realist must also be a substance dualist.

5.2- An Initial Explanation of the Appeal of Dualism

We finished the last chapter with a discussion of the physicalist's efforts to establish that various aspects of our mind consist in some physical state. We saw that there are two potential elements to our concepts, a compositional component and a functional one, and that most physicalists have tried to motivate the view that our mental concepts are purely functional. So, does our concept of the mind include features that have a compositional profile or not? Despite the efforts of physicalists, I think that the answer is still clearly "yes." This can be seen from the strong intuitive appeal of standard thought experiments in philosophy of mind. When Jackson says

> Tell me everything physical there is to tell about what is going on in a living brain, the kind of states, their functional role, their relation to what goes on at other times and in other brains, and so on and so forth, and be I as clever as I can be in fitting it all together, you won't have told me about the hurtfulness of pains, the itchiness of itches, pangs of jealousy, of about the characteristic experience of tasting a lemon, smelling a rose, hearing a loud noise, or seeing the sky[177]

he is not expressing some idiosyncratic concept of the mental, he is latching onto our ordinary understanding in a way that people are consistently drawn to. Similarly, although he won't like my use of the example, when Searle says that

> There is more to having a mind than having formal or syntactical processes. Our internal mental states, by definition, have certain sorts of contents. If I am thinking about Kansas City or wishing that I had a cold beer to drink or

[177] Jackson, 1982, p. 127

> wondering if there will be a fall in interest rates, in each case my mental state has a certain mental content in addition to whatever formal features it might have[178]

he is again capturing something central to our understanding of the mind that every undergrad can latch onto. In both cases, the lesson to be learned is that we know that there is more to the mind than what it does. Our mental concepts aren't mere functional role concepts; they have a compositional profile. It's not enough to be pain that the right behavior is caused at the right times, it also has to *feel* a certain way. It's not enough to be able to regurgitate the right sounds to understand or to believe, there also needs to be the right meaning or content to those states. So our mental concepts do seem to have a compositional profile to them, and not merely a functional one. And, contra what Searle must think about semantics, it is simply obvious that our understanding of meanings and feelings do not commit us to their material constitution; quite the opposite, given that these things have a constitutive profile and not merely a functional one, it certainly seems that they are immaterial. That is why Jackson's Mary case counts as a *dualist* thought experiment in the first place, and that is why we have to tell students, much to their surprise, that Searle is actually a physicalist.

It is due to the strength of our intuitions of distinctness when considering these mental states that physicalism initially seems wrong to just about everyone. The mental and the physical fall into distinct conceptual categories. It seems obvious that mental states must feel a certain way or have a certain content because we are aware of these feels and meanings as their most salient features. And it also seems obvious, at least at first glance, that arranging matter in various ways can't make it constitute a subjective feel or a meaning. Although this seems clear

---

[178] Searle, 1984, p. 31

enough, it will be worthwhile to look more closely at our intuitions about these states to see exactly where the difficulties lie.

5.3: Examining our Intuitions

One classic representation of the intuition of distinctness comes from Leibniz:

> It must be confessed, moreover, that *perception*, and that which depends on it, *are inexplicable by mechanical causes*, that is by figures and motions. And supposing there were a machine so constructed as to think, feel and have perception, we could conceive of it as enlarged and yet preserving the same proportions, so that we might enter it as into a mill. And this granted, we should only find on visiting it, pieces which push one against another, but never anything by which to explain a perception.[179]

Leibniz claims that thinking can't arise by virtue of physical interaction among objects or parts of objects. Alvin Plantinga takes up this idea with regard to content. He describes the current materialist position as follows:

> From the materialist's point of view...a belief will be a neuronal event or structure...But if this is what beliefs are, they will have two very different sorts of properties. On the one hand they will have *electrochemical* or *neurophysiological* properties ('NP properties', for short)...But if the event in question is really a *belief*, then in addition to those NP properties it will have another property as well: it will have a *content*. It will have to be the belief that *p*, for some proposition *p*.[180]

[179] Leibniz, Monadology, 17
[180] Plantinga, 2007, p.107

This propositional content is an additional, compositional feature of our belief states. Notice that our conception of this feature is not one that is in any way similar to the neurophysiological properties, and that our conception of beliefs has no such features as components. In addition, our conception of nuerophysiological properties not only don't include the presence of propositional content, we can't even understand how to add them in. As Plantinga goes on to explain,

> The fact is, we can't see how [an NP property] *could* have a content. It's not just that we don't know or can't see how it's done...the case under consideration is different. Here it's not merely that I don't know how physical interactions among neurons brings it about that an assemblage of them has content and is a belief. No, in this case, it seems upon reflection that such an event could *not* have content. It's a little like trying to understand what it would be for the number seven, for example, to weight 5 pounds, or for an elephant (or the unit set of an elephant) to be a proposition....We can't see how that could happen; more exactly, what we can see is that it *couldn't* happen. A number just isn't the sort of thing that can have weight; there is no way in which the number seven or any other number could weigh anything at all....similarly, we can see, I think, that physical activity among neurons can't constitute content.[181]

There have been efforts to explain content in terms of the causal relations that hold between the state in question and the role it has in affecting our behavior and the behavior of those around us. The difficulty is that we have a sense of content separate from these causes, as is illustrated in the Searle example, and which is grasped in cases of genuine understanding. Bon Jour describes this, saying "one crucial feature of such conscious thoughts is that when I have them, I am in

[181] Plantinga, 2007, p.109-10

general consciously *aware of* or consciously *understand* or *grasp* what it is that I am thinking about,"[182] adding that these features can't be accounted for adequately on externalist or functionslist grounds because

> it is an undeniable fact about conscious intentional content that I am able for the most part to consciously understand or be aware of what I am thinking about 'from the inside.' Clearly *I* do not in general do this on the basis of external causal knowledge: I do not have such knowledge and would not know what to do about it if I did.[183]

This content we see from the inside, which Plantinga calls "original" (as opposed to "derived"), is what seems to be missing in all of these accounts. It seems to be an additional, compositional, and non-physical feature of beliefs and various other mental states. And it is this component that remains missing in efforts to explain content, and that motivates the claim that when it comes to the production of original content of beliefs, "it isn't just that we can't see how it's done...It is rather that we can see, to at least some degree, that it can't be done, just as we can see that an elephant can't be a proposition, and that the number seven can't weigh 7 pounds."[184]

To see this point more clearly, we should look back at Searle's objections to functionalism. Seale claims that in the Chinese room, we wouldn't understand Chinese, and so syntax isn't sufficient for semantics. But what is it, exactly, that is missing in our efforts to understand? Here, I think that Searle proves far more than he meant to. Because what is missing is a certain sort of internal state in which you are aware of the content of the symbols being passed back and forth. When we understand something, there is an intrinsic feature of our mental states (its content or meaning) and it is by grasping or being aware of this feature that we

[182] BonJour, 2010, p.16, emphasis original
[183] BonJour, 2010, p.18, emphasis original
[184] Plantinga, 2007,p. 113

understand what the symbols mean. When we are merely manipulating symbols, it's "all dark inside," and the meaning of the sentences isn't present. Since a computer can only manipulate symbols, it can't help but always be dark inside. The talk about syntax may help us envision this, but it is the contrast between the internal state we recognize when we understand the meaning attached to the symbols and the one we recognize when we are merely parroting the expected result that does all the work. Like other thought experiments in philosophy of mind, the function of the example is to get us to pay attention to some feature through introspection. Most contemporary accounts of content would have us believe that content is either diffused into the functional patterns of the states that possess it or else that content is actually external to the mind. But in this thought experiment, we are drawn to focus our attention on content as an intrinsic features of our minds that we can simply see in introspection. If we can see this feature fully present as an intrinsic component of our minds, then this fact is incompatible with either of these accounts of content. And when we attend to this feature it appears fairly obvious to most observers that this feature cannot be part of the physical world.

In this thought experiment, like most examples used to motivate dualism, once we see what we're supposed to be attending to, there is no more work to be done to convince most people that the feature isn't physical. Searle's example doesn't hold appeal because it details the causal processes of language use and shows that mere syntax, try as it might, will fail to live up to semantic content. It holds appeal because it draws attention to content as a compositional feature of our beliefs, and of other mental states. The intuition that understanding doesn't consist in functions is not explained by the fact that functions aren't quite up to the task, it's explained by the fact that understanding includes a grasp or an awareness of a component of our mental states, namely their content, which is an independent feature, existing over and above the activity

that happens to be going on with the particular mental state of which the content is a component. The difficulty is more fundamental than a debate over what computers can do; it's an instance of a general fact that compositional features can't be reduced to functions, and intrinsic features don't permit externalist accounts. Something has to be there doing the work, and the components of a state are simply a fundamentally different type of thing than the work a state does. Functionalist and externalist accounts of content are missing this component, and, *pace* Searle, this component is not and cannot be some neural structure in the brain. It has none of the components of such structures, and such structures lack the fundamental nature of mental contents.

Perhaps Searle, or others, may wish to say that I've gone too far here. They may think that even if we're aware that content is something over an above the functions being fulfilled, this doesn't mean that we understand it to be immaterial. But I don't think can be right. All of the work in examples like Searle's is done by drawing our attention to the relevant component of our mental lives. If our grasp of their natures in virtue of drawing our attention weren't sufficient to allow us to understand something of their compositional nature, we would not have strong intuitions of distinctness like we have seen above. We aren't aware of the chemical structures of things, even if we know that they have one, but it is in virtue of this that we have no *a priori* intuitions about the possibility of composing things out of the relevant chemicals. If focusing our attention on these features of our mental lives didn't include some sense of their nature, we should expect a similar widespread agnosticism about the possibility of constructing them out of electrical activity in squishy gray stuff, not a strong intuition of distinctness that takes great effort to overcome.

I have focused on content here because I believe it is the less obvious of the two cases. When it comes to qualia, the case is even stronger. When Jackson first sought to convince people that qualia weren't physical, his "argument" consisted entirely of drawing our attention to clear examples of them. Fred and Mary made for nice discussion, but all they help to do is draw our attention more clearly to what it is that is missing in the reductive story, and what it is that we are aware of in introspection. Attending to sensations themselves is all that is needed to see a clear difference between them and any physical state. As Reid put it "If matter [is] what we conceived it to be, it is equally incapable of thinking and of acting freely."[185] One reason dualism is often difficult to offer deductive proofs for is because acceptance of dualism isn't based on an argument. These thought experiments aren't best thought of as arguments for dualism, they're best thought of as tools for pointing to certain features of the mind. They let us focus on these features more clearly and fully, but once we have, the "argument" consists of nothing more than pointing to the feature and saying "That! That thing right there! That is left out of your story, and that thing simply isn't physical." The best explanation for why dualism comes so easily to so many of us is that we are aware of and understand our own minds well enough to draw a basic and fundamental distinction between its nature and the nature of the physical world. Because it is a feature of a basic means of conceiving of the nature of what we experience in introspection, dualism seems to be a basic feature of common sense, not a theoretical addition to our belief set.

5.4- "The Thing Itself" and the Force of the Compositional Intuitions

In the last three sections, I have looked at two characteristic features of the mind and tried to draw our attention to the strength and nature of our anti-materialist intuitions. Some people,

[185] Reid, 1969, p. 356

however, have claimed that an account of the mind cannot be complete until a description of the substance itself is given. Thus Daniel Dennett complains

> There is the lurking suspicion that the most attractive feature of mind stuff is its promise of being so mysterious that it keeps science at bay forever.
>
> This fundamentally antiscientific stance of dualism is, to my mind, its most disqualifying feature, and is the reason why in this book I adopt the apparently dogmatic rule that dualism is to be avoided *at all costs.*[186]

Similarly, Paul Churchland asks

> can the dualist tell us anything about the internal constitution of mind-stuff? Of the nonmaterial elements that make it up?...compared to the rich resources of explanatory successes of current materialism, dualism is less a theory of mind than it is an empty space waiting for a genuine theory of mind to be put in it.[187]

Indeed, even after arguing against materialism, BonJour makes the same basic point, stating

> As far as I can see, the main appeal of substance dualism is that the account of the supposed mental or spiritual substances is far too vague and sketchy to provide the basis for any very clear argument that such substances could *not* be the locus of consciousness.[188]

This request for a description of the type of "stuff," over and above the features that are present in the mind seems to me misguided. To see this, consider our understanding of physical substances. Reid presents the issue as follows:

> What is body? It is, say philosophers, that which is extended, solid, and divisible. Says the querist, I do not ask what the properties of body are, but what is the thing

[186] Dennett, 1991, 37-8
[187] Churchland, 1988, p.19
[188] BonJour, 2010, p. 21, emphasis original

> itself? Let me first know directly what body is, and then consider its properties. To this demand I am afraid the querist will meet with no satisfactory answer; because our notion of body is not direct, but relative to its qualities. We know that it is something extended, solid, and divisible, and we know no more.
>
> Again, if it should be asked, What is mind? It is that which thinks. I ask not what it does, or what its operations are, but what it is? To this I can find no answer; our notion of mind being not direct, but relative to its operations, as our notion of body is relative to its qualities.[189]

Here Reid points out an important fact that our grasp of the nature of substances amounts to a proper grasp of a substance's characteristic features and abilities. To understand the relation between the mind and the rest of the world, including the bodies they inhabit, we will need something more than an insistence on accounting for "the thing itself". Rather, what we need to do is get as clear as we can on the characteristic features of the various types of substances and analyze our intuitions about the co-instantiability of the various features across categories.

When we start to look at the intuitions drawn out above in light of this, the connection to substance dualism and the full significance of these intuitions can be brought to light. To see it, it will help to start with an objection to the use of the intuitions we have been examining offered by Peter Van Inwagen, who writes

> For it is thinking itself that is the source of the mystery of a thinking physical thing. The notion of a non-physical thing that thinks is, I would argue, equally mysterious. How any sort of thing could think is a mystery. It is just that the thing that does the thinking is physical, for we can form mental images of the

[189] Reid, 1969, 7-8

> operations of a physical thing and we can see that the physical interactions represented in these images--the only interactions that can be represented in these images--have no connection with thought or sensation, or none we are able to imagine, conceive or articulate. The only reason we do not readily find the notion of a non-physical thing that thinks equally mysterious is that we have no clear procedure for forming mental images of non-physical things.[190]

Here, van Inwagen claims that the failure to capture how it is that mental states are physical lies in the mysterious nature of the states themselves, not in a special mystery of combining mental states with physical substances. The complaint seems to be that making minds non-physical doesn't help us understand how minds have the relevant properties; it only seems to do so because we are freed from having to try to introduce thoughts or feelings into our mental images of the substance in question because in the immaterial case we have no such mental images.

However, I don't think this can really be the explanation for why we have difficulty understanding the idea of physical substances thinking while not having difficulty conceiving of thinking substances thinking. To see why, consider the intelligibility of the idea that numbers or properties have thoughts. Taking these as abstract objects, we can follow the tradition of conceiving of them as non-physical features of reality. These abstracta, however, have no more ability to be sensibly attributed mental states than do physical substances. The number seven can't enjoy being prime, or believe it's greater than three, or wish it ate nine any more than it can weigh six pounds, or be twelve feet long, or travel to Detroit. Merely being non-physical isn't sufficient for permitting us to grasp, or trick ourselves into imagining that we can understand, the mental attributions we seem to have no difficulty assigning to God, angels, or demons, even as

---

[190] Peter van Inwagen, 2002. p. 176

we understand these things to be non-physical. This shows that we must have some conception of the nature of minds themselves that allows us to attribute mental states to them, but not to abstracta, and not, as we have seen above, to physical concreta.

I think the above considerations demonstrate something fundamentally important about our ordinary conceptual framework. It's not just that we have some mildly interesting intuitions about the incompatibility of minds and bodies. Rather, built in to our ordinary framework for understanding the world, there exists a set of distinctions between types of substances that are revealed in our intuitions about the impossibility of co-instantiation of their characteristic properties. Our sense of minds as a fundamentally different type of thing from physical objects and from abstracta *consists in* our ingrained sense that the attribution of the characteristic features of the mental to objects belonging to the other realms is a conceptual mistake. Features such as thoughts, sensations, and emotions can't be properly predicated of physical matter or of abstract objects. Furthermore, these boundaries of the non-co-instantiability of certain characteristic features constitutes our understanding of the nature of substance types. Given this, asking for an additional characterization of the substance over and above an identification of these characteristic features will do nothing more than lead us down that long and fruitless road of searching for the "thing itself" when the thing itself is sitting there in plain view. We naturally divide the world into at least the categories of physical objects, minds, and the abstract. Because of this, property dualist views and physicalist views of the mind are both similarly mistaken in their efforts to account for the nature of mental properties as features of a material substance. Both violate rules on the possibility of co-instantiability of the relevant properties by predicating them of substances that cannot possess them. Having this as the basis of such distinctions shows an ingrained, pre-theoretical commitment to substance dualism as a basic

conceptual means of carving the world at the joints. Such commitments are hallmarks of key components of our common sense view of the world.

5.5 An Epistemological Problem for Functionalism

As we saw, problems for physicalism grow out of arguments that draw our attention to the apparently compositional features of our minds. When we attend to these features we recognize them as either the aspects of our subjective experiences that are featured in introspection or as the objects of understanding when we engage in thought and discourse. Functionalists have been attempting to show that these apparently compositional features can actually be diffused into the functional characteristics of the mind. The natural response of most physicalists to my insistence that we place great weight on these intuitions is to try to downplay their significance or force. Here, I present an argument showing that such attempts create a serious epistemological problem for functionalists in accounting for our knowledge of the functions of these states in the first place. The best account of how we learn these functions will include the view that we recognize specific mental states fulfilling these functions, and that we make use of these intrinsic features to track them and thereby learn these functions.

Functionalism seeks to strip the mind down to its powers described abstractly, and leave it open as to what satisfies that abstract description. In this way, they can keep a central aspect of what it is we know about ourselves, but an aspect that should be amenable to physical reduction. The functionalist will agree with the Cartesian that the mind's activities are knowable to us, and that the roles of our mental states in those activities are familiar and can be understood. If we don't have a good sense of the causal role or the other functions being fulfilled by a particular mental state then learning about the apparent causal relevance of any particular physical state can't create a basis for thinking that the physical state is realizing the appropriate

causal role of that or of any specific mental state in any instance. Therefore, if we don't have adequate means of securing a fairly robust understanding of the functional nature of the mental states to be reduced, then we will lack an adequate basis for accepting any particular offering of a reductive base since we won't be able to see that it's adequately fulfilling that well-defined role. In other words, since the goal of the functionalist is to reduce the mind to these functional roles, we need to have firm knowledge what happens in our minds in order to know what states play the relevant roles. The hope is that if all we need for the mental is that these roles should be satisfied, then we can claim ignorance as to their nature and leave it an empirical question what the nature of the states satisfying those descriptions are.[191] And it is thought that neuroscience is answering that question clearly in favor of materialism.

However, there is a serious puzzle about the coherence and tenability of this approach. It seems that we learn how the mind operates by observing it operate in certain ways. And we appear to have this ability because we can track specific mental states as they fulfill these roles during introspection. But our introspective awareness of these states seems to be grounded in our recognition of the relevant compositional features that we have discussed above. Without an ability to detect and recognize these features, it is unclear how we could learn about the functional roles in the first place, and therefore it is unclear how we could ever evaluate any proposed reduction. But if we can recognize and follow such features, then the effort to undermine the importance and legitimacy of such introspective recognition doesn't seem like a viable strategy. That is, the main problem I have with the epistemic position that functionalism finds itself in that it is unclear how we can have an ability to track and individuate our mental states to learn about the functional roles in the first place if we don't do so by introspectively

[191] For a clear articulation of these points see Armstrong, 1980

recognizing intrinsic features of those states and thereby see those states playing the relevant roles.

Learning a functional role requires learning the pattern of the relations between mental states themselves and between mental states, sensory inputs, and behavior. There seem to be two general ways to learn about such relations. The first way is *a priori*. One could consider various patterns and realize that it is possible for them to be realized by something or other. But this does not appear to be how we learn about these functional roles. Since we know these are fulfilled in ourselves most directly, we seem to learn about them through introspection. The other way to learn about such a pattern is through observing various features of the world standing in the appropriate relations, and abstracting from this to the relevant pattern involved in the functional role. In this case, we need to be able to introspectively observe which relations are holding between mental states in order to be able to learn the causal profiles that the functionalist seeks to reduce all mentality to. But we can't just be aware of the relations all the way down. We need some intrinsic features to latch onto in order to be aware that the relations are present. That is, we need to be able to recognize intrinsic features of our mental lives and carefully individuate them in virtue of *something* if we are to learn what relations they bear to one another. But according to the functionalist, mental states are nothing more than the relations that they bear to one another; it's relations all the way down, (or as it were, all the way up). But pure relations aren't proper objects of awareness. One cannot observe a pure relation, one can only observe it by witnessing things standing in that relation. Therefore, we need some qualitative feature to latch onto in order to recognize that the relation is occurring.

By attempting to reduce all of our mental states to functional roles, the functionalist seems to eliminate the possibility of having the qualitative features that are needed in order to the

learn the relations involved in the functional profiles of our mental states in the first place. This difficulty is particularly serious when it comes to relations that are purely internal to the mind. When the mind is related to our behavior, or to the effects brought on by the environment around us, we at least have features of one of the relata to latch onto. But purely internal mental relations seem to utterly lack the characteristics needed to track the relations involved in the functional roles in question. Much like the epiphenomenalist in facing the self-stultification problem, the functionalist seems to have removed our ability to learn about the nature of the most central aspect of our mental lives on their view. The above argument can be formalized as follows:

1. An understanding of the nature of a pattern, n, (such as a functional role) can be obtained only through either *a priori* reasoning that reveals n as a possible pattern, or through observation of particular entities following n.

2. If functionalism is true, the patterns involved in functional role descriptions of mental states are known at least to an extent sufficient to allow functionalism a hearing as an account of those states. Such patterns, however, are not knowable *a proiri*.

3. So the patterns involved in the functional roles of mental states must be known by observation of particular entities following those patterns. (from 1,2)

4. Some functional roles had by mental states are purely internal to the mind (they are not directly and exclusively related to observable behavior, but also or even only to one another).

5. To observe these internal functional roles, we would need to observe entities within the mind fulfilling these functional roles. (from 3,4)

6. Such entities must be locatable and identifiable in virtue of some characteristics that extend beyond their functional characteristics in order to be observed as entities that

satisfy those functional roles (we can't see entities fulfill a role if we can't see the entities at all, and we can't see the entities at all unless we see them as exhibiting some intrinsic characteristics).

7. Therefore, a functional characterization of the mental must leave out some characteristics of our mental lives (namely those that allow us to track the entities that fulfill those roles in order to understand the roles in the first place). (from 6)

8. Therefore, functionalism is false. (from 7)

If this argument is successful, then we cannot be functionalists. But functionalism seems vital to any serious reductive project. We reduce one state to another by discovering the functional profile of the state and ascertaining what lower-level phenomenon realizes the relevant role in the functional profile. Even if we cannot have strict deductive connections that allow us to tie these states to one another as the non-reductive physicalist claims, we must at least have a sufficient grasp of these roles for it to be reasonable for us to think the ontological identification is plausible.

The above argument seems sound to me, but I am somewhat hesitant to fully endorse any argument that has this structure. The reason for this is that it depends upon a dichotomy of available options for learning about functional roles. With such arguments, there is always a fear that the dichotomy is false. In addition, I suppose, unlikely as it seems to me, the functionalist may be able to come up with some combination of relations that could explain why we are able to witness these pure mental relations or they could develop a reasonable *a priori* method for learning this states. For my own part, I don't see a promising avenue of reply along these lines. Either way, however, I hope it is clear that there is a serious epistemic worry about the coherence of functionalism which does not befall the interactive dualist, who is capable of admitting our

knowledge of these roles through our awareness of our immaterial mental states standing in the relevant relations. This adds to the increasingly difficult task the functionalist is trying to fulfill.

5.6: Concluding Remarks on this Chapter

There has been a lot more attention paid to our intuitions of distinctness about things like qualia and content since Chalmers published *The Conscious Mind* in 1996. In the time leading up to it, however, the framing of the issue shifted drastically. Initially, it was thought that dualism was the default view, and that the apparent conceptual commitments to it had to be overcome or explained away in order to make physicalism plausible. It is a peculiarity of philosophy that the road to overcome it has lasted so long, and so many have tread upon it, that it seems to have been forgotten that the burden of proof was on the physicalist. Somehow, without ever discharging their initial burden, physicalism has come to be the default view in philosophy of mind. Indeed, the presumption given to physicalism is currently so strong that even those who have given up on the possibility of ever discharging it still feel comfortable affirming their physicalist credentials.[192] This strikes me as a huge mistake. Initially, physicalists knew that if it came down to it, and mental states really were both genuine states of the world and had a compositional nature, then the game was over. That's why they started down that long and winding road. But the fact that everyone followed them down it (and cultivated a derisive attitude toward dualism in the process) means nothing if at the end of they day going down it can't accomplish the requisite task. Reframing the issues in philosophy of mind from within a common sense approach brings this into stark relief by reshaping the significance of our apparent conceptual commitments.

[192] See, e.g., Colin McGinn, 1989

The goal of this chapter has been to survey how the data about compositional features of the mind appears from within a common sense perspective. From within this approach, the data is seen to speak clearly and consistently in favor of dualism. It appears to include a natural acceptance of the existence of simple, non-physical features of our minds, and a natural tendency to categorize substances in a way that divides the mental and the physical and resists the acceptance of cross-category predication of the characteristic features of each. I have also tried to show how several serious challenges face physicalists in their efforts to account for these features of the mind. Physicalists certainly can and have made efforts to overcome these problems and that try to make sense of these features. However, it is worth noting that given the approach of this work, even if they become far more successful, they will be a long way from justifying physicalism. The fact that the initial data supports dualism requires physicalists not only to provide a reasonable alternative, but to also justify the claim that we should switch to it once it has been developed. Even if an alternative is available, there is no reason to adopt it unless there is something seriously wrong with the most natural and direct view that appears to follow from the data. If the physicalist can be cut off at the pass by showing that there is insufficient reason to prefer such an alternative even if it was, quite surprisingly, eventually developed, then one can bypass the need to examine such alternatives in detail to be justified in maintaining a commitment to dualism. In the next few chapters, I will seek to do just this. In chapter 6, I will argue that the functional or causal profile of the mental, which physicalists have traditionally thought was their stronghold in the field is actually a point in favor of dualism. I will argue that the physicalist's account of mental causation is importantly deficient when it comes to agency, and that this creates even more serious problems for their efforts. In chapter 7, I will try to show that there is no internal inconsistency in a dualist metaphysics that would

require moving to an alternative, and that the features of the mind we discuss form a natural and compelling view of the mind as a distinct substance. Finally, in the last two chapters I will argue that the objections to dualism fail to provide good reason to think we should abandon the view once we start with it. Given this, much of the detailed work on examining specific versions of physicalism will be rendered moot in deciding which view is best supported by the available evidence in philosophy of mind.

The explanatory gap looms as large now as ever. The claim that sensations consist in arrangements of matter makes little more sense now than it ever did. The lack of progress in this research program calls for some explanation. The one the dualist has offered from the outset is that we can't make sense of the claim that these features are constituted out of arrangements of matter because it is a conceptual impossibility that they could be; that the intuition of distinctness isn't grounded in an inadequate grasp of our mental states that can be done away with through further information, it is grounded in an adequate grasp of our mental states that includes the information that they are distinct from any physical state. If this is right, then the physicalist isn't asking us to adapt our mental concepts to new information that reveals surprising features of what our minds are like, they are asking us to abandon our everyday understanding of the mental for one that is more palatable to scientific discovery. Furthermore, on the most popular and promising versions of physicalism, it is not even clear how we could ever come to know about these states in the first place, let alone understand them in a way that would produce such strong intuitions about their nature. In the absence of compelling reason to expect great future progress in these endeavors, the initial, straightforward proposal of the dualist certainly seems reasonable.

Chapter 6: The Causal Argument Against Physicalism

"It is not more evident that mankind have a conviction of the existence of a material world, than that they have the conviction of some degree of power in themselves, and in others; every one over his own actions, and the determination of his will: A conviction so early, so general, and so interwoven with the whole of human conduct, that it must be the natural effect of our constitution..."

--Thomas Reid, *Essays on the Intellectual Powers of Man*, Essay VI, Ch. 5

Overview of Chapter 6

Recent objections to physicalism have typically focused on qualia and content. Since these are the most salient features of the mind, these objections pose a serious challenge to physicalism. I claimed at the end of the last chapter that in light of these considerations, the dualist account of our intuition of distinctness is the most reasonable opinion to adopt, absent some compelling reason to think physicalists will eventually be able to overcome these problems. Currently, even most physicalists grant that there are still large problems to be faced accounting for the apparent distinctness between the mental and the physical. However, they are far from giving up, and they have, at first glance, what sounds like a compelling reason to think they should continue to try. This reason is that they are convinced that only the physicalist can give a conceptually and empirically adequate account of the mind's causal influence on the physical world. They believe that the causal profile of the mind already is, or somehow must be fulfilled by the brain, and that therefore the two have to turn out to the be the same thing. They think the causal facts are inconsistent with dualism, and that if either the functional or compositional story has to change, we should give up on the compositional one rather than embrace epiphenomenalism. If the physicalist is right about this, we should certainly consider

this revision of our understanding of the mind. However, I think that our understanding of the functional profile of the mind not only hasn't been shown to already be satisfied by the brain, I think it simply cannot be satisfied by it, or by any physical thing. The physicalist faces conceptual difficulties in adequately accounting for the nature of the mind's influence on the physical world that are at least on par with those it faces in accounting for its apparent nature. The mind doesn't interact with the world in the way that physical objects interact with one another, and therefore physicalism's claim to be able to subsume the mind's influence under the umbrella of general physical causation is precisely what is wrong with it.

To see this, we should look at the nature of agent causation. Despite efforts to explain the nature of agent causation in the past[193], it is still widely believed that the view is empty or at least, as Daniel Dennett calls it, "a frankly mysterious doctrine."[194] Here, I provide a clear, positive account of agent causation. In the process, I clarify why an agent causal view is incompatible with standard versions of physicalism. Most physicalists will be quite happy to accept this consequence. However, this attitude is misplaced. The inconsistency of physicalism with agent causation is a much higher cost than physicalists realize, and it is a cost that threatens to infect their most promising efforts to provide an adequate, comprehensive account of the mind. Given these costs, the tenability of physicalism as a version of mental realism is in doubt.

The chapter proceeds as follows. First, I explain why standard versions of physicalism are limited to event-causal explanations of human behavior, and how this conflicts with libertarian views about free will and moral responsibility. Next, I provide a positive account of agent causation that captures the important distinctions the libertarian seeks to draw between human behavior and events in the rest of the world. Then I examine some recent empirical data,

[193] See. e.g. Reid (1969), Chisholm (1964), O'Connor (2000)
[194] Dennett (2003) p. 100. See also van Inwagen (1983), Nagel (1979) for similar claims about the mysteriousness, inexplicability, or vacuousness of agent causation.

and conclude that, in conjunction with the intuitive appeal of the account of agency offered in the previous section, we have good reason to believe that folk psychology endorses an agent-causal view of the mind's influence on the world. Finally, I look at the impact of this conclusion on the current projects physicalists are engaged in. Specifically, I argue that in addition to creating another high cost to physicalism by further separating it from common sense, the understanding of agency offered here seriously undermines the legitimacy of functionalist efforts to account for things like qualia and content. Given these combined problems, we have good reason to believe that the only tenable versions of physicalism will be ones that are willing abandon mental realism. Since few physicalists are willing to do this, this creates a serious problem for their view.

6.1: Physicalism and Causal Explanation

To begin, we should examine the sorts of causal explanations available to physicalists. We use the word "cause" in a variety of ways, but for our purposes I am interested only in efficient causation. An efficient causal explanation will explain where the power to bring about an effect comes from. For the physicalist, efficient causal explanations (just "causal explanations" from now on) are event-causal explanations. If all there is in the world are different arrangements of inert matter, then if you want to explain one of those arrangements, there is nothing to appeal to besides the other ones. Since physicalism claims that everything is constituted out of various arrangements of matter, if the view is correct, ultimately it must be facts about how this matter is arranged at one moment that explains how it is arranged at successive ones. Physical causal explanations (other than at the basic level) are generally given in terms of *ceteris peribus* causal generalizations.[195] In such cases, the explanation for why

[195] For a clear explanation of the role of such explanations in reduction see Fodor (1989)

one event occurs is that, other things being equal, an event of that type had to occur once events like the ones preceding it occurred. It may turn out that some events cannot be explained in this manner, and it is open to physicalists to endorse an indeterministic picture of the world according to which causes are not always sufficient for their effects. In this case, however, there is no further explanation the physicalist can give. All efficient causation on a physicalist picture lies in the nature of the preceding events, so if these aren't sufficient for their effects (as many versions of quantum mechanics would have us believe), then the gap in explanation of a true causal claim must arise from some degree of chance.

Whether the physical world always has available to it explanations in terms of the causal sufficiency of prior events is, I take it, an open empirical question among physicists. But if all there is to efficient causes can be exhausted by the nature of previous events, then if such explanations are not available, the remaining gap can only be filled in by chance, and the resulting explanation will be in terms of the statistical likelihood of the effect rather than a strict law explaining its occurrence. These observations about physicalism's causal commitments support the following two claims:

(1) Physicalism requires all true causal explanations to be event causal explanations.

And

(2) All event causal explanations take the form of either the *ceteris peribus* causal sufficiency of the preceding state of the world (CS explanations) or of the statistical likelihood of the effect given that preceding state (SL explanations).

There are straightforward reasons for thinking that such accounts are incompatible with views of the world that accept free will and moral responsibility. If mental states had a causal profile amenable to CS explanations, then we should think that, once they occurred, some effect was

guaranteed. However, most people believe that, even given one's total mental state at a time when she makes a decision, it is still in one's power to do otherwise.[196] As John Searle puts it, we have "an experience of voluntary action where the experience of freedom, that is to say, the experience of the sense of alternative possibilities, is built into the very structure of conscious, voluntary, intentional human behavior."[197] This is not compatible with a causal explanation given in terms of causal sufficiency. While such explanations may apply in some circumstances, these instances appear to be cases of compulsion. In genuine cases of compulsion, the will is insufficient to overcome the force provided by our motives. In these cases, the presence of these mental states is sufficient for the production of their effects. But it is precisely because this is true that we don't assign responsibility to compulsive behavior (assuming, of course, that we don't hold the agent responsible for acquiring the compulsion). What this suggests is that in the ordinary case, where we do hold people responsible, we believe the overall mental state that precedes their action is insufficient, on its own, to produce that behavior. We view the will as an additional source of the power to bring about these effects, and view this contribution of the will as free and independent of the causation involved in the sequence of events that precedes our choice. It is this that lies at the heart of Nagel's sense that "something in the idea of agency is incompatible with actions being events, or people being things"[198]

In addition to the sense that CS explanations eliminate the role of the agent in producing behavior, it is equally clear that genuine responsibility requires more than mere indeterminism can offer. According to an SL account of the profile of our mental states, the presence of the state would permit a set of possible outcomes, one of which would happen to follow. The fact that it would merely "happen to follow" is a serious problem for this view. We don't think that

[196] See Nichols and Knobe, 2007
[197] Searle (1984), p.98
[198] Nagel, (1979), p. 455

mental states just happen to produce one of a set of possible outcomes, we think that we choose among the outcomes for a reason. The libertarian view is not just that the future is open, but that it is subject to the causal contribution of the will. Though he thought we were mistaken to do so, Holbach rightly claimed that we each "[regard our] will as the *premum mobile*, the original motive of his actions."[199] This additional contribution of a choice on the part of the agent, and the rationality of the action that accompanies it, is missing in an SL account of mental causation. Mere indeterminism, or statistical likelihood accounts of our behavior make them chaotic or random in various ways, but they do not make them *ours*.[200] No account available to the physicalist captures this sense of the causally independent, original contribution of the agent to her own behavior emphasized by libertarians. So, there is good reason for any advocate of libertarian views about free will to accept the following claims as well.

(3) Free will is incompatible with mental causal explanations taking the form of CS explanations.

and

(4) Moral responsibility is incompatible with mental causal explanations taking the form of SL explanations.

Subsequently, this leads us to the conclusion that

(5) Agency is incompatible with event causation, and, therefore, with physicalism.

---

[199] Holbach (1770)

[200] Timothy O'Connor (2000) puts the point as follows: "Even though the causal indeterminist account allows for the real possibility of different courses of action, any of which would be 'controlled' by the agent in the minimal sense of being an 'out flowing' of the agent, it's not 'up to the agent,' something he 'has a choice about,' just which potential cause will be efficacious in any given instance and so which action will actually occur. It is, rather, a matter of its falling under a statistical or quasi-statistical tendency that governs the general pattern of behavior in types of circumstances over time, and this probabilistic tendency is not something the agent has any choice about." -p.29

Although arguments like this one for agent causation are not entirely new,[201] a pressing problem for libertarians has been their perceived inability to provide a reasonable account of how mental causation works that doesn't suffer the same fate. A. J. Ayer, for example, thinks this incompatibility shows that there is something wrong with libertarianism. As he puts it

> But now we must ask how it is that I come to make my choice. Either it is an accident that I choose to act as I do or it is not. If it is an accident, then it is merely a matter of chance that I did not choose otherwise; and if it is merely a matter of chance…it is surely irrational to hold me morally responsible for choosing as I did. But if it not an accident that I choose to do one thing rather than another, then presumably there is some causal explanation of my choice: and in that case we are lead back to determinism.[202]

As Ayer and many others see it, there is no escape from the event causal view of the world, and so the libertarian cannot get what she wants. In response, many libertarians have claimed that causation in these cases is produced by an agent, and not by any event. However, many people have claimed that their efforts to explain what this means have proven empty. I think this is an unfair criticism of the work of agent causation theorists, but either way I will seek to disprove this contention by providing a clear, positive account of the nature of agency. This account should, in turn, provide further support for premises (3) and (4) in the earlier argument.

6.2: Three Distinctions Between Agent and Event Causation

Perhaps the main reason people reject agent causation is that they do not understand what it is. Even those sympathetic to it express this concern. David Lund, for example, says "it must

---

[201] This argument, for example, shares many similarities with Chisholm's argument for libertarianism in his (1964)

[202] Ayer, (1969)

be acknowledged that the concept of agent causation is far from clear. Perhaps it cannot be rendered sufficiently clear and distinct to dispel entirely the sense of mystery that envelops it."[203] However, I believe this is inaccurate. There are at least three clear distinctions between causation in instances of agency[204] and in the ordinary case. By drawing these distinctions, I hope to dispel the mystery and provide clear content to the notion of agent causation. Before explaining these features it will be informative to examine the sorts of causal explanations we pursue, and the point at which we typically find an efficient causal explanation satisfactory. In the case of event causation, we typically aren't fully satisfied with a causal explanation unless we can discover some general causal fact of which our particular case is an instance. Sometimes, these generalizations may be statistical in nature (ie: events of type A can bring about ones of types B, C, or D, with the following propensities...), but we think that there is bound to be some covering law that explains the particular case. In the most basic terms, we think event causal explanations are successful when they explain particular cases by showing that the effect occurred because, well, "that's the way the world works." This law of nature approach to understanding the causal efficacy of particular events has been wildly successful over the past few hundred years, and has served as the cornerstone of vast reaches of scientific discovery.

These amazing successes have led many people to think that we should extend such an approach to our understanding of human behavior. It has lead psychologists to pursue general facts about how people work that can be used to account for particular instances of human behavior. This approach to grasping the nature of agency, however, is actually quite unnatural. The law of nature approach to causal explanation serves primarily as a replacement for a

[203] Lund, (2005), p. 375

[204] For a good discussion of various attempts to account for a libertarian version of free will within an event-causal picture, and the difficulties that such accounts face, see O'Connor, (2000), esp. Chapter 2. I generally find myself in agreement with O'Connor's criticisms of these alternative pictures.

fundamentally different type of causal explanation: the sort that appeals to agency. Seeking, then, to explain agency itself in these terms is a surprising extension of the approach. The demon-haunted world of primitive religion did fall away to scientific progress. One important point that seems to have been lost in this lesson, however, is that there must have been some reason we thought the world was demon-haunted in the first place. The fact that one sort of causal explanation gave way to another in our empirical investigations seems to require that we possess an understanding of two fundamentally different types of causation.[205]

When we can find no clear power in the events that preceded it for bringing about an effect, people have a natural inclination to posit spirits in the world to provide an explanation. At this point, however, no further explanation of the power to produce the effect seems necessary.[206] This demonstrates that we conceive of agents as sources of power in themselves. We do not conceive of the effects of agency, whatever they may be, as requiring any general structural fact about how the world works to underlie or explain the power in the cause to bring about the effect. We conceive of agents as capable of producing effects of their own accord, and therefore the introduction of a sufficiently powerful agent can serve, in and of itself, to account for the efficacy involved in producing an effect.

This fact about how people conceive of agents helps explain many facts about the history of belief. It makes it comprehensible not only why, but more fundamentally how it could have occurred to people to think that there were spirits pervading our world. One example of this is found in the longstanding appeal of the cosmological arguments for God's existence. Surely any

[205] This fact seems to go pretty far toward pre-futing the claim that agent causation is incomprehensible. The fact that a style of causal explanation is naturally and frequently appealed to must give it some presumption of intelligibility, even if we haven't yet fully articulated its nature.

[206] The word "further" here is important. We may think that an alternative explanation may be unapparent yet superior, but if we accept the agent-causal one, the question of how it was *possible* that it could occur is no longer present.

first event must be among those that cannot be explained by the nature of a preceding one, and so, under the suggestion that there was such an event, we will appeal to that other source of causal power we know of: the agent (and what an agent it would take to bring about an effect like that). Another example can be seen in the events that inspired the Japanese belief in "the Divine Wind." Several cultures have personified nature, and seen its acts as representative of the will of the gods in their lives, but few have as compelling of a story to tell as the Japanese do in their personification of the causes of typhoons that have saved their civilization from conquest on multiple occasions. Two Mongol invasions, either of which would likely have led to conquest were ended by typhoons occurring in the same bay seven years apart, once outside typhoon season, and in each case in an area that otherwise isn't know for frequent typhoon activity. The fact that the people of Japan saw in this divine intent is hardly difficult to understand, but reflects again the natural tendency to explain apparent intent by the existence of a spirit or mind. Perhaps my favorite example can be found in the beliefs of the Trobriand Islanders. The staple diet of these people includes the consumption of large amounts of a particular yam. This yam produces similar effects to certain birth control methods. As a result, the people who live here never had regular experience of the conjunction of sex with pregnancy. This left room for speculation about the potential causes of pregnancy among the islanders. In the absence of the natural explanatory link found in every other society, the people of this culture formed the belief that pregnancy was caused by spirits carrying embryos into women.[207] Even for such routine and natural events as pregnancy, the absence of a clear event-causal explanation quickly lead to the positing of spirits as a cause. This history of positing spirits as causes makes sense only because we conceive of the effects of agency, and of the nature of agent causation, in a fundamentally

[207] This was initially discovered by Malinowski and discussed in his (1929). For those inclined to think these beliefs particularly naïve or otherwise unrepresentative of human thought it is worth remembering that most of Western Civilization has believed that this type of thing happened at least once in human history, around 2,000 years ago.

different way than we conceive of ordinary event-causal processes. Agents have power in themselves to bring about effects, to initiate causal process, to impose their will upon the course of nature. The effects they produce are not to be explained by general facts about the ways of agents, but rather by the particular reasons and motives of the agent in question.

I hope this discussion makes it clear that people are naturally possessed of two fundamentally distinct notions of efficient causal power. It will be helpful, though, to be more specific about exactly where the differences lie. There are at least three clear and important differences between these types of causation. First, while event causation involves the occurrence of particular instances of general causal facts about how the world works, agent causation is singular, and explained in each case by a particular exercise of the power of an agent. Second, when an event serves as a cause, its power to produce its effect is derivative of some more general causal fact, whereas when an agent serves as a cause, the power is original in the agent herself. Finally, event causation is conceived of as blind, while agent causation is conceived of as deliberate, creative, and intentional. Though there may be some grand designer behind general facts about causation, particular instances of event causation are not thought of as serving particular aims, or satisfying any specific intention. Agents, however, act for particular reasons. We create our behavior to suit our needs; it does not flow from a blind, undirected process. We shall examine each of these distinctions in turn.

Our first distinction is between the generalizability of event causal facts and the singular nature of agent causation. To help see this distinction, let us consider a pair of examples. Suppose you walked outside and felt a very strong wind blowing against you, perhaps almost knocking you down. Suppose you then looked across the street and saw that the branches on the trees weren't moving. What would you think? Well, you might think a very strong fan was

blowing on you, or maybe that there was something very weird about the tree. If you looked around and couldn't find any unusual source of the wind, and knew that the tree was normal, you might very well begin to doubt your own senses. You may think you were dreaming or hallucinating. You might think someone was playing a trick on you. What I imagine you would never think, however, is that, just this once, the fact that wind moves the branches of a tree simply failed to hold. The general causal fact about the effects of wind on trees is so ingrained in us that we would go so far as to doubt our own eyes before we would question the generality of this relation between cause and effect. This seems to be part of a more basic fact about how we understand the relationship between physical events in the world: When one type of event stands as an efficient cause of another, this is grounded in some general, structural fact about how nature works. The pursuit of laws of nature as a source of the explanation of particular causal facts seems ingrained in our ordinary sense of the nature of event causation.[208]

This generalizability, however, does not extend to human action. To see this, consider a different case. Suppose you were to invite someone into a room and ask her to pick a number between 1 and 100. Suppose you were to do this a number of times, perhaps wiping her memory each time to ensure that the situation was as similar as possible in each instance. Now, suppose she were to pick 24, 62, 97, 12, and 73. I don't think we would find anything odd about the selection of this sequence, or, really, of much of any random sequence. Indeed, it would take some very regular pattern to feel compelled to pursue any further explanation as to the origin of the sequence. Even given the same inputs in terms of events, when it comes to the behavior of agents, we don't think general causal facts will explain what we do. Rather, we think that in each instance, the person simply made a choice and acted on it. This singular nature of the

[208] Thomas Reid articulates this natural tendency as follows: "all natural philosophy grows…[from] an eager desire to find out connections in things, and a natural, original, and unaccountable propensity to believe, that the connections which we have observed in time past, will continue in time to come." Reid, 1764 (1997), p. 41.

causes of human behavior is a striking difference from how ordinary events occur. In the case of physical events, it doesn't even enter our minds to question the generality of causal explanation, in the case of agents, it doesn't typically enter our minds to look beyond the particular agent's choice to understand her behavior.

The next, closely related difference between these types of causation has to do with the source of the power in the cause to bring about its effect. In the case of an event, the explanation for why it has the power to produce its effect lies in the general structure of physical reality. In particular, it is general facts about the causal profiles of various features involved in the event that provide it with the power to produce an effect. When, for example, a baseball breaks a window, general facts about objects with the baseball's mass, traveling at the relevant speed, combined with facts about the structure of glass, serve to guarantee that the window will shatter. This fact contrasts sharply with the nature of agency. As we saw above, it is precisely when the preceding events seem to lack features that can provide them with the power to produce their effects that we find it necessary to appeal to agents. The agents work as independent sources of power. This notion of agents as original sources of causal power is expressed by Reid as follows:

> ...a free action is an effect produced by a being who had power and will to produce it; therefore it is not an effect without a cause.
>
> To suppose any other cause necessary to the production of an effect, than a being who had the power and the will to produce it, is a contradiction; for it is to suppose that being to have power to produce the effect, and not to have power to produce it.[209]

[209] Thomas Reid, 1813 (1969), p. 332

Indeed, it is not so much determinism that is the enemy of freedom, as the application of a style of explanation that seeks to draw this power over our own behavior out of the agent, and back into the general structure of nature. Unlike in event causation, where events draw their power from general facts about how the world works, the power is already there in the agent. In cases of agency, the relevant properties don't serve as sources of power, they become relevant to the causal story when the agent uses them as reasons for exercising her power in one manner rather than another. This leads us into our final distinction between event and agent causation.

The final difference we will look at between the two types of causation is that event causation is blind, while agent causation is intentional. It is here where the relevance of these distinctions to the causal profiles of our mental states becomes most clear. When a feature of an event is relevant to producing the effect, the relevance is explained merely by the presence of the feature. Events aren't guided in selecting their outcomes by the properties they have within them; they blindly produce their effects merely in virtue of the presence of those properties. In the case of our mental states, however, the exact opposite is true. Mental states become causally relevant by being taken by the agent as a reason for action. The general structure of how the world works mandates nothing about the future in virtue of the mere presence of a mental state. Rather, we take them as reasons to choose one future rather than another. This choice needn't occur in the moment. We can structure future responses to occur automatically in the presence of a certain mental state. How we choose to do so, however, is extremely malleable. We can, if we choose, seek to cultivate the virtue of charity by, say, structuring our responses to an awareness of need by making it serve as an immediate reason to provide for that need to the best of our ability, and by structuring our responses to our feelings of sympathy to drive us toward unreflective kindness. We could also, however, choose to take these same states of awareness

and emotion as an immediate basis for avoidance or learned callousness. Their presence requires neither; their relevance and the responsibility for how they influence our behavior lies in how we choose to make use of them. The presence of our mental states gives us a sense of what options exist, and provides motivation for adopting one over another. Their mere presence, however, guarantees nothing about the effect that follows. This is a key difference between the roles that mental properties play in producing our behavior and the roles that physical properties play in explaining the power of an event to produce its effect. In an event, the mere presence of certain properties mandates the passive production of an effect. In agents, our mental states motivate us by providing us with reasons to actively produce certain behaviors. However, not only do mental states fail to produce behavior without the additional contribution of the agent, but which action the mental state provides a reason for producing is subject to an agent's evaluation of what its relevance should be. For this reason, mental states have no unique profile in terms of potential outcomes, and depend for their causal significance on their ability to influence how an agent directs the power that lies within herself.

We can see now that an agent-causal account of the role that our mental states play in our lives is not one that is amenable to event-causal explanations. Agent causation includes a view of the role of reasons that precludes them from being efficient causes of our behavior in any case where we act responsibly. It also includes a view of them according to which, although an agent may act on those reasons, she is still an independent causal source of her own behavior. As we saw, not only are we able to decide which features will be relevant to the behavior we choose to create, we can even determine how the features will be relevant. The causal profiles of our mental states aren't writ large in the structure of the world, they are plastic, and can be created and altered by the will of the agent who has them. Neither of these facts is compatible with a

story of causation according to which the mere presence of these features in the preceding state of the world is all there is to the explanation for why the effect occurs. If mental states were amenable to CS explanation, then the agent couldn't freely decide whether or not they would be relevant to what happened when they were present as motives. If they were amenable to SL explanations, then there would nothing more to the causal story after explaining what futures are compatible with their presence, and seeing which one actually occurred. But not only is there more to the story in terms of the contribution of the active power of the agent, but the very facts about which futures they will motivate us to bring about are flexible, and significantly within our control. The account of the role of mental states in agent causation, then, is not amenable to the types of causal profiles available to physicalists.

6.3: The Experimental Data

Those who give strong weight to intuition should think that the intuitively compelling account of the agent causation provided here creates a serious difficulty for physicalists. Certainly, the common appeal of such a view of causation should influence the views of common sense philosophers. However, I do not need to rely solely on such data. Recent empirical work, though not univocal on the issue, offers further support for the contention that people naturally believe that mental causation is fundamentally different from event causation. Psychologist Paul Bloom claims that "common sense tell us that our mental life is the product of an immaterial soul."[210] One of Bloom's primary reasons for thinking that people are natural dualists is precisely this distinction between agents and other sorts of objects in the world. Bloom notes, for example that babies learn to expect agency from people and animals, but are surprised to see it in ordinary physical objects. The capacity of children as young as 12 months old to differentiate between actions that exhibit intentionality and agency from behaviors of objects that

[210] Bloom, (2004), p. 226. It should be noted that Bloom claims this despite rejecting dualism himself.

lack those types of causal impetus has been well documented in recent work by child psychologists.[211] This suggests that people have a natural tendency to distinguish these two types of causation from a very early age.

In addition, certain x-phi experiments deliver similar results. With respect to the question of whether or not the universe is deterministic, Nichols and Knobe discovered that "across conditions, nearly all participants (over 90%) judged that the indeterministic universe is most similar to our own."[212] They go into more detail with respect to moral choices and physical events, reporting that

> In a set of experiments exploring the lay understanding of choice, both children and adults tended to treat moral choices as indeterminist (Nichols 2004a). Participants were presented with cases of moral choice events (e.g., a girl steals a candy bar) and physical events (e.g., a pot of water comes to a boil), and they were asked whether, if everything in the world was the same right up until the event occurred, the event *had to* occur. Both children and adults were more likely to say that the physical event had to occur than that the moral choice event had to occur. This result seems to vindicate the traditional claim that ordinary people in our culture believe that at least some human decisions are not determined.[213]

However, they go on to note that "experimental study has not been so kind to the traditional claim that ordinary people are incompatibilists about responsibility," where results are mixed. They describe the results as follows:

> These initial experiments [resulting in compatibilist intuitions] replicated the finding (originally due to Nahmias et al.) that people have compatibilist intuitions

[211] See Bloom, (2005), Chapter 1, esp. pp. 14-19
[212] Nichols and Knobe (2007, p. 669)
[213] Ibid, p. 667

> when presented with vignettes that trigger affective responses. But they also yielded a new and surprising result. When subjects were presented with an abstract vignette, they had predominantly *incompatibilist* intuitions. This pattern of results suggests that affect is playing a key role in generating people's compatibilist intuitions.[214]

Nichols and Knobe are inclined to think that the incompatibilist intuitions are more reliable indicators of common sense belief in this area, because they are less likely to be negatively affected by affective response. I won't replicate their arguments here, but will point out that they seem correct to suggest that

> It appears that people have *both* compatibilist *and* incompatibilist intuitions. Moreover, it appears that these different kinds of intuitions are generated by different kinds of psychological processes. To assess the importance of this finding for the debate over moral responsibility, one would have to know precisely what sort of psychological process produced each type of intuition and how much weight to accord to the output of each sort of process.[215]

In light of this, we should look at which set of intuitions are more likely to be reliable. I believe we have good reason to afford the abstract, incompatibilist intuitions much greater weight than the compatibilist ones offered in concrete cases.

To see why this is, we should look at some recent work in cognitive psychology from Daniel Kahneman. Kahneman contends that many experiments intended to elicit intuitions are unreliable indicators of belief due to a problem he calls "substitution." This problem occurs when respondents are incapable of providing an immediate response to the question asked, and

[214] Ibid, 671
[215] ibid, 681

therefore substitute it for a different, related question that is easier to answer. Kahneman discovered this problem in the area of probability, where people were shown to be very unreliable at providing accurate answers to even simple problems. Kahneman claims that in such cases, because the task requires engaging in reasoning rather than accessing an immediate response, "when called upon to judge probability, people actually judge something else and believe they have judged probability," and goes on to say that people "often [make] this move when faced with difficult target questions, if the answer to a related and easier heuristic question comes readily to mind."[216] Because of this, there are limits to what sorts of questions one can ask in an effort to reveal people's intuitions about the target question itself. If we ask a question that demands cognitive labor, we will often end up with an answer to a separate question that requires less work to answer. Such an answer will not provide an indication of someone's intuitions about the question at hand. Even worse, if the heuristic substitute is a subjective question about one's attitudes, it likely won't reflect an intuition at all. Instead, it will simply report one's introspective awareness of how they feel about a related issue.[217] This gives us reason to be skeptical of x-phi experiments whose target questions require the subject to engage cognitive labor to answer, since the results may not reflect an intuition about that question, and may not reflect any intuition at all.

The next thing to notice is that moral utterances serve multiple purposes. In addition to expressing evaluative judgments, moral utterances also serve expressive and prescriptive functions. Given these functions of moral speech, it would make sense for us to discourage or

[216] Kahneman, (2011), p. 100. I think this fact actually provides us with good reason to take philosopher's assertions about common sense belief as often at least as reliable as experimental data obtained from such experiments about intuitions. Since often what is actually stated doesn't reflect an intuition, and philosophical questions are often difficult to answer unthinkingly, there is good reason to be skeptical of the value of most studies of people's intuitions about philosophical questions.

[217] Most of the examples Kahneman gives in his (2011) follow this pattern of substituting a subjective question for an objective question. See, esp. p. 101.

express disapproval toward specific actions that we don't want people to engage in, even if, in the particular case, the person isn't responsible for what they did. Therefore, it is to be expected that in order to fulfill these non-epistemic, yet still vitally important aims, we will often disapprove of actions that people aren't morally responsible for. The desire not to fail as sources of moral instruction by making it appear as if those features are acceptable, and the desire not to seem unsympathetic by failing to express the proper emotions, each create a cognitively difficult situation when people are asked to offer immediate responses to concrete moral situations, at least if these purposes diverge from moral judgment. The complex function of ethical utterances is more likely to create problems in more concrete cases, since aspects of the cases are likely to elicit strong desires to express emotional responses, or to express disapproval of certain features of the event.

The fact that answers to concrete moral questions can serve multiple, potentially incompatible purposes creates a serious challenge to the legitimacy of empirical tests of responses to those questions as indicators of moral judgments. Questions that require cognitive labor typically result in substitutions for questions that can be more easily answered. Questions whose answers serve multiple, potentially incompatible purposes create ambiguity about how to answer, and therefore create a need for cognitive labor to ensure an accurate response. Since, particularly for questions of moral responsibility, two easily substituted questions are "are you displeased with this situation?", calling for an expressive role in your moral response, and "should we encourage people to act like this?", calling for a prescriptive role to be most prevalent in your response, substitution problems in this context are likely to lead to statements that fail to offer a declarative moral judgment at all, instead serving as efforts to express emotions or exhort others to avoid certain activities. Given this problem for experiments testing

our intuitions in concrete moral cases, the value of the data in these cases is highly suspect, and is more subject to error the more concrete or complex the situation is. Judgments about more abstract moral claims, however, are less likely to elicit emotive or prescriptive replies, and so are likely to be more reliable indicators of belief.

Since moral statements offered in concrete cases are often unreliable, while judgments about more abstract cases are more trustworthy, the apparent conflict between people's responses to the two types of cases is best resolved by accepting the judgments made in more abstract cases. Since the more abstract judgments vindicate the view that we are natural libertarians, there is good reason to think that our best empirical studies to date confirm the view that we think of causation very differently in cases of agency. The reason we have a robust sense of agents as sources of the power to freely create their own activity is because this is how we naturally conceive of human behavior. Given the ingrained nature of this view about agency in everyday judgment about human behavior (as expressed in the opening quote from Reid), rejecting it requires rejecting a heavily integrated and natural component of a widespread conception of mental causation. This creates a much higher cost than is typically recognized for those who reject agent causation. The costs, however, do not end there. In addition to the direct costs of rejecting agency, there are indirect costs stemming from how rejecting this aspect of mental causation changes the prospects of the most promising efforts by physicalists to account for other features of the mind, such as qualia and content.

6.4: The Extent of the Problem

If the above is correct, then physicalists cannot capture our intuitive understanding of the nature of agency, and therefore have left out a vital aspect of mental causation from their accounts of the mind. Those who afford folk psychology strong evidential weight should think

this is a serious problem. Many physicalists, though, are likely to bite the bullet here, and think this is so much the worse for the ordinary view of mental causation. However, this response fails to appreciate the importance of this point. Rejecting our common sense understanding of agency carries a much greater cost than physicalists typically realize. It does so both because rejecting agent causation is more problematic than they realize and because doing so undermines the reasons we have to believe that a successful account of the realization base of those states could be found.

To see why the cost of rejecting agent causation is higher than is typically supposed, we should consider the question of where the notion of agent causation could have come from. It is certainly possible to imagine a species with a natural understanding of a type of causation human beings don't have an intuitive conception of. For example, one could imagine a species that naturally detects magnetic attraction between objects. For such a species, the fact that two magnetically charged objects could attract or repel one another wouldn't be as surprising or confusing as it is to human beings who first encounter it. One could further imagine this species misappropriating this causal explanation in a variety of instances. For example, they might think it explains why rain is drawn to the ground. If it was integrated into their mating rituals, they may think that sexual attraction in other species was also caused by instances of this magnetic attraction they couldn't detect. But what one could not plausibly suppose is that a species like this existed *in a universe where no magnetic forces ever existed.* A natural and automatic form of causal explanation ingrained in a species has to come from somewhere. The most obvious source would be particular instances of such causation that they are directly aware of, coupled with a natural inclination to posit such instances in new cases where other causes weren't apparent. Causal conceptions that occur *ex nihilo* are rather bizarre theoretical posits for any species.

Under the supposition that agent causation is an empty view, *ex nihilo* creation of a conception of "some cause I know not what" isn't implausible. But, given the account of agent causation I defend here, and given the fact that it appears to identify a positive, well-defined, and naturally appealed to conception of a distinct type of cause, the supposition that this understanding of agent causation arose *ex nihilo* is no less exotic and implausible than the supposition that a natural understanding of magnetic attraction as a form of causation could arise in a species that exists in a universe where no magnetic forces exist. The rejection of the existence of agents as causes therefore strikes me as rather idle speculation positing implausible origins of a prominently used conception of agency in the human species. This, in turn, strikes me as a large and improper theoretical extravagance, one that Ockham's razor should count heavily against. The far more plausible and natural hypothesis in this instance is that this conception of agency arises automatically in us and reflects instances of direct awareness of exercises of such agency in ourselves, as the Reid quote at the beginning of the chapter claims. Denying the existence of this type of causation therefore incurs a heavy theoretical cost.

In addition to this theoretical cost for physicalists, the rejection of agent causation creates a problem for the project most physicalists are engaged in. Physicalists are trying to convince us that every particular instance of a mental state is realized in a state of the brain. In order to show this, they need to be able to show that the various features of the mind can be realized by the brain, and then argue that this account of their nature is superior to dualist accounts. Although some recasting of our understanding of the mind is understandable in these efforts, there is a point where divergence from ordinary understanding is too great for such a project to be plausible, or worth maintaining. As Michael Huemer points out

> the reason why we believe, for example, that water=H_2O is, roughly, that (i) we have independent (that is, pre-theoretical), direct awareness of the presence of water, and (ii) the theory that this substance of which we are aware is composed of H_2O molecules helps explain many of its observable features (again, features that we know of independently of our scientific theory).[218]

Even if conceptual analysis isn't the end of the story, it does sets limits on legitimate accounts of realization. The reason that we don't say, for example, that epilepsy realizes demon possession is because the nature of what we believe is actually happening is simply too different from what was supposed to be happening in such cases for one to serve as a realization base for the other. To put it another way, there must be some way to differentiate interesting ways of understanding the underlying nature of something from efforts to replace our understanding of it with a concept that better fits one's metaphysics[219]. When we move from one to the other, we exit the realm of realism and enter the realm of error theoretic replacement. For the physicalist's project to succeed in defending a realist account of realization, she must therefore first recognize that we have an independent, pre-theoretical understanding of the nature of our mental states, and then show that the elements of this understanding are best accounted for by the hypothesis that mental states are realized by neural states. Absent this, we will not have a reason to believe that the account is one of the mind rather than an account of what is actually there instead of a mind; that is, rather than a realization account, we will have an error theory.

By now most physicalists have realized that a view of the mind that fails to take into account things like qualia and content are simply too different from folk psychology to be

[218] Huemer, 2005, p. 85

[219] For a good account of how these sorts of limits function in our understanding of works of art, and a good defense of applying this standard in general, see Thomasson (2006).

acceptable.[220] When it comes to the "hard problem" of phenomenal consciousness, however, their efforts to provide an adequate account have been far from successful.[221] The physicalist's promissory note as we wait for a satisfactory account of the salient features of introspection has been justified by their assurance that the mind's causal functions are fulfilled by the brain. That is, it is because the brain is doing the causal work that we are supposed to believe it can do the rest. However, in light of the work in this paper, we have serious reasons to be skeptical of this approach. There are two connected reason for this skepticism.

The first reason the earlier work in this chapter undermines physicalists' efforts is that it shows that their account of the mind's features is far more deficient than commonly realized. When it comes to the efforts of physicalists to explain things like qualia and content, no account has been widely embraced, even amongst physicalists. The longer physicalists fail to provide such an account, the more likely it is that they reason they haven't is because the project isn't viable from the outset. As for the promissory note of causation, the claim that physicalists can adequately account for causation is grounded in a dramatically truncated understanding of the nature of mental causation; one that leaves out our sense of the mind as a locus of free choices and of independent exercises of agency. Conceived of so as to include these features, the claim that the mind's causal influence is realized by the brain is far from established, and this promissory note loses much of its value. Given the inadequacy of the account of causation and the absence of a compelling account of qualia and, arguably, of content, it is unclear how closely the physicalist's conception of the mind adheres to a pre-theoretical view about its nature. Based on their inability to account for the salient features of introspection, and their elimination of our

[220] Of course, not all physicalists recognize this need. See, e.g., Dennett (1988)
[221] Some physicalists have completely given up on this project. Colin McGinn (1989), for example, thinks that it is impossible to understand how such features could be physical, but thinks we should believe they are physical anyway, while others, such as Jaegwon Kim (2005), have decided to settle for something "near enough" to physicalism instead.

sense of agency from the causal activity of the mind, the physicalist may well have crossed the line that separates versions of mental realism from error theories about the mind's true nature. This is more likely to be true if we have little reason to think physicalists will be able to eventually provide a conceptually adequate account of things like qualia and content in light of the fact that it leaves out aspects of our pre-theoretical view about mental causation.

The second reason to worry about the physicalist's prospects is that the inadequacy of their account of mental causation leaves us little reason to think they can use that conception to ground a recognizable account of the nature of qualia and content. One reason for this is that this deficiency changes the content of the project. The standard physicalist claim is that the causal nature of the mind, as we pre-theoretically understand it, is being fulfilled by the brain, and that therefore we will one day be able to see how the rest of its features are realized there as well. This claim provides legitimate reason to think their efforts might be successful. However, the claim that if we accept an event-causal analog of our pre-theoretical understanding of mental causation we will someday be able to see how this account allows us to understand these features is far less compelling. Any prospect for understanding things like qualia and content *in terms of* how the mind functions requires a close conceptual relationship between the functions of the mind and those other features we are trying to understand. Without this connection, I see little reason to think that developing a better understanding of those functions could help us see how the other features could be physical. As it stands, the contention that such a connection exists is far from obvious[222]. What the physicalist is now seen to be suggesting, however, is not that there is a sufficiently cohesive view of the mind's nature to allow us to one day appreciate this truth.

[222] Indeed, some physicalists have already given up hope that this is true. Papineau, for example, thinks that "there are no a priori associations between concepts of conscious states and specifications of causal roles" (2002. 39-40). Papineau instead rests the positive case for physicalism almost entirely on the issue of causation. Given the above concerns about how he conceiving of those causes, however, I don't see how such a limited defense can meet the conceptual burden for endorsing a realist version of physicalism.

Rather, they are claiming that there is a sufficiently clear conceptual connection between these apparently intrinsic features and an *incomplete or substituted account* of their functional nature, and that seeing *this* connection will allow us to understand how those features of the mind can be physical. But whatever reason we might have had to think that pre-theoretical concepts about distinct features of the mind form a sufficiently cohesive unit to allow this reductive effort to succeed, I see no reason at all to suppose that the conjunction of certain pre-theoretical mental concepts together with a replacement theory of mental causation will possess sufficient conceptual cohesion to allow us to see how a functionalist account of the realization base of the mind could be true. Since functionalist views have provided the most promising efforts to account for such features, this poses a serious problem for physicalists in their efforts to convince us that they can solve the problems they currently face.

Concluding Remarks

Physicalists want us to believe that mental occurrences that seem nothing like synaptic firings in squishy gray matter are actually nothing more than that in any particular instance. I believe part of the reason they have been so successful in convincing people that this is possible rests in the fact that so many people believe that mental causation can be better understood within a physicalist framework. The lack of success from dualists in motivating the acceptance of a viable alternative account of mental causation has certainly helped in this regard. By providing a positive, detailed account of agent causation, and by showing that this is an important part of our pre-theoretical or folk psychological view of mental causation, I hope to have shown that there is real competition on this issue, and that physicalists should not feel comfortable resting on causation as a source of strength while we wait for the rest of their project

to unfold. Indeed, I hope to have shown that the deficiency in their account of mental causation seriously undermines the tenability of that project as a whole.

Any effort to establish an adequate realization base for the mind has to start somewhere. If it isn't going to start by explaining the features we seem to know about through introspection, and it isn't going to start by accounting for the mind's causal influence on the world as we understood it prior to those efforts, then it is unclear how the project can get off the ground. Until there are new grounds for thinking such a project can succeed, I see little reason to think physicalists will be able to provide a conceptually adequate account of how the mind is realized by the brain. But without such an account, physicalism lacks adequate fidelity to common sense to qualify as a version of mental realism. The only types of physicalism left on the table, then, would be either error theories that seek to replace our understanding of the nature of the mind with a drastically different conception of its nature or eliminativist views that seek to abandon our belief in the mind altogether. Since physicalists are typically unwilling to abandon mental realism, this is a heavy price for them to pay. Given that physicalism requires such a drastic revision of our understanding of one of the most familiar and accessible features of our world, it's a price I don't think anyone should be willing to pay.

Chapter 7: Dualism as Common Sense

One possible objection to the view that dualism is the common sense view of the self would arise if our pre-theoretical understanding of the mind failed to provide a coherent metaphysical picture, but instead consisted of an aggregate of disparate intuitions. Having described some the relevant features of the mind, I believe we are now in a position to see that this isn't the case. In this chapter, we will briefly examine two remaining features of the mind and show how these features help us tie together the work in the previous chapters into a comprehensive and cohesive view of the mind as an immaterial substance. The goal here isn't to provide full-fledged defenses of the arguments in favor of these views, but rather to show how these features commonly appealed to by dualists as evidence for dualism fit together with the other features we have looked at to form a cohesive whole. This should help assuage the concern that our natural understanding of the mind consists of disparate or inconsistent elements. In turn, this should help us see why dualism is an integral part of common sense, thereby providing it with a strong presumptive position in the field.

7.1 Disembodiment Intuitions

One of the features of the mind we have not examined yet is the sense of the mind as a distinctive part of our selves. One sees this view of the mind throughout societies, and it is an integral component of most views of the afterlife. Whether the soul is supposed to go to eternal paradise or damnation, get a new body through reincarnation or resurrection, or stay in the realm of perfect forms contemplating abstract truths for eternity, most cultures have found it natural to think of the self as something that can survive the death of the body. In addition, popular culture is full of stories where minds travel from one place to another without the body. Movies like

Freaky Friday and The Change-Up are based almost entirely on this idea.[223] Science fiction, fantasy, and cartoons all frequently have spirits possessing people, objects suddenly becoming animated or obtaining personalities, minds being taken over by outside influences, astral projections of the self out of the body, ghosts and angels visiting people or witnessing what is happening in the world, and so on. One thing about such stories is that even the most adamant materialist usually has little trouble understanding what is supposed to be going on in these stories or in these religious views of the afterlife. Conceptually, we typically have no problem with the idea that minds could exist without bodies, or that spirits can produce effects in the world once they are present in a body.

This sense of the self as a separable aspect of a person has been defended by philosophers as well. Rene Descartes is the most famous advocate of this position, arguing for it in his *Meditations*, stating

> Because, on the one hand, I have a clear and distinct idea of myself in so far as I am only a thinking and unextended thing, and because, on the other hand, I have a distinct idea of the body in so far as it is only an extended thing but which does not think, it is certain that I, that is to say my mind, by which I am what I am, is entirely and truly distinct from my body, and may exist without it.[224]

This passage from Descartes captures a recurrent idea among dualist philosophers; that the mind and body are distinct and in principle separable aspects of persons, and that the central aspect of our identity is found in the mind, or soul. One of the first serious defenses of this position came from Plato. In a conversation with his friends the night before his death, Socrates is quoted as saying "Death is, that the body separates from the soul, and the soul, separated from the body,

[223] A full, long list of such stories can be found at http://en.wikipedia.org/wiki/Body_swap
[224] Descartes, 1968, p. 156

exists by itself apart from the body. Is death anything but that?"[225] One of Socrates' primary arguments appeals to the simplicity of the soul. Socrates believes the soul will not dissipate upon death because it is "to the composite, which is by nature compounded, to which dissolution is proper," [226] while the soul is a simple, uncompounded substance. Unlike physical objects that have proper parts, the soul is typically thought to be a single, unified substance with no proper parts. The simplicity of the soul is one of the things that differentiates it from the body in the opinion of Socrates and most other dualists in philosophy.

In both of these cases, appeals are made to our modal intuitions about the body and the mind or soul. They rest on the idea Alvin Plantinga describes as follows, "I take it that we human beings have the following epistemic capacity: we can consider or envisage a proposition or state of affairs and, at least sometimes, determine its modal status—whether it is necessary, contingent, or impossible—just by thinking, just by an exercise of thought."[227] The idea behind the modal argument for the distinctness of the mind and body is well summarized by Stewart Goetz as follows,

> The modal argument goes as follows: I can conceive of myself existing disembodied; therefore, it is reasonable to believe that it is possible for me to exist disembodied. Because it is a necessary truth that no physical body is able to exist disembodied…it follows that I am a soul and not identical with my (or any) body. Thus, I am a soul and dualism is true.[228]

Here Goetz describes the central idea behind the arguments from disembodiment for dualism. Although there have been several objections to this style of argument, given an epistemic

[225] Plato, 1984, p. 467
[226] ibid, 482
[227] Plantinga, 2007, p. 2
[228] Goetz, 2005, p.44

position that gives strong presumptive weight to our intuitions, and given the widespread and natural intuitive appeal of these ideas, there is good reason to think that this argument captures something important about our ordinary conception of the mind.

In addition to these defenses in philosophy, one shouldn't forget that for a long time people took idealism very seriously. Idealist views of the self such as Berkeley's don't even make sense without the assumption that the mind is an immaterial substance capable of existing without a physical body. Both within philosophy and without, the view of the mind as an immaterial substance capable of existing without a body is a persistent part of how people view human nature. This further supports the idea that people are natural dualists, and that common sense views the mind or the spirit as a substance that exists in addition to the body.

7.2 Disembodiment and Agency

These intuitions about disembodiment motivate a substance dualist view of the mind in and of themselves. However, I believe they also help explain why it is that the mind is capable of the causal activity described in the previous chapter. There are some recent efforts to defend much of what I say about free will and agency within a materialist framework. Trenton Merricks seems to endorse this sort of view,[229] and Timothy O'Connor has gone to great lengths to develop such a position in impressive detail.[230] However, views like these face the difficulties that befall all accounts of radical emergence, and also fail to have the advantages of a dualist metaphysic. We will start with an example of these advantages first and then move on to problems with material agency.

One thing to notice about the appeals to agency made in human history is that they usually don't include any requirements about the nature of an object that limits the application of

[229] See Merricks, 2001, esp. Ch. 6
[230] See O'Connor, 2000

agency to the object. Even something as remote from consciousness as "a pair of pale green pants with nobody inside them" has been personified in children's literature,[231] and is represented as something capable of engaging in intentional acts, and of having a mind and a personality. This natural tendency to suppose the possibility of agency in any body is best explained by the view that agency comes with the addition of a spirit or immaterial substance to an otherwise unchanged object. The idea that agency can be an add-on to any object does not fit well with a view according to which it is an emergent feature of a special configuration of matter.

In addition to losing some of the intuitive appeal held by dualists, emergence views face other problems. This first and most pressing problem for these views is that they violate what appears to be a basic principle of metaphysics: that a whole is nothing over and above its parts in combination. According to these views, the self is capable of producing actions that are not the result of the various parts of the self or how those parts are combined. The emergence view requires new causal powers that are wholly inexplicable in terms of the various parts of an object. It requires that these powers be able to produce changes in behavior in which the laws of nature that apply to the constituent parts of those objects fails to hold in certain contexts for no ascertainable reason. The idea that some wholes possess novel powers that counter-act or override the standard facts about causation governing the constituent elements requires a rather peculiar metaphysics of causation, and adds new laws of nature to the universe that are not only hard to account for, but that also lead to objects wholly unlike those that exist as the result of any other arrangement of matter.

It is here where we see the clearest advantage of the dualist metaphysic when it comes to free will. We already naturally think of immaterial minds as sources of active power; this is

[231] See Seuss, 1961

what makes spirits such appealing sources of causal explanation. A rich conception of a fundamentally different kind of substance that engages in fundamentally different sorts of causal interactions provides a much more comprehensible loci for the sort of causal power available in agency than a human organism does. In addition, it provides us with a much cleaner metaphysic. On this view, what fundamentally exists are simples of different sorts interacting with one another in various ways. Composite objects consist, at any given time, in a certain collection of simples engaging in appropriately cohesive interactions to maintain the functions that allow us to recognize them as a cohesive unit and apply the term "object" to them. We will look more closely at such a view in chapter 9 when responding to certain concerns about apparent materialist intuitions. For now, however, it should be enough to indicate that dualism has theoretical advantages over compromise views that try to maintain a radically distinct view about the causal nature of agency while maintaining an otherwise materialist framework.

In a paper objecting to naturalist views is meta-ethics, Michael Huemer ended with the sentiment that "if you're going to be a moral realist, you might as well go the whole hog with intuitionism."[232] His point was that efforts to garner some sense of naturalistic credibility in areas traditionally conceived of as involving non-physical elements of reality fail to sound plausible while inviting a host of new problems. I believe this lesson is right, and that it should be extended to philosophy of mind. The fundamental difference between the pictures of the world held by naturalists and dualists grows out of a difference in the methodological frameworks that underlie the respective worldviews. Mitigating efforts lose plausibility and credence because they separate themselves from the apparent and straight-forward ontological commitments of the different philosophical frameworks. Liberty with a timid ontology won't get you what you really want, and even if it could, it invites an array of new and difficult problems with little in the way

[232] Huemer, 2000

of plausible answers or clear avenues of response. The battle for genuine freedom shouldn't be fought in the trenches after ceding the war over method. It should be fought on the front lines, focusing on the big picture questions about what methods make us best able to see the world as it truly is. Once we have fought this fight, we should simply allow the ontological pieces to fall where they may.

7.3 Subjectivity and Rationality

Another feature of the mind we should examine is the subjective nature of experience. When we have experiences, one of the central features of the experience is the sense that it belongs to us. We see the world from a certain point of view, and are aware of certain things from the inside that others can't directly observe. The knowledge we have of our own mental states comes from a unique perspective. We are aware of what is happening in our own minds in a way that we are unable to use to know what happens in the minds of others. Subjectivity consists in the perspective or point of view from which you have experiences. Experiences are had by subjects of experience. In addition to this fact, there is a general sense that in the absence of a mind nothing can have such experiences. Subjective experiences are a type of event that minds can be involved in but that no other type of substance can have. When we conceive of material objects, if we want to imagine them as subjects of feeling or desire or as aware of what is happening to them, we must personify the object, and we do that by imagining that it possesses a mind in addition to the matter that composes it. The nature of things that experience the world and things that do not appears fundamentally different and naturally invites us to suppose that there is something in addition to the matter that composes the object whenever we attempt to conceive of something as a subject of experience. According to Paul Bloom, if you want to create a story with a supernatural being, you should first start with the notion of a soul, and then

have it exist either by itself or embodied in an unusual way. According to Bloom the entire catalogue of supernatural notions "involve the attribution of mental states" and, despite usually possessing abilities that go beyond our own in such attributions "none of this violates our foundational understanding of how souls work."[233] This fits quite naturally with a dualist metaphysic since it makes it clear just what is added to matter in order to make what happens to it result in experiences. The ability to participate in a new type of event arises only with the addition of a mind. The sense of a self as a subject of experience, then, seems connected to the sense of the mind as an independent and additional substance that has those experiences. As I will discuss next, this subject of experience also presents a sensible location for the compositional features discussed in chapter five, while helping us understand the relevance of those features to permitting intelligent exercises of the power an agent possesses discussed in chapter six. We will turn to these points now. In doing so, we should see a compelling picture of the mind as an independent substance and also see more clearly how this view of the self is integrated into our conception of human beings as rational and moral agents.

We should begin by examining some ways that experiences differ when we consciously attend to them and when we do not. David Armstrong presents a case of this by discussing the difference between paying attention to the road while driving and driving on auto-pilot. Armstrong describes the difference in these cases by pointing out that "we cannot directly observe the minds of others, but each of us has the power to observe directly our own minds, and 'perceive' what is going on there. The driver in the automatic state is one whose 'inner eye' is shut: who is not currently aware of what is going on in his own mind."[234] There are other familiar examples of these phenomena. For example, when focusing intently while writing, if

[233] See Bloom, 2004, p. 212-13 for a full discussion.
[234] Armstong, 2005, p. 148. Armstrong believes that this is physically manifested in a "self-scanning mechanism in the central nervous system."

you are lucky, you may be made aware of an existing sensation of pain when someone rubs your shoulders. The physical state was obviously already there, but subjective awareness of the state arises when we attend to it. Notice that attending to the relevant states automatically includes recognition of the sensations that they produce. Sensation appears to be an intrinsic and perpetual aspect of the experience of consciously attending to something. We can easily differentiate experiences where we are consciously aware of a sensation from ones where we are not, and one primary way we do so is by feeling the experience's effect on us from the inside.

We can respond on autopilot to unconscious experiences, but we typically respond more deliberately and intelligently when we attend to the state and become aware of its qualitative aspects. In part this is because in addition to the presence of a sensation, there is typically a way the sensation is presented that provides us with additional guidance. Sensations often draw our mind's attention to their causes. In addition, they are typically presented in a way that allows us to gage their significance as well as to guide us in what to do about its presence. Appearance states include sensations as well as indications of the relevance of the state. When something is frightening, there isn't just a sensation, that sensation is presented as important and guides us to avoid or move away from the object in question. In a complex creature that has multiple competing stimuli, sensations appear to help us know what stimuli to attend to first, and help guide us with respect to how to respond to them.[235] The discussion of appearance states as epistemic seemings is an example of this pattern.

In addition to this, when we consciously attend to something, we are typically aware of the content of the state in question. The ability to understand and intentionally respond to content seems to be the largest part of intelligence. While I claimed in chapter two that content

[235] A fascinating example of how we deal with and unify complex states can be seen in the "McGurk effect." In this case, the effort to synthesize visual and auditory clues creates an inescapable auditory illusion. You can experience it for yourself at http://www.youtube.com/watch?v=G-lN8vWm3m0

by itself isn't sufficient to explain belief production, the content of our beliefs and experiences is certainly relevant to other things we do with them. For example, once we form a belief, we often realize that the content of the belief holds a connection to the content of some desire we have, and use this fact in evaluating the relevance of the belief for fulfilling or resisting the desire in question. So, this awareness appears to allow us to act in ways that are responsive to that content as an independent component. Meanings permit a sort of interaction with our environment not available to creatures that merely respond to stimuli. Indeed, I believe it is this that is missing in the Chinese room. When we understand a language, our responses occur not just to the syntax but also to the content itself, which is recognized as an additional feature of the language that we must recognize in order to understand. When we understand the content of our experiences, we can see how it connects with the content in other experiences and in other ideas. We can focus on certain aspects of the content and connect it with the content of other states in order to better understand what is going on now, to integrate it into in complex plans or deliberations, and to engage in processes of reasoning that allow us to see its consequences or significance. As Alvin Plantinga puts it

> We ordinarily think true belief leads to successful action because we also think that beliefs cause (part-cause) actions, and do so *by virtue of their content*. I want a beer; I believe there is one in the fridge…we think it is by virtue of the content of that belief that it causes me to go over to the fridge; it is because this belief has as content that there is beer in the fridge that it causes me to go to the fridge rather than, say, the washing machine.[236]

This responsiveness to content as an independent feature of our how we engage in behavior makes perfect sense on the picture of the mind we are endorsing here. When one views a state's

[236] Plantinga, 2011, p.353

content as consisting in its ability to permit appropriate responses, one misses out on what it is about these states that makes the response appropriate. Rather than being limited to a basic response structure produced by the neural event that is correlated with having states of awareness, the mind can understand the significance of the content and rationally respond to the content itself. Rather than simply reacting in the default way the neural pattern would otherwise respond in a creature that lacked that awareness, intelligence and understanding open up an array of possibilities. Seeing content as an independent feature of awareness that permits rational responses fits perfectly with what we have seen so far. The content allows us not just to respond to a state, but to understand what it is we are responding to. Without seemings, such understanding couldn't produce rationally motivated behavior, but without the understanding permitted by the presence of the content, rational evaluations and comparisons; that is, reasoning that is responsive to content and requires understanding, couldn't occur. Placing these features in combination allows the mind to understand what is going on, to take that content into consideration by appreciating its current relevance, and then to reach rational conclusions about what to do in response to the content of these states.

Something as simple as deduction is rational only because we understand the relationships between the contents of the relevant propositions. Rational responses to our environment require both an ability to respond to content as such, and a reason for responding to that content in one way rather than another. This, of course, would greatly enhance the survival prospects of any creature that could use these states of awareness as a basis for directing its behavior. The causal independence of an immaterial mind provides us with a source of power for interfering with the course of events that would have otherwise have occurred. Through an original contribution of the will as a source of active power, we can make rational recognition

effective in producing free and rational action. The ability to engage in rational responses to our environment brings together the various features of an immaterial mind to form a coherent picture of it as a source of rational engagement with our environment. Plantinga claimed that without this responsiveness to content we would have no reason to think that our behavior was likely to be rational or our beliefs to be truth-apt.[237] This may be right. However, one needn't reach Plantinga's view that divine interference is therefore necessary to provide us with reason to trust the reliability of our beliefs or underwrite the rationality of our behavior. If one accepts the existence of immaterial minds, there is an obvious evolutionary advantage to their presence precisely because it permits the rationality a material self would be unable to provide. This isn't to say one couldn't also endorse this view from a theistic standpoint. If God wanted to create creatures who rationally engaged with their environments and could therefore engage in free and morally significant action, this would appear to be the way to do it. Rational action requires a stable environment to allow us to understand the consequences of our behavior together with an independent source of power to bring about effects that are our own. The fact that this view of the self makes sense of the value of and reason for the existence of rationality on both evolutionary and theistic views of the purpose of human existence should make it a friendly addition to either view of the origin of human nature.

7.4: Putting the Pieces Together:

These features of the mind fit together into a coherent picture of an immaterial substance as that part of us that allows for rational and moral action. Content, sensations, and appearances come together in our subjective experiences of the world. This gives creatures engaging with their environment the ability to understand the relationship they hold to that environment and a rational basis or motivation for responding to that environment in light of the meaning or

[237] Plantinga, (2011), Ch. 10, section V

significance of their environmental stimuli. This capacity for intelligent understanding of our relationship to the world is not, however, left as one possessed by a rational but helpless observer. Instead, possessed of an inherent power to create behavior, the mind interacts with the body to produce rational responses to our environment. Sensations or qualia help draw our attention to and recognize the importance of various events happening around us. Understanding creates a level of long-term and contextually complex rationality of responsiveness that otherwise would not exist, and the will offers a means of doing something about those states of understanding by providing us with a new source of power to act. Together, these create a picture of the mind as a substance with power in itself to understand and to be motivated to interact with the goings-on of a particular body from which it derives its experiences of the world. Let's return briefly to the claims made in chapter 5 about how we individuate substance types. We can now fill in the characteristic features of mental substances that differentiate them from other types of substances. First, mental substances can have qualia and content, while other types of substances cannot. Second, mental substances possess inherent active power to produce effects, whereas abstracta lack power of any sort and physical concreta posses only derivative and undirectable power. Third, mental substances are properly thought of as substances in virtue of their ability to exist independently of other substances. Finally, mental substances uniquely participate in a special type of event: subjective experiences. No other type of substance is able to directly participate in this type of event. Together, these features also explain the view of human beings as rational and moral agents who are capable of engaging in deliberate intentional action. Since this is the view of one another that seems presupposed in everyday interactions with other agents, this view of the self seems to qualify as a integral part of the common sense world view.[238]

[238] I believe this goes a long way to assuaging the moral concerns about dualism disconnecting us from life

Contemporary versions of dualism have struggled to gain prominence in the philosophical literature. I believe part of the reason for this is that most contemporary versions of dualism are odd chimeras, slapping rogue immaterial features onto an otherwise materialist metaphysic. By trying to concede as much as possible to the physicalist while holding out for their pet projects, dualists have divorced themselves from the picture of the self that has been embraced so readily by so many for so long. Dualists should be brave and unapologetic in their views. We should forthrightly acknowledge that we are saying the world is fundamentally different than physicalists believe it to be. A bit of qualia or agency stuck onto an otherwise physical world isn't like a menorah stuck on a Christmas tree, it's more like a ham steak stuck on one. The aspects of the dualist's view fit together and create a view of the self that acknowledges the fact that human beings are fundamentally different than the rest of nature, and declares that the world is not the limited place physicalists claim it is. This fact should be embraced. It is only together that these features of the mind provide a compelling alternative to physicalism, because it is only seen together that they create the image of the self that resides in common sense.

7.5 Is Interactive Dualism too Interesting?

When discussing the content of the first chapter of this work with others, one common response is a sort of ad hominem along these lines: "you say we should be suspicious of interesting views, but interactive substance dualism is an interesting view, so shouldn't be we suspicious of it?" My answer to this question is that dualism isn't really interesting. Although the view is interesting in philosophical contexts given that it is widely rejected in the field, it isn't interesting in the more general sense of going beyond or countervailing our experience of

sometimes raised in response to more ascetic versions of dualism. The mind's value lies primarily in its ability to allow us to live meaningful and rational lives. Disconnecting from life in order to pursue merely intellectual pursuits is hardly a decent or rational response to this opportunity.

the world. Interactive substance dualism is the combination of the trivial facts that (1) introspectively accessible mental states aren't physical, that (2) they are causally relevant to behavior, and that (3) they are aspects of a substance whose modal nature is distinct from the modal nature of our physical bodies. These views are either simple consequences of introspection or else consist in basic modal intuitions. Their combination therefore doesn't rise to the level of a theoretical hypothesis. Instead, it is just the consequence of direct observation and conjunction. It is an odd feature of philosophy that the same views are often attacked as both highly simplistic and naïve and as overtly mysterious and incomprehensible. Dualism is often thought to be both a persistent aspect of common sense and a bizarre and incomprehensible view that posits strange substances and incomprehensible forms of causation. I will address the concerns over the mysteriousness of dualism in the subsequent chapters. For now, I would like to explain quite simply why I find dualism uninteresting.

In presenting the case for the view that dualism is part of common sense and that physicalism is fundamentally at odds with common sense ontology, I have tried to be as fair to the physicalist as possible. I tried to explain in a detailed way exactly what features of the mind draw one to think of it as an immaterial substance, and how these aspects of the mind fit together into a cohesive whole. I have tried to show how this conception fits into a sense of people as rational and moral agents, and how this shows that a materialist view of human nature is at odds with basic assumptions about human nature we carry with us in life. I have done this because it is necessary. Physicalism dominates contemporary philosophy of mind, and contemporary intellectual life in general. I have tried to treat with it with as much respect as possible. But for a moment, I would like to step away from this attitude of respect and explain why I think this should not have been necessary. If nothing else, this effort should help physicalists see just why

it is that so many people have such a hard time accepting their view. Here is how the debate seems at least to this dualist, and likely to many others. Although physicalism requires respect because of its broad acceptance among those I am engaged with in debate, but the view itself has never actually *deserved* it. Here, briefly, is why:

In introspection, we are aware of many of our mental states. We can look within ourselves and see what is going on. In our experiences, we are often aware of what is happening inside of us in addition to what is happening outside of our minds. In addition to what the experience is of, there is a way it feels or a "what-it's-like"-ness that is a distinct feature of the experience. When we focus on these features of our lives, we are aware of intrinsic aspects of their nature. Think, for a minute, about a experience of joy. In such an experience, there is a certain acuteness, a level of intensity, and a distinctive phenomenology or flavor to it that is a recognizable part of the sensation. This is, indeed, the aspect of joy that most draws our attention and that we are most readily familiar with in experience. Of course, there are also physical events occurring at the same time. Often, there is activity in or uplifting of your chest and your heart rate and breathing likely change to some degree. Muscles move in your face and your posture often changes as well. But these physical events are incidental. What makes joy what it is lies the aspects of the sensation itself. Focus for a moment on a particular instance of that sensation. Focus in particular on those features of it such as its intensity and its experiential flavor. According to physicalists, those features aren't anything more than various physical processes transpiring in your body at that particular moment. Unless one is a behaviorist, it won't be the list above, but according to all physicalists, that experience and those features of it ultimately consist in nothing more than matter arranged in a complex way. Physicalists say this despite the fact that the experience of having the sensation appears to be a complete and distinct

thing in its own right, separate from any physical goings-on. They say it despite the fact that the experience appears to have its own features, and a nature that is quite different from the nature of any of those physical events. They say it despite the fact that it does not appear to be made up of little bits of stuff, and despite the fact that it appears not to have any mass or to be the sort of thing you could put in a test tube or cut up into parts or push along thin fibers of tissue in the form of synaptic firings. Somehow, despite the massive apparent difference in nature between these particular instances of the experience of joy and the particular physical events occurring in your body, the one thing is supposed to be nothing more than the other. That is what physicalists are asking us to believe. And that is simply crazy. Two things that are so very unalike and so obviously different on their face simply cannot be one and the same. We could accept many different theories on the underlying physiology of, say, dogs. But we could and should never grant the hypothesis that their true nature shows they can't have fur, never bark, have no legs, breathe through gills, and live their lives swimming in the sea. Such a theory simply cannot be true.

And why is it that they are asking us to believe that their theory is true? Well, primarily because when we have the experience of joy we see certain physical events happening in the body. They appear to co-occur, and the physical stuff seems causally related both to the physical events that brought about the emotion and to the physical events, such as your subsequent behavior, that follow the experience. And isn't it much neater is we just say there's one thing there instead of two? Well, my dad delivered mail for more than two decades. When he would go out to deliver mail, he would drive around in a mail truck on his routes. Wherever he went to deliver mail, the truck was right there with him. We know that part of what makes someone a postal carrier is that he meets the description of "going around town and distributing mail to

various locations." But there is something else that also fits this pattern: the truck. But one would have to have a seriously diminished understanding of what a postal carrier is, and would also have to think that only metal detectors or views from the sky were legitimate ways to detect a postal carrier, to look at the pattern and say that mailmen are actually large, heavy, metal objects with four wheels and internal combustion engines. This could be a revolutionary new account of the nature of mailmen if one were to simply ignore all the evidence that let us know they are human beings and to only accept evidence that could detect trucks. But I knew my dad, and I'm quite certain that he was not a truck. I recognize his features and the ones he lacks, and I know that, although a particular truck was there when he delivered mail around town, at no point was my dad made of metal, have four wheels, or have an internal combustion engine. The fact that he happened to overlap in some of his interactions with the world with such an object doesn't make them the same, and doesn't even make it possible that they could have been the same. And when I introspect, I am also aware of things, such as particular sensations of pain. I am aware of various features they have, like their intensity and acuteness, an intrinsic negative quality, and a certain ineffable "what it's likeness" that goes beyond these features. I am also aware that they aren't composed of matter, don't have any mass, and don't move quickly through a bunch of squishy grey stuff. You can tell me all you want about how certain physical events are present when the pain is there, and how those events seem to be relevant to my behaviors such as saying "ouch" or getting a band-aid, but I've introspectively seen my sensation of pain just like I've visually seen my dad, and I know first hand that what I've seen in introspection is not that thing you see accompanying its causal role in your laboratory. Point to any instance of brain activity and try to say it is acuteness or intensity, that it is intrinsic badness or has a feature of being like something, that it is an experience had by a subject and felt with anguish and a

longing to fix or remove it, and I will look at you like you have three heads. These things are simply different. Asking me to believe otherwise would be like asking me to believe that my dad was actually a large, heavy truck.

Physicalism is so manifestly contrary to our experience of ourselves that it would take a massive and widespread intellectual bias against all forms of introspection and against the legitimacy of common sense in general to have ever gotten to a point where it was even taken seriously, let alone widely accepted. When engaged in a defense of a view like this, people often find a way to move past this point of initial resistance and to think that they are doing serious work trying to figure out a great mystery. Brilliant minds spent centuries, for example, engaged in the opposite task of developing complex and fascinating theories about how things like your brain and the rest of the physical would could actually be nothing more that mental features of reality, and the array of idealisms that came out of this effort were astounding in scope and in creativity. But none of them ever had a chance of being true. And none of them ever should have been taken seriously as legitimate ways of conceiving of reality or of human nature. The same is true of physicalism. Brilliant minds have justified to themselves ways of looking past the obvious falseness of their view, and are currently engaged in amazingly complex and brilliant attempts to defend something that can't possibly be correct. At least, I think this is how the debate is seen by many dualists, and is a large part of why the debate appears so intractable. If this is correct, then the effort is also a tragic waste of rare genius.

7.6: Concluding Remarks

We have seen that there is good reason to believe that a dualist metaphysic is integrated into our ordinary understanding of human nature. Given the work of the earlier chapters, this should provide dualism with a strong presumptive status in philosophy of mind. In fact, given

that physicalism is best understood as an error-theoretic replacement view, the work of the first few chapters provides us with strong reason to think that physicalism should be viewed as a non-starter. Nonetheless, since several arguments have been presented for physicalism, and since the degree of conviction in the common sense approach is likely to be less than full and universal, it is worthwhile to carefully examine the reasons offered in support of physicalism and in opposition to dualism. We will turn to this in the remaining chapters.

Section III: Objections and Replies

Chapter 8: Causal Objections

8.1: An Overview of Section 3

So far I have laid out the main considerations in favor of the views that our ordinary, common sense view of the world ought to serve as the foundation of our philosophical views, and that our ordinary, common sense view of the mind commits us to substance dualism. Together, these show that dualism deserves the presumptive position in the field and that to overturn dualism the physicalist will need powerful arguments against it. If I am right that the physicalist position is a drastic departure from common sense, the strength of these arguments will have to be sufficient to justify an extensive error theory about human nature and a significant revision of our ordinary beliefs. I have contended that such revisions are virtually impossible to justify, and I think dualism should be accepted without much hesitancy at this point by those who embrace this method. However, since I know that many readers are unlikely to share my evaluation of the amount of work it would take to overturn the presumption, I want to take on the objections to dualism on a more equal footing than I think the preceding work entitles them to. I will seek to show that the objections to dualism are far weaker than it is generally assumed, and that taken together they fail to defeat even a moderate presumption in its favor.

Objections to dualism fall into two main categories. The first, and by far most influential of the categories, is the causal objections. These objections come in both *a priori* and *a posteriori* forms. The *a priori* versions range from Princess Elisabeth's initial concerns raised in correspondence with Descartes, to contemporary concerns such as Jaegwon Kim's pairing problem. These versions seek to show that there is something fundamentally problematic or

perhaps even incoherent about the understanding of mental causation the dualist offers. Meanwhile, the *a posteriori* versions seek to show that even if the notion of immaterial causation is coherent, the current empirical data shows that in fact no such causation takes place. These range from objections appealing to particular scientific discoveries, such as the conservation laws or results from neuroscience, to more general constraints grounded in the claim that physics justifies causal closure, to more moderate claims that we have sound inductive grounds for thinking that physics can provide a complete account of all physical events without ever having cause to appeal to immaterial minds. A combination of these types of objections are also raised against agent causation. All of these causal worries will be responded to in the remainder of this chapter.

Non-causal objections to dualism come in two main forms. First and most prominently are objections grounded in considerations of parsimony. It is generally assumed that simpler theories should be preferred, and that this serves as a significant advantage for physicalism. Second, there are those who attack introspection and intuition as sources of evidence and use this as a basis for preferring a naturalistic approach to philosophy of mind. Perhaps more importantly given the aims of this work, some have claimed that endorsing dualism is contrary to common sense, and have argued that our ordinary sense of the self takes it to be that thing we see in the mirror and not some immaterial mind. These objections question the contention that dualism offers the default, common sense view of the self. These objections will be considered in the final chapter.

8.2 The Anti-Interactionist Intuition

Many people seem to have an intuition that there is something odd or problematic about immaterial causation. Although I do not share this intuition, it has been around for centuries,

dating at least back to Princess Elisabeth's objections to Descartes. Elisabeth describes her concern as follows:

> it seems that how a thing moves depends solely on **(i)** how much it is pushed, **(ii)** the manner in which it is pushed, or **(iii)** the surface-texture and shape of the thing that pushes it… The first two of those require contact between the two things, and the third requires that the causally active thing be extended. Your notion of the soul entirely excludes extension, and it appears to me that an immaterial thing can't possibly touch anything else.[239]

Others have shared this concern. For example, Daniel Dennett asks of mental states, if "these, *ex hypothesi*, are not physical; they are not light waves or sound waves or cosmic rays or streams of subatomic particles. No physical energy or mass is associated with them. How, then, do they get to make a difference to what happens in the brain cells they must affect, if the mind is to have any influence over the body?"[240] For many people, the attempt to conceive of immaterial causation leaves them confused at best, and certain of its impossibility at worst.

The first thing to say in response to this intuition is that it is far from universal. Most people have little trouble thinking of God, for example, bringing about events, and this involves a cause from an immaterial being. Nor is it uncommon for people to posit immaterial causes of currently inexplicable events, as we discussed chapter 6. The tendency to posit spirits as causes shows that the idea of immaterial causation is far from incomprehensible to most people. The second thing to notice about these concerns is how they are typically framed. Usually, the objector wants to know *how* mental causation works. Thus Searle asks "are we supposed to think that our thoughts and feelings can somehow produce chemical effects on our brains and the

[239] Princess Elisabeth of Bohemia, 1643. Quoted in Anscombe and Geach 1954
[240] Dennett, 1991, p. 34-5

rest of our nervous system? How could such a thing occur? Are we supposed to think that thoughts can wrap themselves around the axons or shake the dendrites or sneak inside the cell wall and attack the cell nucleus?"[241] Descartes believed that part of the reason that people struggled to conceive of the mind correctly was because they insisted on trying to imagine it visually. As he puts it, "it seems to me that those who wish to use their imagination to understand [the soul] are doing just the same as if, to hear sounds or smell odours, they attempted to use their eyes"[242] A similar phenomena appears to be happening here. When people ask how mental causation occurs, they seem to be asking for an account of the *process* that explains the mind's effect on the world. But this betrays a mechanistic style of causal explanation that underlies the confusion. As Plantinga puts it, "To ask, "How does a self produce thought?" is to ask an improper question. There isn't any how about it."[243] the production of ideas and the initiation of behavior don't follow processes that one can visualize or imagine. Acts of mental causation are singular exercises of active power. I see little reason to think that such exercises of power should or even could have the sort of systematic structure that would be present in a process. The idea that every type of causation involves a process of moving from one event to another leaves out active power as an explanation from the start. It is not hard to see, at that point, why those who are wedded to it struggle to understand what is going on.

Finally, it is worth pointing out how weak this style of objection is, particularly as a grounds for changing one's mind. There are bound to be aspects of any theory that have not been fully worked out, and outstanding problems that, to one's opponents, may seem unsolvable. William Demski, for example, claims that cells have structures that cannot be explained by

[241] Searle, 1984, p.17
[242] Descartes, 1968, p. 17
[243] Plantinga, 2006, p.117

standard evolutionary theory and require the intervention of a divine mind. Biologists have, to my knowledge, yet to fully account for at least some of these features. At any rate, there are certainly some biological features that are puzzling or difficult to understand the origin of. But this has not led biologists to throw up their hands and think they have to fundamentally restructure their understanding of evolution. Nor should it. Some problems in an otherwise well-established theory typically shouldn't motivate rational people to abandon that theory. Dualists have been rather neglectful in providing an account of mind/body interaction. I hope that some of what I said earlier in this work helps to mitigate that. But lacking a fully developed account of the interactions between two substances doesn't tell us that one of those substances can't really be what it appears to be. A lack of certainty about an aspect of a view should invite investigation, not revision.

8.3: Kim's Pairing Problem

Jaegwon Kim has developed a far clearer explanation of the problem alluded to by others. Kim's formulation of the difficulty is known as "the pairing problem." This problem focuses on understanding why it is that, if you and I were to simultaneously decide to raise our arms, my willing would raise my arm, and your willing would raise yours, instead of the other way around. When simultaneous physical event produce similar effects, we can identify the causes due to the spatial relationships between cause and effect. However, with immaterial minds, there are no such spatial relations to appeal to. In the absence of such relations, how are we to understand the relationship between the mental occurrence and the physical one such that my mental states cause the physical effects in my body, rather than in yours?

When Kim presents this problem, he sometimes presents it as an epistemological worry, and sometimes as a metaphysical one. For example, he asks at one point "…how would we

ascertain causal relationships in [a Cartesian] world?"[244], and at another "But do we have any understanding of such purely mental causal chains?"[245] These epistemological worries about our ability to identify or to understand the nature of mental causation are distinct from metaphysical questions about what is actually occurring in such cases, but Kim raises these sorts of concerns as well. Kim sets up the problem by noting that for a dualist "a causal account of ownership [over our bodies and our actions] seems the most natural option for substance dualists," and therefore claims that a mere appeal to the fact that it is my hand or your hand that is raised won't suffice to explain this connection. He then asks a clearly metaphysical questions about the relationship of ownership, stating

> What I am asking for is more basic: If God "united" my mind and body to make a person, there must be a relationship R such that a mind stands in relation R to a body if and only if that mind and that body constitute a unitary person. In uniting my mind with my body, God related the two with R. Unless we know what R is, we do not know what it is that God wrought. The word "union" remains a mere label… (78)

Although these are distinct questions with distinct answers, the answer to the epistemological question should help shed light on how to answer the metaphysical one.

Asked in a very straightforward way, the question "how do I know that I am the cause of my own behavior?" is a very simple one to answer. We know it is ours the same way we would no matter what view of the mind was correct: we are aware of ourselves as the source of the behavior in introspection. A sense of self is present in one's kinesthetic awareness of our body's movements. This is how we know. It's not like the average person is waiting with bated breath

[244] Kim, 2005, p. 84
[245] ibid, 82

to see if the neuroscientist can tell him that he was responsible for eating his breakfast this morning. Our knowledge of our ownership of our behavior is a basic part of our experience of ourselves in everyday life. The epistemological question is particularly bizarre in the case of agency. In this case, what Kim appears to be asking is "how can you know what you are doing?" This hardly seems like a deep philosophical mystery. If an account of causation suggests that it should be for dualists, that seems to suggest a problem with the conception of causation they are using, not one for those who reject that conception.

The epistemological issue, however, isn't the interesting or important one. The more fundamental problem for dualists is to provide a sensible, comprehensible account of how an immaterial mind and a physical body could causally interact so as to be united into a single person. As Kim puts it, the dualist needs an account of the R relation that unifies body and mind. The full details of such a view would require time to develop, but it is reasonable to hope that dualists could at least provide a framework for understanding how such a relationship might work. In order to begin to provide that understanding, I will sketch what I take to be the most plausible account of this relationship.

On my view, the R relationship is a combination of two things: experiential awareness and sovereign control. Descartes speaks to the need of a close relationship of both sorts, saying

> "it is not sufficient that [the soul] should be lodged in the human body, like a pilot in his ship, unless perhaps to move its limbs, but that it needs to be joined and united more closely with the body, in order to have, besides, sensations and appetites like our own, and in this way to constitute a true man."[246]

The close relationship between the body and the mind, needs to be one that explains the minds awareness of what is occurring in the body, the sense of control it has over the body, and the

[246] Descartes, 1968, p.76

sense of self that arises from this combination. When physicalists ask for an understanding of the "upward" causation from body to mind, they seem to think of this relationship as similar to other types of event causation. They tend to think of it as a sequence of events, one in the body and one in the mind, and the first somehow pushing the mind around in a way that produces the relevant experience. I don't think this is the right way to think of the events, and that this helps explain much of the confusion. One of the basic features of a mind is the ability to be aware of things. Minds seek out and attend to things, and in doing so they become aware of certain aspects of the world. Among the things the mind is aware of are the happenings within and directly involving a particular body. We have both a proprioceptive and a kenesthetic awareness of a certain chunk of the physical world, and are in a state of constant awareness of it. In addition, our awareness of the rest of the physical world is mediated through the world's effects on various parts of the body, such as the eyes, ears, nose, and skin. The direct awareness we have of our bodies and of the world from the point of view of our bodies is a constant feature of the mind. However, in this experience it is wrong to think of the mind as a passive receiver of signals, like an air-traffic controller awaiting calls in from jets and sending back replies. The mind doesn't sit back and await incoming signals, it reaches out and pays attention to what is happening. Experiential awareness is a function of the mind's ability to pay attention, and paying attention is an activity. There is a certain low-level of attention that is present for the body as a whole and is felt in proprioception. But there is also a more focused and deliberate attention that occurs and that occupies most of our conscious awareness. Sometimes, the presence of unusual occurrences in the body, such as a sudden wound, is accompanied by a form of awareness that is intense and acute, and draws the mind's eye. Other times, we decide more

freely to think about or pay close attention to certain things or ideas that we believe to be important. In either case, however, the mind doesn't appear to be a wholly passive participant.

Awareness appears to be a natural and automatic result of drawing attention to something. In most animals, a basic level of attention and a heightened attention to certain types of sensations that accompany certain effects on them is likely the limit of the mind's awareness. This limit likely also limits how well their minds are able to direct their causal power. In humans, however, in addition to attention, there is intelligence. An intelligent mind is not only aware of the content of their experience in a brute way, an intelligent mind can abstract out, and shape and form that content in ways that allow us to understand and better respond to it. Intelligence consists not just of awareness but of a capacity to conceptualize content and thereby understand its meaning. This permits us to respond to content in novel ways. We need not merely learn to move toward someone when a certain sound is made like a dog does when it hears its owner calling its name. We can understand people's words, see how their meaning connects with other things, and decide what response is best as a result. Awareness combined with a capacity to conceptualize creates the possibility of intelligent understanding, which creates the possibility of rational action in the form of responding to content in ways that are sensitive to its meaning. In all of this activity of the mind, however, there need not be, nor do I see any reason to think there is, a new power of the body at work, sending signals up to a mind. Bodies are dumb, unaware hunks of matter like any other when left on their own. It is the mind that is active in attending to, understanding, and experiencing the world through a particular body. Minds are naturally and automatically aware anywhere they draw their focus. For a mind to be united with a body in the sense of having it be affected by what is happening in the body, all it needs to do is start to be directly and experientially aware of what is happening in it and to

it. For it to be intelligently aware, it must be able to conceptualize that awareness in a way that permits understanding of the content. "Upward" causation, on this view, is a misleading phrase. Minds don't wait for bodies to push them around. Minds seek out awareness of what is happening in bodies, and are influenced by the content of what they are aware of when they pay attention. Using the theological metaphor, half of what God would need to do to unite a mind and a body is to say to a mind "look here, pay attention, I give you constant awareness of what happens in this flesh, in these bones, in this brain and these organs. Attend to it lovingly, for it is yours and it is you."

Continuing the metaphor, the other half of what God would have to do to make it truly ours is to add "and I give you control over it. The power that is within you I grant the authority to exercise and direct toward controlling and guiding this body of yours. I grant power over this portion of my world, this bit of flesh, to you. You are its steward and its very self, take care of it and use it well." A mind with active power and a body it is directly connected to could understand how to direct that power in governing and controlling the actions of a body. As I discussed in chapter 6, I believe that minds have original power as part of their nature that they freely exercise in controlling a body. Power is the sort of thing that people have over certain things and to certain degrees, but exercises of power require nothing more than a knowledge of one's ability to exercise it, and the strength of will to use it. A few things should be said to help explain how the mind and body interact in exercises of agency. Physicalists often suggest that if the mind was to interact with the body it would do so by pushing around or effecting the energy in some part of the brain. While it is certainly true that changes must occur in the brain in the process of acting, we don't view our actions as ones that occur there, nor should we think of agency in so limited a way. We typically view our intentions in far less specific ways. We trust

our body to fulfill our will, and don't usually pay attention to the details unless our constant attention is required. For example, after one has learned to type, one needn't look at their fingers or their keys and try to deliberately strike each key stroke. We learn how to do something without thinking, and then general intentions suffice to bring about large scale bodily changes. As I think these words, they come onto the screen without any attention paid to the actions leading them there. So, thinking of the mind's exercises of agency in most cases, it would be wrong to look at a moment where it started and say the mind's influence was there. The mind doesn't tip over a single domino from its hidey-hole in the brain and hope the rest fall down without a hitch, it typically brings about an entire sequence of events altogether in producing action. The mind brings about a large event containing lots of physical processes through a single act of will, sometimes monitoring it to make sure it goes through, and frequently trusting the body to work right in bringing it about so it can move its attention to other things. So, when looking for the mind's influence on the body, it is misguided to look at a certain point where it acts and then fades into the background. The mind's control is far broader in scope and far more continuous in its effort than is typically imagined by those looking for it in the brain. To say a mind is active isn't to say it has a switch that can be turned on, allowing it to shoot a jolt of energy into the brain. The mind, as it were, is always on. Its involvement with the body is regular and continuous. So, when looking for its influence, we should be looking for it everywhere, not in isolated pockets. Dualists perhaps too often invite the view of the mind as a detached thinking machine, trying to interfere with the events in a distant physical body. But active minds are persistently aware of and engaged with the bodies they inhabit.

Some may worry that this agent causation renders mental causation brute. Typically, an appeal to brute causation is viewed as ad hoc, and is therefore viewed as a serious cost.

However, in this case, brute causation is simply the consequence of the natures of awareness and agency. In cases of awareness, beyond paying attention, there doesn't appear to be anything the mind does that makes it aware of something. Awareness appears to be an automatic feature of attending to something. Given this, asking for a process to be inserted where none appears to go appears ad hoc, rather than the suggestion that there is no such process. As for acts of agency, those were previously described as singular, generative, and intentional. Singular acts of causation have to be brute in one sense, since they aren't explained by any more general pattern. In addition, acts of creation or generations of new causal chains are the sort of thing most naturally understood as brute. For one to be the "prime mover unmoved" of our own actions, the causal chain must originate with us, and originating causation requires that the initiation not itself be a step in some previous causal chain. The nature of agent causation explains why instances of it won't be explainable in terms of a more basic pattern and would therefore have to be brute. In either case, bruteness is simply a consequence of the nature of the occurrence, rather than an additional cost for the dualist to bear.

A second concern some might have with this account is the worry that it begs the question. Kim claimed that one couldn't use an account of the unification of the mind and body to explain mental causation on the dualists view because "a causal account of 'ownership' seems the most natural option for substance dualists," and therefore the account would be circular. In response to Kim, I think that this worry takes an improper view about the nature of objects. On my theory of the nature of objects, their composition is a matter of unified behavior. What makes something an object is the fact that its parts possess a level of causal interaction with one another that is sufficiently unified with one another and sufficiently dis-unified with the surrounding environment to allow for us to recognize it as a distinctive part of the environment.

I will discuss this in a little more detail in chapter 9 when I provide a fuller account of personhood. For now, I will simply point out that given such an account, there is no reason that composites cannot be composed of both matter and immaterial simples. If what makes something an object is a matter of possessing a functional role characterized by causal unity, then it shouldn't matter which parts of reality are engaged in that unity. Composition of a mind and a body would occur just in case the behavior of a mind and the behavior of a body became appropriately unified for the combination to count as one object. This helps explain part of Kim's confusion about the mind and the body's relationship in the first place. The idea that the question of causal unification and the question of composition are separate is mistaken, and so there is no difficulty of accounting for one by accounting for the other.

8.4: Neural Correlates of Mental States

I have argued that the *a priori* objections to dualism either rest on a dubious or question-begging presupposition about the nature of causation or else lack a clear explanation of where the difficulty is supposed to lie. Nevertheless, many people still have a strong sense that interactive dualism must be at odds with the work being done in physics, in nuero-science, or psychology. Although it is often left as a vague appeal or an unargued for assumption, sometimes this concern is offered along with specific examples. I will look at these purported conflicts, and respond to them in the remainder of this chapter.

One of the most common replies I hear when I tell people that I am a dualist is that neuroscience has already proven that mental states are just brain states. The evidence offered for this is typically that scientists have discovered correlations between the presence of various mental states and the occurrence of brain states or other physiological states that are present at the same time. The number of correlations discovered is impressive, although less complete than

is often believed by philosophers. We have learned a great deal about which parts of the brain are active when we are engaged in certain mental activities, such as reasoning, feeling emotions, regulating emotions, and so on. We have also learned a great deal about ways that various physical states, such as chemical reactions and physical stimuli produce certain emotional responses in us, and how to regulate those responses through drugs. For example, we know that dopamine levels affect happiness, and that we can affect happiness levels by affecting dopamine. Many people are deeply impressed by this knowledge, and think that it shows that physicalism is true and empirically established. Many non-scientists are also impressed and surprised. As Paul Bloom describes it

> The natural conception of the brain by children, even after science education, is that it is a tool we use for certain mental operations. It is a cognitive prosthesis, added to the soul to increase its computing power.
>
> I doubt that this understanding is much different from that of many adults. Much excitement has been generated by recent studies showing increased neural activity—part of the brain "lighting up" in a scanner—when subjects think about religion, or sex, or race…[247]

Bloom goes on to say that "people often seem fascinated by the mere fact that the brain is involved at all."[248] Upon learning of the brain's influence, many people start to question their view about the mind. Educated physicalists often turn to reductive theories. Most people I encounter adopt either the hybrid view discussed by Bloom, or an eliminativist view, thinking that the claim that love is a chemical reaction no different from the one caused by eating lots of chocolate means there really is no such thing as love. The eliminativism is typically limited to

[247] Bloom, 2004, p. 200-201
[248] Bloom, 2004, p. 201

those identities they hear have been discovered, but people often feel, as Thomas Wolfe describes it, as though their soul just died.[249] Wolfe himself doesn't adopt the sadness others feel at this suggestion, instead expressing in a blasé tone that dualism is "also known as the 'ghost in the machine' fallacy, the quaint belief that there is a ghostly 'self' somewhere inside the brain that interprets and directs its operations."[250] But whether they are saddened and worried by their loss of soul or merely condescending in their newfound "knowledge," the average response to these findings appears to be that science has shown there is no mind or no soul, and that it's all just chemicals instead. Whether we are dealing with the fears or pomposity of those who misunderstand the science, or the level-headed intelligence of those who work in it and are often misunderstood, I can't find much reason to think that any of these responses have merit.

Although we didn't know most of this information in the past, the presence of such correlations should be relatively unsurprising given interactive dualism's claim that there is regular causal interaction between the mind and the body, and the discovery of the fact that the brain is vitally important to most of the things that happen in the body. What the dualist is proposing is that the mind is in a constant state of interaction with a physical body. The mind's experiences are typically of what is happening either directly within or else right around the body it inhabits, and most of its important roles are directed towards helping the body intelligently interact with the environment. The relevant empirical result is that the brain is central to the performance of most bodily functions and is at least connected to nearly all of them through the central nervous system. The fact that the mind's activities should be well correlated with such a central regulatory aspect of the body is not a surprising fact. In addition, thinking is often regulatory and future-directed, and when it is not, it typically results in direct implementation. In

[249] Wolfe, 2000, p. 89
[250] Wolfe, 2000, p. 97

the last case, the mind's influence requires more or less constant interaction with the brain, and in the first two, since such actions usually occur unconsciously in the future as a result of such planning, there need to be physical states that maintain the efficacy of the intentions. The mind isn't like a CEO who is detached from the workforce and just passes down directives, it's entire experience and value lies in its intimate connection to the activity of the body. When one claims that thinking shouldn't need to occur in close contact with a brain one is likely thinking of the mind in a vastly diminished way. Cognition isn't just something you do sitting in an armchair, it is an active aspect of everyday life.

Given the continual interaction of the mind and the body, the fact that there would be regular activity in the parts of the body most relevant to behavior and awareness isn't at all surprising. The fact that neural states are correlated with mental states also isn't really news. As far back as the 18th century, it was well known that interfering with optic nerves affected vision, and could even stimulate hallucinatory experiences.[251] This didn't lead to a widespread rejection of dualism. Not until reductive projects became intellectually fashionable did the presence of correlation in this area tend to encourage belief in a reductive account of the mind. Physicalists often view the discovery of a correlation between the mind and the brain as a discovery of an identity between that activity and that brain state. But if correlation doesn't imply causation, it certainly doesn't imply identity. The mere presence of physical correlates is expected on an interactive dualist metaphysic. Given this, the discovery of such correlates certainly doesn't entail the falsity of the view. What the physicalist needs here are findings that are contrary to what the dualist would expect, not findings that are fully compatible. Some findings have suggested a more direct contradiction here. In particular, many findings about how behavior is

[251] See, e.g. Reid, 2002

produced seem to conflict with the account of agency developed earlier. We will turn to those after considering more general concerns with the agent-causal picture developed in chapter six.

There is a related position here that should also be addressed. Many people think that the expertise of neuroscientists and psychologists in their areas grant their view about the nature of the mind strong presumptive weight. That is, when neuroscientists and psychologists are engaged in their work, they typically view their results as establishing the true, underlying nature of our mental states. If we are to accept their expertise, many think we should also accept their account of the nature of the things they are studying. In many areas this is plausible, but when it comes to controversial philosophical assumptions of scientists, I can't see why we should think this. When Bishop Berkeley was alive, he was an expert in optics. He knew more about how vision worked than most of his philosophical colleagues, and he certainly knew far more about how it worked than the average layperson. He also claimed that physical objects were mental images, and that the idea of a material substance was incoherent. Does his expertise in how we see physical objects make idealism somehow more plausible? Should we now think, due to the fact that he has more scientific knowledge than most of us about how vision works, that we have all been deluded throughout our lives in thinking we could see physical things? I can't see why we should. The fact that he has a broader base of empirical knowledge about how vision works doesn't mean that we are obligated to accept his philosophical view that the facts about vision he helped unearthed were actually ones about relationships between ideas in a non-physical mind. For the same reasons, I can't see why I should think that just because neuroscientists know a lot more than most people about the physical events that occur in the brain, we should all accept their philosophical view that the facts they are unearthing are about relationships between actual mental states which are somehow part of the physical world. The background philosophical

interpretation of empirical data brought to one's investigation are not part of the dataset we should feel compelled to accept, even when it is shared by many of the experts in a discipline.

8.5: Objections to Agent Causation:

Many of my replies in this chapter have appealed to comments about agency discussed in chapter six. Many of the reasons given to think that the concerns at hand are illegitimate rested in the idea that agency works in a fundamentally different way that the sort of causation conceived of by opponents of interaction. However, many people have expressed other sorts of concerns with agent causation itself, or even with libertarianism in general. Given this, we should spend some time considering whether these objections undermine my appeals to agency in my replies in this chapter. Among the objections to agent causation are *a priori* concerns such as van Inwagen's rollback argument, and a widely accepted account of the causal role of motives, and *a posteriori* concerns deriving from experiments suggesting that decisions are made before we think they are. We will look at these in turn.

8.5.1 Van Inwagen's Rollback Argument

Peter van Inwagen has described an objection to agent causation in the form of the rollback argument. Van Inwagen's argument takes much of its flavor from the objection Ayer raised that was discussed in chapter 6. The argument invites us to look at indeterministic worlds in a very abstract way. If there is a world with a history, it will consist of a sequence of events. If any of those events are instances of agency, they will include agents with particular features at particular times. If the future is open, then there will be more than one possible future sequence relative to different behaviors performed by an agent, and different psychological states the agent could undergo. And if we "rollback" the universe to a given point in time, it is possible that it could go a different direction this time than it did the previous time. But, says van Inwagen,

none of this explains freedom. While asking us to consider a world of immaterial angels, he points out that these angels will undergo changes in their psychological states over time. He asks "And what is responsible for the way an angel changes its states with the passage of time? One possibility is that it is something structurally analogous to the laws of physics-- something that stands to our angels as our laws of physics stand to electrons and quarks."[252] In such a world, he claims that either those laws or deterministic or indeterministic, but in either case, we are not truly free. Like Ayer, he thinks that in either case, whatever changes happen aren't really up to us, and for the same basic reasons. Moreover, van Inwagen claims that in a world with real beings changing their features over time, there is no real option to explain changes other than something analogous to the laws of nature. He claims that merely recognizing the presence of beings with natures that change over time "we must imagine the inhabitants of this world as being subject to something analogous to the laws of physics."[253] Interestingly, van Inwagen says that he thinks there must be something wrong with this argument. He cannot however, see what it could be, and therefore concludes that metaphysical freedom is a mystery. Let's see if we can't unravel that mystery here.

Van Inwagen claims that there must be, in any universe with real things existing over time, something analogous to the laws of physics that explains what happens in this world. It's not clear in context exactly how he understands laws of nature. On the picture of laws of nature we have been using here, laws explain events by explaining how events with certain features have the power to bring about subsequent events, or one of a range of subsequent events, with various statistical likelihoods for those events. The idea is that the universe follows rules connecting events of certain types with events of other types in regular ways, and that this

[252] van Inwagen, p. 191
[253] ibid, 192

explains both why things happen and why there is a comprehensible structure to the universe at all. Contra van Inwagen's claims, the mere presence of possible futures isn't enough for laws of nature to explain events. Something in the nature of the event prior to the branching options must explain why there are those options and what the likelihood of each one is, otherwise there is not enough regularity to the structure for there to be genuine laws and thus they cannot explain the future. The limits of possible futures needn't be deterministic, but they need to have structure connected to the nature of the present events if the present event is to explain the next one. So let's look at the problem from this perspective. In this case, in order for our behavior to be caused by statistical laws, there would have to be some identifiable frequency of various outcomes that explains the likelihood of our behaviors given the current state of the events transpiring within us. Using the rollback scenario, it might be thought that it is possible to figure out, at least hypothetically, what those frequencies would be for a given psychological state. Simply rollback the universe an infinite number of times, and then see how often each future was chosen. Then you can see what the likelihood was for each possible future, and get a statistical structure that explains the subsequent events. In this version, the rollback argument would show that for any choice people have to make, there will be a frequency with which it is made. Since event causation can suffice to explain effects even if the law in question is statistical, the suggestion would be that the presence of a statistical regularity shows there is a law of nature being followed, and therefore omits the need for or value of the agent added to the picture by the agent causation advocate.

To begin to see what is wrong with van Inwagen's position, it should be noted that the fact that there will be a frequency with which behaviors are caused if the universe is "rolled back" is very different from the claim that the frequency is all there is to the cause of the

instances of behavior. To see this, let us consider an alternative version of the infinite rollback scenario. Suppose that someone were allowed to produce effects in two distinct universes with identical physical laws, and that time would roll back in each universe an infinite number of times to the same point. Suppose the person is deciding whether or not to eat some chocolate ice cream. I see no reason why one could not choose to eat chocolate ice cream 80% of the time in one universe, but only 60% of the time in the other. Even given the same psychological states at the time, there does not appear to be any structure to these states that necessitates that each possible future have an identical likelihood in each universe. If the statistical fact was an independent law of nature that explained the effects, then the frequencies with which the agent chooses each effect would have to be identical in otherwise identical universes. To put it another way, for the statistical regularity to be a sufficient cause of the behavior, the regularity would have to be the product of a pre-existing law of nature built into how the universe works and connected to the nature of the events that serve as causes. But as the dual universe alternative shows, the regularity isn't part of the fabric of the universe, it is a consequence of a set of choices the agent makes. This shows why the agent is relevant to the production of the behavior. What this shows is that the frequencies are *the product of* the exercise of active power in the agent in the universes, not *the explanation for* the effects in and of themselves. If the universe were rolled back a million times we would create the frequency with which events occur by exercising our inherent active power in various ways. The power is in us as an original source of causation, it is not a product of the features we have at a given time, and thus is not limited in its production of effects by any laws of nature, statistical or otherwise, connected to those features.

Against an advocate of free will who accepted event causation, the rollback argument might be successful. But van Inwagen's opponent in this case isn't someone who doesn't know

that repeated sequences have outcomes with various frequencies of occurrence, his opponent is someone who thinks that agents are the cause of their own behavior, and that the general structural fact about reality isn't relevant to its production. His opponent is someone like Reid, who says

> ...a free action is an effect produced by a being who had power and will to produce it; therefore it is not an effect without a cause.
>
> To suppose any other cause necessary to the production of an effect, than a being who had the power and the will to produce it, is a contradiction; for it is to suppose that being to have power to produce the effect, and not to have power to produce it.[254]

The exercises of power in an agent are not usually random. They are done for reasons that can be stated and understood. But neither are they exercises that obey or have to obey any specific statistical regularities. The psychological states that motivate us do not have the same explanatory structure as physical states that cause events, because the psychological states do not have a nature that is connected in a regular manner with the structure of the universe so as to produce a future in light of their mere presence. At the end of the day, van Inwagen's rollback argument can be seen as little more than a non-sequitor. It takes the leap from the tautological claim that there will be a frequency with which various futures occur to the claim that frequencies are all there are to causes, when this is precisely what the agent causation advocate denies. The rollback argument would only suffice in a world where there was no such thing as active power. In a world where there is causal power original in agents to bring about effects, statistical facts about frequencies that would occur in hypothetical scenarios don't suffice to provide an event-causal explanation of the effects of free will.

[254] Reid, 1969, p. 332

8.5.2: The Influence of Motives

One common objection to libertarian views of human agency comes from the influence of motives on human behavior. People commonly explain behavior by appealing to their motives or their reasons for action. The idea that we act freely suggests to many people that there is no explanation for our choices, or that they must somehow be random. Thus, A. J. Ayer says, "But if it is not an accident that I choose to do one thing rather than another, then presumably there is some causal explanation of my choice: and in that case we are led back to determinism"[255] In addition, given the relevance of motives to behavior, it is often thought that the motives eliminate the room for free will because the motives cause behavior in the same way that other events are caused. Indeed, many compatibilists claim that it is only because behavior can be caused by motives that we can really be responsible in any sense. This view about the influence of motives has a great deal of historical endorsement. Holbach,[256] Hobbes,[257] Hume,[258] and several other philosophers all seem to believe that human behavior is the inevitable result of the balancing of the influence of the motives we have, and that free will of the sort I have been describing is an illusion. Whether they redefine free will as Ayer does, or endorse hard determinism and Holbach does, this view of human behavior as the direct result of our motives is widely accepted among those philosophers in the naturalist tradition.

Despite its historical traction, I see nothing that can reasonably be said in favor of this view of the influence of motives once we accept the view that agents possess original, active power. Nor do I see in it any reason to deny the claim that we have such power. As Reid puts it:

[255] Ayer, 1969
[256] Holbach, 1770
[257] Hobbes, 1968, esp. Part I
[258] Hume, 2000, II, 2, iii

> Contrary motives may very properly be compared to advocates pleading the opposite sides of a cause at the bar. It would be very weak reasoning to say, that such an advocate is the most powerful pleader, because sentence was given on his side. The sentence is in the power of the judge, not the advocate. It is equally weak reasoning, in proof of necessity, to say, such a motive prevailed, therefore it is the strongest; since the defenders of liberty maintain that the determination was made by the man, and not by the motive.[259]

Motives certainly provide force or impetus toward certain actions, and typically provide us with a reason to act in one way rather than another. To assume that such motivational force is the entirety of the causes of human behavior, however, is to make us slaves to those motives in ways that, as a plain matter of fact, we simply are not. We can control, direct, and add force to our motives through the exercise of power within ourselves. "Is there no such thing as willfulness, caprice or obstinacy, among mankind?" Reid asks. "If there be not, it is wonderful that they should have names in all languages. If there be such things, a single motive, or even many motives, may be resisted."[260] People often speak of the power to resist motives; of people refusing to follow the force of reason, of people ignoring their own good judgment, or overcoming illicit desires. We praise or condemn them not for happening to have stronger contrary motives lying around, but for exercising the control they have over their behavior to overcome the force of these motives. In addition, we often act without any motive present. "I do many trifling actions every day, in which, upon the most careful reflection, I am conscious of no motive."[261] The essence of the agent causation view is that humans have abilities that everyone recognized we had until this odd theory of human behavior reared its ugly head. Perhaps some

[259] Reid, 1969, p. 287
[260] ibid, 285-6
[261] ibid, 285

hidden cause is lying at the heart of our behavior, but cluttering our mental landscape with hidden motives merely because someone insists they must be there is surely not the right way to go about understanding human nature. Attention to how we actually behave runs counter to this general theory of human action that has so little in its favor to begin with.

To help clarify this response, we should begin with a point raised by Thomas Reid. Reid begins his response to this concern by asking us what we should expect to find if the agent-causal story was true.

> Let us suppose, for a moment, that men have moral liberty, I would ask, what use may they be expected to make of this liberty? It may surely be expected, that of the various actions within the sphere of their power, they will choose what pleases them the most for the present, or what appears to be most for their real, though distant good. When there is a competition between these motives, the foolish will prefer the present gratification, the wise, the greater and more distant good.
>
> Now, is not this the very way in which we see men act? It is not from the presumption that they act in this way, that we reason from their motives to their actions? Surely it is. Is it not weak reasoning, therefore, to argue, that men have not liberty, because they act in that very way in which they would act if they had liberty? It would surely be more like reasoning, to draw the contrary conclusion from the same premises.[262]

In any rational being, we should expect that they would often act for reasons, and given the similarity in our natures, these reasons will often be comprehensible across persons. The empirical data, then, seems fully compatible with an agent-causal account of the influence of motives. Worse for this objection, if motives were causes in the same way that ordinary physical

[262] ibid, 291

properties were causes, then we would expect very different empirical data than we actually have. Presumably, if in each instance of behavior a motive serves as an efficient cause, then we should suspect that there is something about the cause that explains why it produced the effect it did. The standard account of why a particular motive caused a particular behavior offered by defenders of this objection has been that our behavior is always produced by our strongest motive. Given this, there would have to be some other means for measuring strength that would fit the empirical data for this account to be persuasive. As Reid puts it

> How shall we know whether the strongest motive always prevails, if we know not which is strongest? There must be some test by which their strength is to be tried, some balance in which they may be weighed, otherwise, to say that the strongest motive always prevails, is to speak without any meaning.[263]

On a straightforward account of how to measure the strength of a motive, however, the claim that our behavior is caused by our strongest motive is clearly false. The most natural measure of strength for a motive would be its phenomenal salience. But quite often we act on weaker, less vibrant motives than those we feel most strongly. The very essence of self-control is an ability to overcome strong motives in favor of prudential or moral concerns. A view according to which the most strongly felt motive always wins would be one according to which there were never any successful efforts to overcome strong desires. But this is obviously false. Similarly, one could contend that the strongest motive is the one we have most reason to act on. But, once again, this is clearly empirically false. People often act rashly and irrationally, ignoring the best reasons they have to act. On the two most obvious accounts of how to measure the strength of motives, the empirical data fits the case for free will, not the case for determination by motives alone. In the absence of a well-developed account of how to measure the strength of motives and a clear

[263] ibid, p. 268

and persuasive reason to believe that such an account leaves out the possibility of free choice, there is no reason to accept this account of the causes of human behavior.

8.5.3: The Timing of Choices and Causes

One of the most popular objections to free will among the educated general public stems from studies done by Benjamin Libet[264] and later improved upon by Soon, Brass, Heinze, and Haynes[265] suggesting that our decisions are actually made before the time we think we are making them. This has fed the idea that free will is an illusion. To see the merits of this, we should look at the results. The original study by Libet involved a participant deciding when to push a certain button. The participant was to state when they had made the decision, and typically reported it as having occurred right before the action. But the study showed a ramping up of brain activity just prior to the time the decision was made. Some people think this shows that the brain was already producing the result prior to the decision. In the other study, the situation was more complex. Participants were asked to push one of two buttons, and to recognize a certain letter on a screen at the time they decided to push the button. In this experiment it was shown that activity in the brain that could predict the outcome began to occur around seven seconds before the participant pushed the button. The activity in the brain was used to determine which button was pushed long before the participant was conscious of having made a decision. This study has convinced many researchers that we are often dramatically wrong about our own control over our actions and about the presence of free will.

There are a few minor concerns with how these experiments are done. In the Libet case, the most obvious problem is that the recognition of having made a decision and the reporting of it takes some small amount of time. There is a feeling of resolution that occurs once someone

[264] Libet, 1985
[265] Soon, Brass, Heinze, and Haynes, 2008

decides, and it is quite possible that the participants are reporting on this feeling, rather than on the final moment of decision-making. The second problem is with the fact that the predictive ability of the tests were nowhere near perfect. The fact that one can predict with some accuracy what will occur does not show that we are unable to choose. The largest problem with these experiments, however, is the failure to appreciate that making a decision takes time. The experimenters seem to assume that the decision is spontaneous and instantaneous. Having been primed to make a decision, however, it is highly unlikely that the entire time frame of decision making is identical with the resolution to act. When we are deciding what to do, we consider possibilities. In addition, when we decide to act on a present desire and need not engage any conscious effort to evaluate or prevent its occurrence, we typically allow it to be effective. Moreover, no one who has ever actually made a decision should think that the entire process occurs at the level of consciousness. We sleep on things, go on with our day, and think about other stuff while in the process of deliberation all the time. Not every aspect of decision-making rises to the level of conscious attention. While exercises of constant control are required for difficult choices or ones where we need to resist a strong desire in order to do otherwise, in those cases we likely need to control activity much longer and consider various options, which could result in multiple sites in the brain being activated in accordance with the various options. But, as Reid points out, "to yield to the strongest (motive), [one] needs only to be passive." A simple decision about what button to push is unlikely to be one we need to agonize over or change our minds about once we start down a certain path. So, the presence of data correlated with the onset of deliberation and the presence of data suggesting that a motive that we choose to eventually act on is present is not the same as the presence of a decision. This is particularly true of the more complex experiment, in which one has to make two decisions; first, which button to push, and

second, when to push it. The process of making a decision and the result of reaching one don't occur instantaneously. The fact that a feeling of having settled on something only occurs once the process is allowed to be finished isn't surprising. None of this shows that one did not have it in their power to change their mind or go in a different direction prior to the completion of the deliberative process.

8.6: The Influence of Nature and Nurture:

Baron d'Holbach claimed as early as 1791 that one of the two primary reasons people reject determinism is because they are ignorant of the causes of their own behavior. He believed that once psychology developed to a point where we unearthed these causes we would see that free will is an illusion and that human beings are just another part of nature[266]. More recently, Thomas Nagel has claimed that the empirical data is finally vindicating this prediction and that

> …as the external determinants of what someone has done are gradually exposed, in their effect on consequences, character, and choice itself, it becomes gradually clear that actions are events and people things. Eventually nothing remains which can be ascribed to the responsible self, and we are left with nothing but a portion of the larger sequence of events, which can be deplored or celebrated, but not blamed or praised.[267]

One of the primary reasons people believe that the room for agency diminishes greatly in light of these observations grows out of the increasing understanding we have of the relevance of genetics and environment to behavior. The "Nature vs. Nurture" debate typically assumes that anything that cannot be explained genetically must be explained by environmental factors. Both of these features are demonstrated to have varying levels of significance on numerous features,

[266] see Holbach 1770
[267] Nagel, 1979

including features of character and facts about success in one's life. If every human action can be fully explained by a combination of their genetics and their environment, then agency really would disappear in the course of physical events. Fortunately, however, the evidence we have is far from such a dramatic conclusion, and has been seriously overestimated by advocates of this line of reasoning.

Before evaluating the strength of the case for genetic and environmental determinism, it is important to understand what the data says about the influence of these factors. The best source of data we have on the influence of these factors come from twin and adoption studies. Studies done on identical twins raised in separate environments allows one to measure the effects of genetics while, to an extent, factoring out the influence of shared environment. Studies done on fraternal twins raised in the same household can maximize the influence of shared environment while controlling, to a certain extent, for genetics. Studies done on adopted children raised in the same environment as their siblings can further control for genetics. The results of these studies have shown a significant influence of genetics on a large number of characteristics, a significant short-term influence of shared family environment, and a minimal influence of shared family environment over the long-run. Evidence of some influence on things as varied as political and religious values, how successful someone is, how intelligent someone is, and even how happy someone is has been found.[268] The relevance of various outside factors on our character and our lives is undeniable.

That said, the relevance of these factors on our lives is far from deterministic. By far the most relevant influence on our lives from these sources is genetics. Genetic effects have been shown to have a significant influence on a wide array of traits. However, while the influence is clearly there, genetic determinism is not in the offing. The relevance of genetics is rarely higher

[268] For a description of some of these influences, see Caplan, 2011, ch. 2-3.

than 40-50% for any given trait one would have thought was influenced by free will. Although intelligence, success, income, and even things like political affiliation are correlated with genetics, the effect is far from deterministic. As Bryan Caplan puts it, "Attacks on genetic determinism are a witch hunt. Witches don't exist, and neither do genetic determinists who understand twin and adoption research."[269] In addition to the fact that genetics don't wholly determine our lives, the aspect of environment most easily studied, namely the shared environment kids have growing up, has far less effect that people think. The relevance of shared environment for most features that would be relevant to this debate are between 0-10%, at least over the long-run. That leaves great ranges of our lives where the effect of nature and nurture are anything but deterministic. Often those who are engaged in this research assume the rest of the difference must be explained by what is called "non-shared environment," bringing a deterministic assumption to that conclusion. As Caplan puts it, however, "Another name for 'unique environment' and 'nonshared environment' is 'none of the above.' If free will isn't none of the above, what is?"[270] The result of these studies, then, indicates that there is ample room left over for free will to have a significant impact in our lives. The fact that our nature has an influence on our lives or that relatives often exhibit similar skills and levels of success isn't news to anyone, even though it is demonstrated more effectively now. The fact that upbringing has little effect on long-term success is news to people, but it isn't news that disconfirms free will, it's news that suggests that our lives are more the product of our own choices than we previously believed. The role of events in our lives is, of course, extensive. But the fear that they have

[269] Caplan, 2011, p. 81
[270] ibid, 82-3

subsumed the role of agency to an "extensionless point,"[271] as Nagel claimed, is far from vindicated by the actual data.

8.7: The Conservation of Energy

Perhaps the most popular argument from the current state of physics to the denial of dualism is an appeal to the conservation of energy. It has been claimed by several people that if there were immaterial minds producing physical effects, this would violate this law of physics. Daniel Dennett, for example has offered this argument, stating:

> A fundamental principle of physics is that any change in the trajectory of any physical entity is an acceleration requiring the expenditure of energy, and where is this energy to come from? It is this principle of the conservation of energy that accounts for the physical impossibility of 'perpetual motion machines,' and the same principle is apparently violated by dualism.[272]

Despite its popularity, the claim that interactive dualism would violate the law of conservation of energy is simply false. The law of conservation of energy states that for any closed system, the total quantity of energy in the system will remain constant through any change. In its most basic terms, what it says is that if you have a certain system of energy, and nothing influencing it from the outside, the energy in it won't magically disappear, and more energy won't magically appear. This is just good common sense. But it has nothing to do with dualism. First, there is no clear reason to think that the influence of an immaterial mind on the physical world would require a change in the quantity of energy in the physical world. I see no *a priori* reason to have any opinion about whether the changes in the distribution of energy that would result from the mind's producing a physical effect would produce a change in the *total quantity* of energy in the

[271] Nagel, 1979
[272] Dennett, 1991, p. 35

system in the process. More importantly, however, if there are immaterial minds influencing the behavior of matter, then the physical world simply isn't a closed causal system. Since the conservation laws are written to apply only to such systems, it can't be a violation of them for minds outside the system to affect the behavior of the energy within it. So dualism, viewed as including an influence from a non-physical mind on the physical world, simply can't conflict with law of physics when those laws are couched in a conditional stating that they only apply to closed systems. But, of course, all laws of physics have to have such a conditional built in. Physics doesn't study anything outside the physical world, so its laws require the assumption that there isn't anything else interfering with the processes they study, and are therefore only intended to apply under such an assumption.

8.8: The Causal Closure of the Universe

I claimed at the end of the last section that dualism requires that the causal processes that occur within the physical world aren't causally closed. But it is often asserted against dualism that it has to be wrong precisely because we already know that the universe is causally closed, and so there is no way an immaterial mind could influence it. Papineau calls this argument "The Causal Argument," and presents it as follows:

(1) Conscious mental occurrences have physical effects.

(2) All physical effects are fully caused by purely *physical* prior histories.

(3) The physical effects of conscious causes aren't always overdetermined by distinct causes.

(4) Therefore, conscious mental occurrences must be physical events.[273]

This is a fairly simple and direct argument for the view that the mind is a physical thing. If mental events have physical effects and only physical events can have physical effects, mental

[273] Papineau, 2002, p. 17-18

events must be a type of physical event. The obvious type of physical event to suppose they must be is some sort of event in the brain, since these are what seem to be involved in the production of the effect.

The account of mental causation I endorse involves a straight-forward denial of premise (2). Premise (2) expresses what is typically called the "causal closure principle." According to this view, there are no outside influences that interfere with the sequence of events that occur in the physical world. Some philosophers seem to believe that this principle is part of contemporary physics. As Papineau discovered when he looked into it, however, this simply isn't true. So, if this claim isn't a part of physics, then why is it so popular? I believe that the reason for its popularity lies in the fact that something like causal closure is a guiding assumption of science. Scientists assume that they can find the answers to questions about why things are happening in the physical world through empirical observation and experiment. They assume this in part because they assume that there will not be anything else influencing the events they study. The fact that causal closure is a guiding assumption of a successful discipline suggests to many people that we should accept the principle.

It makes sense for scientists to assume closure holds when they are engaged in their work. After all, if they believed that the answer couldn't be found under this presupposition, they wouldn't be using scientific methods of study to understand it in the first place. But the fact that practitioners of empirical sciences assume something in order to believe that their work can account for what happens doesn't show that their methods actually can account for it. The study of history, for example, requires that the event studied happened in the past, and that there be some remaining empirical evidence about what happened to help us understand the past properly. So unless you think an event is past, and that there is some record of it, you can't think you can

study it as a historical event. But this hardly means that there is no present, and there won't be a future, or that events that didn't leave traces we can find never happened. Only on the additional assumption that your discipline can answer everything would anyone ever assume that the guiding assumptions of their own discipline were decisive of the limits of ontology. This seems to be what Alvin Plantinga has in mind when he says that "the doctrine or dogma of the causal closure of the physical is not a deliverance of current science: it is more like an article of faith or perhaps a pious hope on the part of materialists."[274] Materialists hope that science can find all of the answers, since on their view, they should be able to do so. If this is true, then causal closure surely holds. But, as BonJour claims "to argue for the truth of materialism or for a strong presumption in favor of materialism by appealing to the principle of causal closure is putting the cart in quite a flagrant way before the horse."[275] Whether or not a discipline's methods can provide explanations for everything isn't a question for each respective discipline, it is a question for metaphysics.

8.9: The Inductive Case for Physicalism

I claimed at the end of the last section that the ability of a discipline to answer all questions about reality was a question for metaphysics. However, it is a question that many metaphysicians answer in the affirmative when it comes to science. Given this, we need to examine their reasons for endorsing such a view. For many, the reasons they give seem to be the ones we have already looked at. There are also some who are motivated by considerations of method. We will return to those considerations in the next chapter. But I think the most compelling reason typically given in defense of this view is that science has been so successful in the past that we should think it can answer everything in the future. As David Lewis expressed

[274] Plantinga, 2006, p. 127-8
[275] BonJour, 2010, p.6

the idea in 1966, "in the future [dualism] will be rendered incredible by the continuing advance of physicalistic explanation."[276] This inductive claim and corresponding faith in the future of science is common in philosophy of mind. Thus Armstrong writes

> If we consider the search for truth, in all its fields, we find that it is only in science that men versed in their subject can…reach substantial agreement about what is the case. It is only as a result of scientific investigation that we ever seem to reach an intellectual consensus about controversial matters.[277]

Earl Connee and Ted Sider similarly claim in defense of determinism that "Our belief in determinism is reasonable because we have all seen science succeed, again and again, in its search for the underlying cause of things,"[278] going on to say that "the threat to freedom comes when we realize that this march [of science] will eventually overtake *us*."[279] Papineau goes into more detail to explain his defense of causal closure as similarly inductive. He claims that the reason the causal closure principle has only recently been decided upon is that it is only in the last fifty years where more and more physiological study has been undertaken that the empirical evidence has been sufficient to justify this claim. According to Papineau in the last fifty years "a great deal became known about biochemical and neurophysiological processes, especially at the level of the cell, and none of it gave any evidence for the existence of special forces not found elsewhere in nature" (275). The claim is that if there were special mental forces having effects on us then there should be empirical evidence to this effect, but since no such evidence has been found, it is reasonable to assume that there are no special mental forces. Papineau admits that new discoveries might be made and science might change to admit the existence of such forces.

---

[276] Lewis, 1966, p. 25
[277] Armstrong, 2005, p.138
[278] Connee and Sider, , p.114
[279] ibid, p. 114, emphasis original.

He also admits that there is no reason to accept materialism in the absence of the completeness of physics. But he thinks that the empirical case in favor of the principle is solid and he can "see no virtue in philosophers refusing to accept a premise which, by any normal inductive standards, has been fully established by over a century of empirical research."[280]

To begin our response, we should recognize what is necessary for an inductive inference to be legitimate. As Michael Huemer points out, "As a rule, the strength of an inductive inference depends upon the degree of similarity (with respect to properties other than the property ascribed in the conclusion) between the thing(s) that the premises concern and the thing(s) that the conclusion concerns."[281] One consequence of this is that induction is far less reasonable when it goes from easy cases to hard ones, especially where we have reason to think it will fail. If one is taking a test with progressively more difficult questions, it would be quite unreasonable to think that getting the first set right shows that you will never miss one. The fact that science has explained a large number of things without appeal to outside forces doesn't show that in cases where we initially had the most reason to think there were outside forces we still won't find any. The mind has always seemed to be something fundamentally different in kind from other sorts of things in the world, so assuming science will one day explain it on the basis of induction is highly dubious. As BonJour puts it,

> inductions are always questionable when the conclusion extends to cases that are significantly different from the ones to which the evidence pertains, and even most materialists will concede that conscious phenomena are among the most

[280] Papineua, 2002, p. 227
[281] Huemer, 2008, p. 228

> difficult—indeed, seemingly the most difficult of all—for materialist view to handle.[282]

In addition to this, the sorts of causes that science has replaced over time have been speculative causes. For example, people believed that God explains why there appears to be design in nature on the grounds that it was the most plausible speculation they could come up with at the time. But when it comes to free will and mental causation, our selves as the cause of our behavior is part of the introspective data, not an idle speculation. Therefore, the assumption that it will be overcome like speculative positions have been in the past appears to be unwarranted. The past success of science does not show that it will succeed in areas where we have the most reason to think it will fail because this case is too different from the other examples to think that induction is warranted.

But what about the specific data Papineau appeals to? Papineau claims that recent empirical data has confirmed this position because it has failed to find immaterial causes where we should have found some. This argument has more force if it is a correct description of the situation. However, I see little reason to accept the claim that scientists would have found, recognized, and acknowledged the need for immaterial causes by this point in the research if they existed. Although empirical research has progressed, given the complexity of the brain, it is far from complete. It is not as though we almost have it all figured out, and we're running out of places to find influence from an immaterial mind. More importantly, the interpretation of the data seems to be filtered through one's metaphysics, not the other way around. As L. A. Paul points out,

> It is a mistake to think that one should first study science and then use it as the guide to one's metaphysical conclusions. Science does try to construct an overall

[282] BonJour, 2010, p. 5

> picture of the world. But scientific theorizing usually uncritically assumes the very organizing principles and deep general truths that metaphysics is concerned to prescriptively develop and understand.[283]

We have seen this play out in debates in neuroscience as well. Most neuroscientists have not recognized the need for immaterial causes. Some neuroscientists who are dualists, however, have come to the conclusion that they are needed. Popper and Eccles claimed that one could not explain human action without immaterial mental causes,[284] and Eccles went on to discuss the need for "psychons" to fill in gaps in the causal explanation[285]. Whether or not they are correct is less important than the fact that one's philosophical presuppositions tend to shape how one interprets the significance of one's findings. Given the fact that most neuroscientists start out as physicalists, the fact that they have not thought that their current lack of understanding proves that there need to be immaterial causes doesn't show that there aren't any. Science is messier than people tend to think. High statistical correlation is typically seen as sufficient for drawing conclusions, even when these stats are compatible with dualist views. In addition, when someone find unexpected data, they will initially think they screwed up, thanks to their underlying assumptions. As Khun would put it,[286] only in a time of crisis would scientists start to view irregularities in data as evidence of faulty underlying assumptions, and look for a paradigm shift in accounting for the anomalous data. Given this, we can't take much from the state of neuroscience at this point in the discipline's history. The conclusions reached seem to presuppose what philosophers seek to establish from it, they seem to vary according to those

[283] Paul, forthcoming
[284] Popper and Eccles, 1977
[285] Eccles, 1994
[286] See Kuhn, 1996

philosophical presuppositions, and the data is no where near complete enough to know if the remaining gaps in explanation have immaterial causes or not.

The causal arguments against physicalism have been by far the most persuasive. People often claim that there is something deeply troubling or mysterious about non-physical minds interacting with physical bodies. However, these complaints are often under-specified, and when they aren't, they typically depend on limited and often question-begging accounts of causation. The inadequacy of such conceptions makes it hard to see why dualists have been so vexed about these claims. In addition to these concerns, many people point to the progress of science either in specific findings or in generalities as a reason to think that anything not grounded in science is somehow suspect. But science depends on the same sorts of uses of reason and is subject to the same influences of presupposition as any other intellectual discipline. Its successes are impressive, but not reason to forgo all other avenues of inquiry. In the absence of a pre-existing commitment to naturalism, the causal objections to dualism hold little force, and certainly don't override a reasonably strong presumption in favor of the dualist position established by the earlier chapters of this work.

Chapter 9: Non-Causal Objections

Overview

In addition to the causal arguments against dualism, there are a few objections focusing on more general considerations. Perhaps the most popular non-causal objection to dualism is an appeal to Ockham's razor. Parsimony considerations are widely appealed to in philosophy, but less attention is paid to what justifies such appeals. Never having had a taste for desert landscapes, I have rarely found these appeals compelling. Fortunately, recent work has been done trying to uncover what lies behind appeals to parsimony. Michael Huemer has persuasively argued that the foundations that justify appeals to parsimony in scientific contexts fail to extend to philosophical discussions such as the debate between physicalism and dualism. With respect to parsimony, I believe his replies suffice. However, I believe there is a related formulation of simplicity considerations that is intuitively plausible. I argue that on this formulation, dualism actually comes out as the simpler view.

Second, there have been several arguments, most prominently put forth by Daniel Dennett, arguing for the unreliability of introspective evidence. Dennett's arguments tend to focus on cases of introspective illusion, which he uses to try to support his view that most of conscious experience is simply a magic trick. These arguments are often set against a foil of a variety of dualism that holds tightly to the idea that introspection is infallible. Against these views, perhaps Dennett could score a victory. However, with the progress in epistemology toward fallibilist accounts of justification, the objections seem weak. Indeed, they seem to serve the same dialogical role as the cases of perceptual illusion served in undermining our reliance on perception in the early modern era. The parallels between the move to idealism from such examples and Dennett's efforts to move to physicalism from his are illuminating and instructive

since they both seem to fail for the same reasons. A rather different attack on introspection comes from Paul Churchland. According to Churchland, the dualist is giving a sort of legitimacy to introspective data that we shouldn't hold for any source of data, and needs to justify this double standard. I explain why the relevant data in the case of introspection is different in kind from other sorts of data and how this changes the question of its significance.

A third sort of objection arises from intuitions about the nature of persons. Although I have argued that common sense is committed to dualism, there are some natural things people say that at least appear to suggest materialist intuitions. This creates two problems for dualists. First, if one thinks such comments are more reflective of common sense, then there is a challenge to the view from within common sense itself. Second, if there is pervasive inconsistency in our view of persons, then the legitimacy of any particular aspect of that inconsistent set would be questionable, and loses much of its presumptive support. I tackle these concerns here by providing a dualist account of persons that accommodates both sets of intuitions.

9.1: Physicalism and Parsimony

Most of this work has been targeted at convincing mental realists that they should endorse substance dualism. However, those unpersuaded by the earlier epistemological work may be unconvinced that mental realism of the sort I advocate is the right way to go in the first place. One of the most common reasons appealed to in favor of eliminativist or error theoretical views is Occam's razor. It is widely accepted that simpler views should be endorsed over more complex views in most circumstances. Since physicalists have fewer things in their ontology, they may have a theoretical advantage here. Those unconvinced by the arguments for the view that physicalism is an error theory will likely appeal to parsimony to support their view, and given the long-standing appeal of Ockham's razor, some may be inclined to use it to justify

abandoning mental realism even if they accept the things I have said about our understanding of the mind. I think this would be the wrong way to balance the considerations, even if parsimony considerations counted strongly against dualism. However, I don't think that parsimony actually favors physicalism in the first place. On most accounts of the value of parsimony, physicalism fails to count as a simpler theory than dualism. On the account of parsimony's value that I find most natural and plausible, dualism actually counts as a simpler theory than physicalism, and therefore advocates of Ockham's razor have more reason to accept it, not less.

Michael Huemer has recently surveyed the various accounts of the value of parsimony as a theoretical consideration. Huemer identifies four main defenses of parsimony, and argues that none of them apply well to philosophical questions, including the debate between physicalists and dualists in philosophy of mind. The first theory he examines claims that the success of science provides good reason to accept the epistemic principles followed by science, including parsimony. Huemer points out that this account may give us reason to value parsimony, but it doesn't explain why parsimony is valuable. In the absence of a further explanation for why parsimony is valuable in science "the assumption that this reason applies equally to philosophy as well as to science is something of a leap of faith."[287] The second defense he considers is called the "boundary asymmetry account." According to this view, parsimony is underwritten by the fact that there is no limit to how complex a theory can be, but there is a limit to how simple a view can be. Given this, to have a combined probability of 1 for the set of possibilities, one must have a decreasing level probability while the complexity increases to infinity. While this account works well for views that have an infinite series of increasingly complex options, it fails to hold for situations where the options are more limited. As Huemer puts it,

[287] Huemer, 2008, p. 229

> the argument [for the boundary asymmetry account] does not apply to disputes about the nature of a given class of entities, or about the relationship between two given kinds of entity, when there are only finitely many contrasting alternatives. In particular, it does not apply to the physicalism/dualism debate, because physicalism and dualism are not naturally viewed as successive steps in some infinite hierarchy of theories.[288]

The third account Huemer names "the numerousness account." According to this view, the main reason that simpler views are more likely to be true is that there are typically fewer of them. The view suggests that to ascertain *a priori* probabilities, we should first divide the probability space into sections that reflect their complexity, and then assign roughly equal probability to each view within the level of complexity. Since there are typically more available theories for higher levels of complexity, equally dividing the areas between views will lead to a lower probability for complex views than for simple ones. As Huemer points out, however, when it comes to issues like physicalism and dualism, there is no clear way to divide the views into categories based on complexity. It's unclear what we are counting when we divide the theories into categories, or how many theories to place into each group. This makes the application of parsimony considerations difficult to evaluate on this account.

The final account Huemer considers is the likelihood account, and is the one he seems to endorse. According to this view, the reason to prefer simpler theories is that they obtain greater support from the occurrence of predicted data than complex theories do. Because a simple theory "is compatible with a smaller range of data, it assigns a higher average probability to the possible sets of data which it allows. "[289] Given Bayes' Theorem, this means that the subsequent

[288] ibid, 229-30
[289] ibid, 223

observation of data compatible with both theories supplies more support for the simple theory. To put it another way, simple views can be more easily refuted by incompatible data, so the fact that the view was not refuted by the data is more suggestive of its truth than a view that would have been harder to refute. According to Huemer, the debate over physicalism and dualism doesn't naturally fit the patterns of debates that can be adjudicated on these grounds. For one thing, physicalists and dualists both have significant leeway in accounting for empirical data. We saw this point in chapter 8 while discussing the ways that dualists can accommodate the empirical data, and how it compares to physicalist accounts. In addition, the relevant data that may be able to decide the matter, such as a comprehensive account of all the causation in the brain which either excludes the possibility of or requires the existence of intervening mental causes, has not been found. Since on this account, simple theories aren't *a priori* more probable than complex ones, but rather obtain greater support from the data we find, non-existent data isn't relevant to the probability that any given theory is true. So, on this account of the basis for parsimony, Huemer claims that we lack the necessary data for parsimony to be a relevant consideration in the debate. While this is true of most things, there do appear to be at least two area of investigation where this doesn't hold, which we will turn to now.

According to Daniel Dennett,[290] what happens in the brain when conscious experiences occur is that various parts of the brain are actively seeking what he calls "fame." What he seems to mean by this is that various parts of the brain, each containing a (or, I suppose he would say "each of which is a") separate feature of conscious experience, are competing to direct the future activity in the brain. According to Dennett, these features of the brain can be aware of one another, and a feature is "famous" so long as it draws enough attention from the other parts to get

[290] See, for example, *Sweet Dreams*, (c) 2005 MIT press, esp. Ch. 6

its way. According to Dennett, our behavior is caused by this process rather than through any executive decision-maker found in the brain that evaluates and chooses from among these features. In addition to a lack of an executive decision-maker in the brain, there also doesn't appear to be a unifier of consciousness in the brain. During perception, more than 30 different areas of the brain are involved in trying to process visual information. However, there doesn't appear to be any location in the brain where all of this information comes together to form the coherent visual experiences we have at each moment. Mental realism about a "self" in the traditional sense of a conscious agent would seem to require, at least for physicalists, that there be some part of the brain that is aware of all the others; a place where, as Dennett would put it, the Cartesian theater would come together in experience, and where we would get to decide which player in the theater got to determine what happens next. So according to Dennett, versions of physicalism that endorse the common sense view that there is a self or executive decision-maker in the brain should predict the existence of some part of the brain that is more in control of behavior than the rest, and that unifies experience for us. According to Dennett, there is no such place in the brain.

Dennett draws from this fact the conclusion that the Cartesian Theater is just a magic trick, and that what really happens is that the parts of the brain seek attention only from one another, and that there is no one in charge of which feature wins. The friend of substance dualism, however, could draw a very different conclusion from this. According to dualists, the self is a simple, unified, immaterial substance that gets to be the audience and the director of the Cartesian Theater. While an empirically informed dualist would think that there would be activity in the brain corresponding to conscious experience and to motivation, they would not expect there to be an executive decision-maker *in the brain* that is running things, or that is

aware of the overall product of processing information into experience, since we wouldn't view the self as a part of the brain in the first place. Therefore, looking at the data Dennett offers, the dualist account is simpler in the way that matters: it predicts, or at least is compatible without alteration with, the data we actually discovered about the absence of a self in the brain. So, with respect to at least this one area of inquiry in the debate, dualism appears to be better supported than physicalism given the considerations that underwrite parsimony.

While I think Huemer's responses are otherwise accurate for the views he considers, there is another epistemic principle which is quite plausible on its face and that could serve to underwrite simplicity of a theory as a virtue. This principle is as follows: don't make stuff up. As a general rule, making stuff up usually doesn't lead to the truth. Sometimes, in order to have a complete theory, you need to go beyond what you know and speculate about certain matters. But as a general rule, speculation is a poor source of evidence. Even experts in a field are usually inaccurate in their initial hypothesis, and as we saw in chapter 1, even when they think they've gotten it right based on their research, they're still usually incorrect. It would therefore make sense to believe that theories that have to make up less stuff to account for the data are more likely to be true. Other things being equal, sets of propositions that go beyond the data are less likely to go wrong the smaller the set is. Since what is being "made up" in this case are proposition, simplicity is a matter of elements of the theory rather than of ontology. However, claims that posit new entities to account for the data would be a special case of claims that go beyond the available data and amount to speculation.

To understand this idea, we should look at Reid's explanation of the Newton's version of Ockham's razor. Newton uses a version of Ockham's razor as one of his "Rules of Philosophizing." That principle goes as follows: "Rule I: We are to admit no more causes of

natural things than such as are both true and sufficient to explain their appearances."[291] Reid viewed Newton's success in science as grounded in laying down principles of common sense[292] as a foundation for pursing investigations in natural philosophy. He said of Newton's Rules of Philosophizing "though they have not the same kind of evidence that mathematical axioms have; yet have such evidence, that every man of common understanding readily assents to them, and finds it absolutely necessary to conduct his actions and opinions by them, in the ordinary affairs of life."[293] Reid viewed proper science as a discipline that followed general principles of common sense, and refused to speculate beyond what those, in addition to careful observation, offered. According to Reid "we may learn something of the way in which Nature operates, from fact and observation; but if we conclude that it operates in such a manner, only because to our understanding, that appears to be the best and simplest manner, we shall always go wrong.[294]" As Reid understands it, Newton's principle is grounded in the idea that speculative additions to our worldview are grounded in a method of belief formation that has historically proven to be very unreliable, and that therefore should be limited as much as possible. Speculating any more than is necessary, and doing so in any case where the data fails to offer clear guidance for careful induction is almost certain to produce results that will eventually prove to be inaccurate.

This basis for a limitation on speculative views is grounded in the structure of evidence. Theories exist in order to explain certain sets of data. Given this, the evidence we have for theories must derive from the data itself. Because of this dependency, theoretical claims are less likely to be true than the things they are supposed to explain. The more speculative and independent of the data a theoretical claim is, the less it draws from and stays connected to

[291] Newton, Principia Mathematica
[292] See, e.g. Reid, 1997, p. 12, 2002, p. 40-41
[293] Reid, 2002, p. 40-41
[294] Reid, 2002, p. 531

sources of evidence, and so the less likely it is to be true. This explains why anything other than careful reasoning from a sound starting point is likely to lead to error. The more theoretical the claims one uses to explain the data, the more likely it is that one's efforts will involve the sort of speculative reasoning that is highly unreliable. Therefore, the more theoretical additions there are in a theory, the more likely it is that a theory is false. In the most basic terms, if you make stuff up, it's not likely to be true, and the more you make up, the more likely it is that some of it isn't true. Given this, we should try to avoid making stuff up whenever possible. This seems to be a sound theoretical basis for preferring simpler theories to more complex ones.

When it comes to ontological parsimony, there is something initially puzzling about it, which lead me to reject it as a principle for a long time. The idea that there should be more or fewer things in the world seems to suggest that there is some magic, special number of things or types of things that there can or should be. That idea, on its face, sounds silly. When it comes to initial data sets, I still think this is correct. We should be glad we can experience as much of reality as we can, and trying to hamstring ourselves by ignoring or denying some it seems odd. However, when it comes to ontological additions in one's theorizing, the same line of thought that justifies theoretical simplicity of speculative claims justifies ontological simplicity of posited objects. Existential claims accepted for theoretical reasons are among the speculative claims we should be skeptical about, and for the same reasons. Therefore, when comparing theories for simplicity, numbers in the data set are irrelevant: what matters is how many additions there are over-and-above the initial data set that we include in our ontology. Dennett openly accepts error theories or eliminativist views of several aspects of the mind. I have argued earlier that physicalism in general is best understood as an error theory. The reasoning here about parsimony leads us to the first premise of an argument against such views:

(1) Considerations of parsimony only apply to theoretical posits, not to elements of an initial data set.

Of course, this premise requires that we have some way of formulating data sets in the first place. When determining what goes into a data set, the same reasons supporting an anti-speculative line of thought suggest that the elements of our data set in our efforts should be based in an ontology that didn't require speculation to obtain. Things we are aware of independently of a need to engage in speculative reasoning don't have the problem of dependence that makes their inclusion in our worldview problematic. Therefore, among the things that it is reasonable to include in our initial data sets in our metaphysical efforts to understand the world as a whole are the elements of our common sense worldview. Or, in other words:

(2) Apparent objects of awareness in ordinary exercises of human cognition aren't theoretical posits.

Together, our first two claims suggest that a properly parsimonious ontology will accord with common sense. This fact poses a serious problem for the use of parsimony in the arguments of error theorists. To see exactly how, we should look at how error theories work.

In an error theorist or an eliminativist view, some commonly accepted claim about what exists in the world is rejected. In either case, there is a need to explain why it is that so many people embraced the existence of the things that these theorists deny. Typically, these efforts focus on the idea that we have suffered some sort of illusion in our ordinary experiences that made us embrace this false view. On this story, we have to remove from the data set a number of items and add to our ontology a set of cognitive illusions. Since belief in the existence of these illusions as well as an appropriate cause for these illusions will be additions to the data set

grounded in theoretical reasoning, they will count as theoretical posits, and therefore will make the theory more complex. That is,

(3) Massive cognitive illusions that get us to mistakenly believe in things that aren't really there and the causes of those illusions are both theoretical posits.

Putting these points together, we get the remainder of the argument:

(4) So, theories that include massive cognitive illusions but exclude apparent objects of awareness in the ordinary exercises of cognition are more theoretically complex than theories that include a larger ontology but no such theoretical posits.

(5) Error theorist views and eliminativist views require positing massive cognitive error (and, in the case of error theories, replacement explanations) at the expense of apparent objects of awareness.

(6) So, error theories and eliminativist views are less parsimonious than realist views in any area that deals with apparent objects of awareness in the ordinary exercise of our cognitive abilities.

Since error theories and eliminativist views are often heavily motivated by parsimony, this argument seriously undermines their credibility. Many people express a preference for desert landscapes, and these views are often seen as the purest expression of this attitude. Most advocates of this preference claim that it is more than a mere aesthetic bias, because it is underwritten by considerations of parsimony. The proper use of considerations of parsimony, however, doesn't lead to a desert landscape, it leads to the rich ontological tapestry of common sense. Fans of desert landscapes, therefore, will have to either find another basis for their preference, or acknowledge its evidential irrelevance and stop appealing to it in support their favored views.

Parsimony, on this account, is a matter of refusing to add more things into your ontology than the data already includes. The data set should include the apparent objects of awareness in all of our ordinary exercises of our intellectual powers. One noticeable thing about this approach to parsimony is that it does not justify a reduction in ontology for its own sake. Therefore, the question of the principle's application to a philosophical problem depends upon the ontology of the data set prior to any speculative additions. If someone is presupposing a physicalist metaphysic or a strongly empiricist epistemology, this would likely lead them to believe that any non-physical entity must be a speculative addition. However, on a more ecumenical philosophical approach to our means of understanding the world, we find a broader ontology in our initial set, and therefore many things that are traditionally viewed as speculative additions are more naturally seen as part of the original data set. What Reid counsels for those seeking to adhere to such a picture is to "admit the existence of what we see and feel as a first principle, as well as the existence of things whereof we are conscious; and to take our notions of the qualities of body, from the testimony of our senses, with the Peripatetics; and our notions of our sensations, from the testimony of consciousness, with the Cartesians."[295] If I'm right to claim that physicalists must be error theorists, this will mean that all versions of physicalism must posit additions to explain the illusion of distinctness, and their theories will be more complex. Even if this isn't true, since, as Lycan acknowledges, "dualists do not think of either Cartesian egos or immaterial properties as explanatory posits,"[296] it will still be true that dualism is no more complex than physicalism in the ways that matter.

On the view I've been advocating, the relevant starting data set includes, among other things, our clearest intuitions, our introspective sense of things, and the common sense view that

[295] Reid, 1997, 210
[296] Lycan, forthcoming

comes out of these. Taking a drastically different approach, Daniel Dennett endorses what he calls "heterophenomenology." According to Dennett, the data of folk psychology can be taken into account in developing a view in philosophy of mind, but not in a direct way. Dennett thinks that people's testimony about their mental states can be listed among the relevant empirical data to consider, but that it doesn't have any special standing in virtue of our direct access to it in introspection. Like all other data to be considered, this data must first make itself inter-subjectively accessible and therefore appropriately scientific-ish before Dennett will allow it to influence our findings. Dennett claims that it isn't our experiences of ourselves from our first person points of view that count as data, but rather "for heterophenomenologists, *primary* data are the sounds recorded when the subjects' mouths move."[297] So, on Dennett's view, to form a belief in an immaterial mind or a soul, one would have to infer its existence from the inter-subjectively available data of the noises people make, which they take to be about their own minds. This distinction in method creates a vast difference in what one takes to be theoretical additions to the initial data set. Dennett is perhaps the most ardent defendant of focusing exclusively on scientific data. His method is more extreme than most, but evaluating it should be illuminating for any view that wants to significantly privilege scientific data.

I have already explained in chapter 3 part of my reasons for disapproving of the fetishizing of scientific data Dennett tends to epitomize, and I will have more to say about it in the next section. For now, it should be noted that if the above account of what motivates simplicity is correct, then Dennett's approach is unfounded. If the initial data set is fixed by the objects of apparent awareness in the everyday exercise of any of our intellectual powers, then Dennett places an unfounded limitation on what gets to count as data. Indeed, what Dennett's

[297] Dennett, 2005, p. 44

approach amounts to is holding a significant range of our initial data beholden to a theoretical constraint emerging from an intense focus on a proper subset of the data to be explained. Dennett's position seems motivated by the idea that the fact that we've done well when focusing on a subset of the observational data in some contexts shows we should just ignore the rest of it, or rather simply discount it as data at all unless it can first be turned into the sort of data he prefers. This is reminiscent of the Cartesian fetishizing of mathematical or deductive reasoning that led to the dismissal of the authority of direct observation of the physical world. Indeed, this is evident in Dennett's own complaints about the sort of approach I am suggesting. Dennett quotes the following assertion from Nagel:

> If the subjective character of experience is fully comprehensible only from one point of view, then any shift to greater objectivity—that is, less attachment to a specific viewpoint—does not take us nearer to the real nature of the phenomenon: it takes us farther away from it.[298]

Dennett claims that this is question-begging because it assumes we can only fully comprehend subjective experience from the inside. It's unclear why we should think that taking subjective experience to be the best source of data on subjective aspects of experience is question-begging rather than obviously the right thing to do. Apparently, Dennett won't take the obviousness of this for granted, however, insisting that if it were obvious, "then even a rudimentary attempt to deny it should expose itself to decisive refutation."[299] The parallels of history here are once again easy to see. It was the refusal to accept the obviously appropriate method of allowing our ability to perceive the external world to have authority in our beliefs *about* the external world that opened up the idea that we should let deductive reasoning reign in that area. In addition,

[298] Nagel, 1974, p. 447. Quoted in Dennett, 2005, p. 36
[299] Dennett, 2005, p.36

once that wasn't simply granted, it was anything but rudimentary to prove that we should trust our senses. Hundreds of years of failed efforts to establish the legitimacy of perception in early modern philosophy shows how hard it is for philosophers to prove that rudimentary mistakes of method are mistakes of method once they give up on the legitimacy of the means by which they can best and most easily do so. The right thing to do in both cases is simply to refuse to allow philosophers to insist upon an unwarranted rejection of certain types of data in the first place, no matter how much they like other ones.

In addition to shaping how we look at the introspective and intuitive data, this shift in method will also affect how we look at the empirical data. This helps further explain one way that dualism is simpler than physicalism. When we looked at the example of accounting for a self or executive decision-maker, we saw that on one account of simplicity, dualism was better supported by the data. In the absence of the data expected by physicalists, Dennett had to make up a story about how we come to suffer a massive, bizarre illusion about the nature of agency. On my view of the data, we can maintain dualism as the correct view and take the data as the expected, or at least not an unexpected output. In addition, by taking the data of our experience of ourselves as a given aspect of direct experience, rather than taking a third-person perspective on the relevance of the testimony of people about such experiences as a source of data to ground a theoretical conclusion, my view explains why we should avoid views that would run counter to this data. Given the less restrictive data set justified by my background epistemology, and the rejection of Dennett's third-person perspective that goes with it, my view is therefore more justified by this version of the appeal to simplicity. This point about Dennett's views on consciousness generalizes to any view that endorses an eliminativist or error theorist account of some substantial element of common sense. In every case, the defender of the more limited

ontological view will have to tell some sort of story making theoretical claims about cognitive errors or illusions that go beyond what is available to us in our initial establishment of the data to be explained. These claims would be theoretical posits that would then make the corresponding views less parsimonious in the elements of a view that matter than their realist counterparts. Despite a sparser ontology, these views are more complicated than the realist views in areas that deal with elements of common sense ontology. If I was correct to claim that this includes a commitment to active, immaterial minds, views that reject such aspects of common sense ontology are less simple than ones that include them.

Most physicalists seem to have the view that if they can simply provide a story that allows us to think of the mind as a physical thing, then considerations of simplicity will allow that view to win out over its dualist counterparts. However, given this account of the value of simplicity, this doesn't seem correct. I argued in earlier chapters that physicalism of any variety is an error theory about the mind. Given this, whatever story the physicalist tells will involve a revision of our belief set grounded in the adoption of theoretical claims that increase the complexity of their view in the ways that matter. If this is right, then even if the physicalist can tell some revisionist story about what the mind is like, considerations of simplicity not only won't help them in getting over the bar to the favored view, they'll actually count against adopting it as an alternative to the common sense realism of interactive substance dualism. This would mean that the physicalist would have to show that there are serious problems with dualist views to justify the additional push into physicalism, even if they could somehow come up with a plausible physicalist alternative. But the objections we looked at in the last chapter don't seem to give us reason to believe there is any serious problem with dualism as it stands. Therefore, the work ahead for the physicalist is daunting.

It is the physicalist, not the dualist who is engaging in speculative philosophy when they claim that the nature of the mind is fundamentally different than it appears to be. Positing a material nature to our mental states is an unacceptable theoretical addition to the data set, and considerations of parsimony should therefore count against it. Only if one assumes that the data set should start out limited to the empirical data alone could one think that it is the dualist rather than the physicalist who is making unwarranted additions to our understanding of the nature of the mind. However, there have been those who argue for the view that we should limit the data, not merely as a methodological presupposition, but due to the fallibility of such sources of evidence. If there are defeaters for introspection or intuition as legitimate sources of knowledge, then one could justify the restriction of the data set and thereby shift the perspective on which position was violating considerations of parsimony. We will turn to this objection now.

9.2:Dennett on Introspective Illusions and Thought Experiments

The existence of illusions of all sorts have commonly been appealed to in philosophy to undermine the legitimacy of various sources of evidence. Early modern philosophy included numerous appeals to optical illusions to undermine the legitimacy of perception, contemporary meta-ethics contains numerous criticisms of ethical intuitions on the basis of cognitive illusions, while other forms of intuition have faced similar objections, and even introspection has come under fire as subject to troubling errors and illusions. Daniel Dennett has probably been the foremost advocate of the view that introspection is subject to numerous illusions and that this undermines the legitimacy of introspective evidence in deciding issues in philosophy of mind, citing several examples of introspective illusions in his works.[300]

Although illusions are often appealed to in undermining our acceptance of various sources of evidence, this effectiveness of this line of argument is largely undermined by three

[300] For several recent examples, see his 2005, esp. chapters 3 & 4.

facts. First, cognitive illusions are present everywhere to a surprisingly large degree. Therefore, if they undermined the legitimacy of sources of evidence, there would be no legitimate means remaining of forming beliefs, since no mechanism of belief formation is immune from such errors. The prevalence of cognitive illusions in all of our intellectual exercises has been well documented in recent work in cognitive psychology. Studies have shown that you can affect how well people perform on math exams by showing people pictures before the exam to reinforce stereotypes,[301] by having them put different letters at the top of the exam page before starting,[302] and by offering them money for correct answers.[303] I take it that this difference in performance doesn't show that mathematical reasoning is unreliable or that numbers don't actually exist. The style of argument appealed to by Dennett and others would be effective if their target was still a view that held that the mind was fully present and available to us or that we are infallible with respect to our own minds. However, the progress of epistemic theories toward fallibilist views has made the standard for how subject to error a faculty needs to be to legitimize this line of response much higher. In addition to this, illusions stand out primarily because they run contrary to ordinary experience. Our expectation of accuracy when we use introspection is so high that the presence of cognitive illusions is surprising. But this shows that the illusions are rare, and the faculty is generally reliable, not that it can't be trusted. Given any reasonably fallibilist epistemic position, the presence of some surprising errors therefore fails to provide adequate reason to doubt the legitimacy of ordinary judgments made using those means of belief formation. The standard for defeat does not appear to be met by examples of error that seem to exist against a backdrop of what still appear to be reliable beliefs about our own cognition.

[301] See Saul, forthcoming
[302] A discussion of these results can be found at http://www.sciencedaily.com/releases/2010/03/100308203306.htm
[303] A discussion of this result can be found at http://www.freakonomics.com/2012/06/26/bribing-kids-to-try-on-tests/

Finally, the track record of appeals to illusions to undermine the legitimacy of sources of evidence is far from stellar. In the end, the mistake is once again like that of the early modern philosophers who sought proof of the external world after realizing that our means of knowing about it sometimes went awry. Early modern philosophers often appealed to optical illusions as reasons to mistrust perception as a source of knowledge. Again, the idea appeared to be that if it could be shown the faculties used to reach conclusions on the issue in question were less than perfect, this would mean that the preferred methods of the philosophers of the day could and should be used instead. But optical illusions no more show that we typically can't trust our senses than introspective illusions show that we cannot trust introspection. Recognizing the fallibility of human judgment is sufficient to guard against our ability to err. General skepticism toward sources of knowledge goes too far. It also leads to the replacement of reliable sources of evidence with less reliable sources.

The senses are not perfect, and neither is introspection. But by and large they get things right. The exact level of reliability of these means of knowing about the world are debatable. Whatever their level, however, they are far more impressive in their results than what people have sought to replace them with. As we saw in the first chapter of this work, reasoning of any form is far from ideally reliable. Even the smartest, best informed, most practiced thinkers within a discipline get it wrong far more often than not. We often refuse to acknowledge our own errors, but the fact remains that at its very best, reasoning in any field is rarely more than a 50/50 proposition. Despite this, people are so confident in their own abilities in these areas that they somehow think that if they attempt it, they can be more reliable using reasoning at getting the right answers than we can by trusting our basic means of knowing about the world. Asking people to systematically trust in the reasoning of brilliant people over their own eyes is an

inexcusable exercise in hubris, and this is true as much of the inner eye of introspection as the outer eyes of our senses.

In addition to attacking introspection as a source of evidence, Dennett attacks the use of thought experiments, or, as he calls them "intuition pumps," in philosophy of mind. Dennett comes close to describing my position when he decries the fact that

> Some participants in the consciousness debates simply demand, flat out, that their intuitions about phenomenal properties are a nonnegotiable starting point for any science of consciousness. Such a conviction must be considered an interesting symptom, deserving a diagnosis, a datum that any science of consciousness must account for…[304]

Dennett's complaint is an attack on a large number of people, apparently. He admits that "There is a powerful and ubiquitous intuition that computational, mechanistic models of consciousness, of the sort we naturalists favor, must leave something out—something important."[305] The fact that an intuition is ubiquitous and holds such a high degree of conviction makes it quite an interesting datum to explain for sure. Dennett, however, is certain that there can't be anything to this "Zombic hunch," as he calls it. He speculates that "a hundred years from now, I expect this claim will be scarcely credible, but let the record show that in 1999, John Searle, David Chalmers, Colin McGinn, Joseph Levine and many other philosophers don't just feel the tug of the Zombic Hunch (I can feel the tug as well as anybody), they credit it."[306] One can't help but get the sense that Dennett sits there dreaming of the day when the court of high science will drag these fools before the bar and force them to answer for their crimes. But like nearly every idle speculation, this daydream of Dennett's doesn't count as evidence.

[304] Dennett, 2005, p. 178
[305] Dennett, 2005, p. 13
[306] Dennett, 2005, p. 14

Dennett has more to offer than this idle speculation, however. We will turn to that momentarily. First it should be noted that Dennett's view can be seen as part of a broader skepticism about intuitions that has spread lately in philosophy. I find this attack on intuitions baffling. I suppose one main reason for this is because I don't think those attacking them really understand what they are suggesting, or what intuitions are. To show this, I would like you to think about the following claim:

(A) It is possible to think about claims.

Now I would like you to evaluate the following claim:

(B) When thinking about claims, it is sometimes possible to evaluate whether or not they are correct.

Now, I would like you to evaluate whether or not the following claim seems correct:

(C) When thinking about claims, sometimes they seem to be correct.

Now, I'm assuming you were able to think about and evaluate these claims, and that you made judgements about them. When you made these judgments, did you do anything epistemically blameworthy? If you said "no," then you seem to accept the fact that intuitions are sometimes legitimate grounds for forming beliefs.

Intuitionists in various areas are often thought to believe in magic capacities or weird, mysterious abilities.[307] Some have even claimed that intuitions are worthless, or that we have no good evidence for them.[308] This is silly. Intuitions are initial intellectual appearances.[309] To have them requires nothing more than the capacity to think, the capacity to evaluate, and for the things you evaluate to sometimes seem to be true. If that's weird or magical, then so are we. If we are not, then neither are intuitions. Moreover, intuitions are the central source of evidence in

[307] See, e.g. Mackie, 1990, esp. pp38-42
[308] See, e.g. Cummins, 1998
[309] Huemer (2005) p. 99-101 for a more thorough account of this claim.

philosophy. The evaluation of any argument requires an exercise of intuition, often in the form of evaluating premises, and always in the form of evaluating the cogency of the reasoning. If intuitions can't help us to evaluate the legitimacy of positions, the entire field of philosophy ought to just give up. A perfectly general attack on intuitions, then, strikes me as a bizarre and highly counter-productive effort for any philosopher to engage in.

However, Dennett may claim that there are certain classes of intuitions that run against this general story. Indeed, he draws parallels between the intuition of distinctness in philosophy of mind and the vitalist intuition in biology, and to the sense of Americans that there is intrinsic value in dollars, but not in foreign currency. Dennett says that "vitalism—the insistence that there is some big, mysterious extra ingredient in all living things—turns out to have been not a deep insight but a failure of imagination."[310] It's not clear how widespread the vitalist intuition was. Descartes, for example, clearly thought that animals could be explained mechanistically.[311] However, there has been a history of debate where people have been sure scientific explanations would eventually fail. People thought this in explaining the motion of heavenly bodies, the origin of design in nature, and in other cases, such as the vitality of animals. Given the fact that often people have had a sense that there is something there to be explained that science can't account for, and the fact that science has often come along and found a way to account for it, one could conclude that bare intuitions that run counter to scientific progress no longer deserve respect. There is an important difference between these cases, however. In the case of the "intrinsic value" of money of the "vital spirits" of animals, the hunch really is just a hunch. There weren't compelling thought experiments setting aside the relevant features of physiology and identifying clear features of living organism left necessarily left out of the story. There

[310] Dennett, 2005, p. 178

[311] See Discourse 5 for a detailed account of Descartes' knowledge of how animals function, and his insistence that their behavior can be explained mechanistically, while human intelligence cannot be.

seems to be something more there, but what it is isn't specified. In the case of the intuition of distinctness in philosophy of mind, this isn't the case. The function of the thought experiments used in philosophy of mind isn't to express a hunch, it is, as I argued in chapter 5, to abstract away from the various features of the mind we might confuse things like qualia or content with and get us to focus on their intrinsic nature. Dennett claims not to see anything there when you spend enough time stripping away all the other features.[312] I have no privileged access to his mind in these efforts, but I know that I see what is there, that Dennett claims to feel the tug of the intuitions, and that even he admits that the intuitions are ubiquitous, suggesting that after hearing the thought experiments almost everyone thinks there is something there still in need of explanation. Given all this, I can reach no conclusion other than he simply refuses to see what is really there.

In addition to rejecting key aspects of philosophical reasoning, Dennett's position on thought-experiments also threatens to be self-defeating, given his views about science. He asks us not to credit the appeal of thought experiments such as the case of Mary, or of Chalmers' zombies, while asking us to embrace a more scientific approach to understanding the mind. But thought experiments such as the one he wants us to ignore are central to sound scientific inquiry. As L. A. Paul puts it

> To the untutored, such examples can seem like pointless and unnecessarily complicated flights of fancy. But as the use of abstraction and idealization in actual scientific practice reminds us, they play an essential and eminently respectable role in theory-building. Models are used to focus on salient entities or structures (especially in the case of abstraction), to show how to apply untested

[312] See Dennett, 2005, Chapter 8

> theories to possible cases, and most importantly to communicate large amounts of complex information in a simple way.[313]

The use of thought experiments in science makes views that minimize the value of such evidence problematic for advocates of a scientific approach to knowledge. The process of focusing our attention on features by isolating them out or abstracting away from the rest of the details generates intuitions about the nature, especially the modal nature, of those features. The use of these tools to generate intuitions in science are not different in important ways from how they are used in philosophy of mind. Discrediting their legitimacy in the philosophical debates threatens to undermine the practices of science that Dennett so fervently insists on following in debates about consciousness.

9.3: Churchland on Introspection

A somewhat different attack on introspection's evidential weight is suggested in the work of Paul Churchland. Churchland's claim is that the dualist isn't simply asking for equal respect for introspection, but for special authority from introspection in comparison to our other intellectual abilities. As Churchland puts it

> [the argument from introspection] assumes that our faculty of inner observation or introspection reveals things as they really are in their innermost nature. This assumption is suspect because we already know that our other forms of observation do no such thing. The red surface of an apple does not *look* like a matrix of molecules reflecting photons at certain critical wavelengths, but that is what it is. The sound of a flute does not *sound* like a sinusoidal compression wave train in the atmosphere, but that is what it is. The warmth of the summer air does not *feel* like the mean kinetic energy of millions of tiny molecules, but that is

[313] Paul, forthcoming

> what it is. If one's pains and hopes and beliefs do not *introspectively* seem like electrochemical states in a neural network, that may be only because our faculty of introspection, like our other senses, is not sufficiently penetrating to reveal such hidden details. Which is just what one would expect anyway. The argument from introspection is therefore entirely without force, unless we can somehow argue that the faculty of introspection is quite different from all other forms of observation.[314]

If this is correct, then the parallels I have been drawing between our trust of perception and our trust of introspection has been misguided. Fortunately, however, I don't find this reasoning persuasive.

There are two things to note about the cases Churchland mentions. In the case of things like colors of surfaces, sounds in the air, and temperature, although their underlying nature is not fully apparent we (1) do have strong beliefs about the category of thing we are talking about, which sets clear limits on acceptable accounts of that nature, at least in certain cases, and (2) although the appearance represents itself simply in the cases Churchland mentions, there is no sense that the things in question don't have a deeper underlying nature that is being represented qualitatively. With respect to (1), when we think of our experience of color, for example, we are not immediately aware of the fact that our experiences are caused by the surface reflectance properties of objects. However, we also don't think they couldn't be caused by that. Color couldn't turn out to be a mathematical equation just as easily as it could turn out to be a surface reflectance property; it also couldn't turn out to be angel wishes or demonic spirits. There are limits on acceptable accounts of the underlying nature of what we are aware of tied to our conceptual categorization of the thing in question. So, while the entirety of the underlying nature

[314] Churchland, 1988, p. 15

of objects of perceptual experience isn't apparent, enough of that nature is to rule out certain accounts of it. Given this, it isn't a case of special pleading to suggest introspection sets similar categorical limits on acceptable accounts of mental states.

As for (2), in the case of introspection, the thing we are aware of appears to be the entirety of the thing that exists. Qualia and content, for example, appear to have their nature fully present, and don't appear to have an underlying physical base. In fact, since qualia typically serve the role of representing the underlying physical states that cause them, it's not clear how they could do so while also being fundamentally different from what they appear to be. The way a color looks can be fundamentally different than the physical basis of what a color really is, but the way it looks can't be fundamentally different than the way it looks. There's just not enough left there for the misunderstanding to arise. The apparent irreducibility of qualia is what makes it impossible to think that it is hiding a deeper, true nature, whereas the surface-level observation of the physical world practically invites investigation into its underlying nature. Sensations or qualia aren't at all like the things they represent. We have obtained data that the thing that causes the sensation of warmth has a certain underlying nature, but not that the sensation itself has an underlying nature. There's no relevant analog here, because the full nature of qualia appears to be right there in front of us. Given this, one can reasonably claim that in allowing introspection to be a source of understanding of the true nature of things we aren't engaged in special pleading. In both cases, we are asking simply that the aspects of the nature of things we are aware of through the use of our faculties be recognized and accepted. It's just that in the case of introspection, more of that nature is given because there is less to the nature of what is observed to begin with.

9.4 Apparent Materialist Intuitions and what to Say about Them

I argued earlier that we have a strong intuition that it is possible for us to exist without our bodies, and that, in addition, we view ourselves as ontologically simple. These facts, at least at first glance, fly in the face of many other things I know about myself. I know that I am over 6' tall. I know that I have brown hair and brown eyes. I know that I kissed my spouse this morning, and that I took my dog out for a walk. As John Foster puts it, " it is also part of the 'common-sense' outlook that human persons have corporeal properties."[315] None of this fits very well with the idea that I am an immaterial, ontologically simple substance. This seems at first glance to pose a serious problem for the idea that dualism is the common sense view of the self. It poses the difficulty in two ways. First, it is entirely plausible that our sense that we walk around and have a height is more obvious than the intuition that we can exist without our bodies, so if the conflict is genuine, then the materialist intuitions may very well win out. Second, even if it's unclear which intuitions should win the day, it would seem that these intuitions, at first glance anyway, demonstrate a fundamental confusion or incoherence in our sense of the nature of persons, and that this undermines the usefulness of such intuitions in adjudicating debates about personhood. John Foster expresses some concern about the coherence as well, speculating that "perhaps we should conclude, then, that, from the standpoint of our Cartesian philosophy, our ordinary concept of a person is incoherent…"[316]

I believe that this apparent conflict within our understanding of persons between our dualist and materialist intuitions doesn't stand up to scrutiny. I want to sketch here what I take to be the common sense ontology of persons. I don't imagine that the account will be beyond reproach, but merely demonstrating that there is a clear and perfectly coherent metaphysics that

[315] Foster, 1991, p.239
[316] Foster, 1991, p. 239

can account for all of these intuitions demonstrates that the charge of incoherence needn't pose a serious threat.

Here then, is a simple sketch of the ontology of persons I take to underwrite common sense.

(1) Selves are essentially simple mental subjects.

This was discussed in chapter 7.

(2) A body is any physical object.

I think that this expansive use of the term 'body' is an appropriate way to use the term in common speech. It also helps explains why we think that, say, Data from Star Trek, or Johnny 5 from Short Ciruit, could be persons.

(3) A physical object is a spatially extended collection of particles whose behavior is properly unified.

This view is controversial in metaphysics, but I think it is the view that best fits ordinary judgments about objects. According to (3), our sense of object-hood is tied to the behavior of certain portions of matter. Psychologists test for a child's understanding of the nature of objects by conducting tests of object-permenance, and these tests seem to demonstrate that we come to understand how to individuate aspects of physical reality by learning about how the world works.[317] We see objects as those portions of the world that work and move around together, while being discontinuous from the rest of the environment. Whatever the implications in metaphysics, this does suggest that our ordinary sense of 'object' is a functional one.

[317] For a good account of how psychologists test for object permanence, see Bloom, 2004, p. 8-11

(4) A self can have a body, and human persons are selves with a particular sort of body[318].

This explains how we go about questioning the moral status of various things. We wonder if fetuses, or animals, or androids are persons despite the fact that we know they have bodies. This uncertainty seems to rest in the question of whether there is a self or a subject of experience there in addition to the body.

(5) A self has a body just in case the behavior or activity of the self is unified with the behavior of the body.

This account seems to make sense of our intuitions about all sorts of cases. It seems to account for why we can disavow certain actions we may engage in when they are caused by reflex or aren't in our control for some reason or other. It also helps explain how we can think up fascinating stories of possession that attribute the behavior of our bodies to others, and why we can conceive of pretty much anything becoming animate and deserving of moral consideration.

(6) We are essentially selves and contingently embodied.

This explains our disembodiment intuitions.

(7) Our names refer to us sometimes as human persons and sometimes as selves.

This explains why it is true of us that we have a height and weight, and yet could still exist without them. It also explains the apparent contradictions in our discussion about the dead, when people often believe both that someone no longer exists and yet that they are in heaven. The account of the constitution of persons given above seems internally consistent and can account for our various intuitions. It also appears to be widely used in human thought. Given the presence of an available account of persons that naturally accords with how we speak of them

[318] I take no stance on whether or not there can be persons who are disembodied. I am uncertain if our ordinary use of 'person' would include angels or demons or ghosts who don't inhabit a body or not.

and is compatible with the intuitions we have been considering, the threat to dualism from our apparent materialist intuitions can be set aside.

9.5: Concluding Remarks

I think at the end of the day the dispute over method is grounded in a fundamental disagreement about the purpose of philosophy. Traditionally, it was thought that the purpose of philosophy was to add to our understanding of the nature of reality by whatever means we have available to us. In recent history, however, this picture of the purpose of philosophy seems to have faded. Currently most philosophers seem to think, or at least practice philosophy in a way that suggests, that it is the role of science to attain an accurate picture of the nature of reality and that the role of philosophy is merely to help articulate that picture. Contemporary philosophy seems to be guided by the notion that it has nothing unique to add to our understanding of the world, and indeed that at the end of the day nothing but science will have anything to add to our understanding of the nature of the world. Seen in this light, it is not surprising that traditional views of reality not grounded in a scientific approach would become unfashionable. However, despite its currently unfashionable state, I think that a move toward common sense philosophy is not only desirable, but also likely to be successful at drawing followers.

Our goals in philosophy are three-fold. We seek, above all else, the truth. But we do not simply want to know the truth, we want to be able to understand it and communicate it to others. We work in the hope that at the end of the day the truth will be discovered and that it can be elegantly phrased and easily comprehended. This hope may be in vain, but it is a worthwhile goal and we should not abandon hope of obtaining it unless we have to. The question before us, though, is how can this goal best be fulfilled? I propose that we should first look at our own

cognitive capacities and the abilities we have for obtaining knowledge and see what picture of the world these common capacities afford us.

We should start with the hope that our intellectual powers are suited to inform us of the nature of reality. We should seek to understand those capacities and to evaluate the phenomenological character of our mental lives as we exercise those capacities. This should include an examination of the apparent nature of objects of awareness, which should furnish us with a metaphysical starting ground. The picture of the world thus obtained, if it can be reasonably accepted, would seem to satisfy our epistemic goals to the highest extent. A picture of the world afforded to us through exercising ordinary intellectual powers should be comprehensible for that very reason. Since it seems to me that elegance of a theory is largely measured by the intellectual ease with which the theory can be comprehended, if a systematic and coherent picture can be drawn from such an examination it should beautifully satisfy our related goal of elegance. Common sense philosophy therefore offers us our best hope of meeting all of these goals. In addition to this, I think there is good reason to think that if it became widely respected, common sense philosophy could lead to widespread agreement. To see why this is, we should look at one last psychological experiment.

The Milgram experiments are one of the most famous psychological experiments of all time. For those who aren't familiar, in the experiment subjects were told that they were testing the effectiveness of negative reinforcement in the form of electrical shocks on people's ability to accurately answer questions. The subjects were not aware of the fact that the test was actually done to see how far people would go in shocking their fellow citizens if they were told to do so by an authority figure, nor that the person being "shocked" was an actor who was not suffering any pain. In the famous variation of the experiment, people were found to be very likely to

continue shocking someone past the point where the person became non-responsive, and frequently to the point where the person would have died if the experiment were real.

While this is certainly a testament to the dangers of blind acceptance of authority, there are two aspects of the experiment that are not often recounted. First, many people who continued to "shock" the person answering questions claimed that they did it because they were told that it was important for science, and they wanted to do the right thing because they cared about scientific progress. Although the results in this case were quite bad, this is a very good thing, and it should provide assurance to those who seek to pursue the truth that a large portion of the population sees great value in the effort. People genuinely value the truth in many contexts, and are willing to respect scientific or intellectual authority quite readily when they do. This helps explain why ideas often filter down and become embraced by much of the public when they think that they have intellectual credence. People want to learn and want to help researchers; they like to feel like they are part of the intellectual community.

Another interesting feature of these experiments was discovered in a variation involving two authority figures instead of one. In this version, the authority figures disagreed about the need to continue. Both figures stood as equal authorities, debating one another about whether or not to continue with the experiment. In this variation, *no one* was willing to complete the experiment. Faced with two legitimate authorities who disagree, people are not willing to go against what they know is right. This experiment gives me great hope as an advocate of common sense philosophy. Today's intellectual community appears to speak almost univocally on the value of science and on the intellectual deficiency of non-materialist views of the self, of morality, or of the world as a whole. This univocal voice is troubling, first of all, because its views are incorrect. It is even more troubling because, appearing univocal to the common

person, those who value intellectual progress find themselves in a position where they have to ignore much of what they know in order to feel like they are doing their part in supporting the truth. However, if an equally legitimate authority can be set up that allows people to continue to believe what they already knew was right, they will, at that time, stick with what they know and trust their own good sense. Common sense philosophy needs to work on developing an intellectually compelling alternative for serious people who care about the truth to see. I suspect that there are many closet dualists out there, likely working in other areas because they did not think they could find a voice in philosophy of mind. Those dualists need to speak up. Only if we can become an independent authority in the minds of those who seek answers in these areas can we overcome the longing felt by so many to believe what intellectuals are "supposed to believe" by making it no longer true that they are supposed to believe it. Having done so, however, the work will be nearly complete. In the presence of conflicting authorities respected to equal degrees, people will naturally default to what sounds true, and that will be common sense.

I would like to finish by noting that I have aimed big in this work. If I have been fully successful, I will have shown that we should accept a philosophical method according to which common sense creates a bedrock for our work in philosophy that cannot be overturned without a Herculean effort. I will have shown that independently of this method, all genuine mental realists are conceptually wedded to substance dualism. I will have shown that even with only a mild presumption in its favor, the arguments against dualism fall far short of providing legitimate grounds for overturning the view. It is unlikely that I will have been fully successful convincing everyone of each of these strong claims. However, even if I have fallen short of the mark in some areas, there is still a strong case to be made for dualism. If I have shown that there are

good reasons to take the views of common sense as the default views in our philosophical efforts, that the common sense view of the mind fits best with substance dualism, and that the arguments against dualism simply aren't strong enough to overturn this presumption, then I will consider the work a success. If each of these more moderate claims are true, then it is still the case that interactive substance dualism is the most rational position to accept in philosophy of mind. Dualism has been described as "not a serious view to contend with, but rather a cliff over which to push one's opponents."[319] Don't be afraid to jump off that cliff. The fall won't kill you, and once you're here, the grass is green, the sun is shining, and you are free to be person you've always known you really are.

[319] Dennett, 1978, p.252

References:

The 21st Century Monads' "The G. E. Moore Shift" at http://people.umass.edu/phil592w-klement/monads/TheGEMooreShift.mp3

Alston, William, (2005). *Beyond "Justification": Dimensions of Epistemic Evaluation*, Ithaca, NY, Cornell University Press

Anscombe, G. E. M. and P. T. Geach (trans. and eds.), (1954), *Descartes: Philosophical Writings*, Indianapolis: Bobbs–Merrill Company

Armstrong, David (1980), "The Nature of Mind" in *The Nature of Mind and Other Essays* (University of Queensland Press).

Ayer, A. J. (1936). *Language, Truth, and Logic*, London: Gollancz, 2nd Edition, 1946.
-------(1969) "Freedom and Necessity" in *Philosophical Essays* (Bedford, St. Martin's Press)

Bergmann, Michael (2006). *Justification without Awareness*, New York, Oxford University Press

Bloom, Paul (2004), *Descaretes' Baby: How the Science of Child Development Explains what makes us Human*, (Basic Books)

BonJour, Laurence, (2004). "In Search of Direct Realism," *Philosophy and Phenomenological Research*, Vol LXIX, No. 2
-----(2010) "Against Materialism" in *The Waning of Materialism* (Oxford University Press)

Bornmann L, Mutz R, Daniel H-D (2010) A Reliability-Generalization Study of Journal Peer Reviews: A Multilevel Meta-Analysis of Inter-Rater Reliability and Its Determinants. PLoS ONE 5(12): e14331. doi:10.1371/journal.pone.0014331

Boyd, Richard, (1988). "How to Be a Moral Realist," in *Essays on Moral Realism*, G. Sayre-McCord (ed.), pp. 181–228.

Caplan, Bryan. (2007). *The Myth of the Rational Voter: Why Democracies Choose Bad Policies*, Princeton
------ (2011). *Selfish Reasons to Have More Kids*, Basic Books, New York

Carnap, Rudolph. (1932). "The Elimination of Metaphysics Through Logical Analysis of Language," trans. Arthur Pap, in A. J. Ayer (ed.), *Logical Positivism*, New York: The Free Press, 60–81. Originally published in German in *Erkenntnis*, Volume 2.

Casscells, W., Schoenberger, A., and Grayboys, T. (1978): "Interpretation by physicians of clinical laboratory results." *N Engl J Med.* **299**:999-1001.

Chalmers, David (1996), *The Conscious Mind*, (Oxford University Press)

Chisholm, Roderick (1964), *The Lindley Lecture*, (© the Department of Philosophy, University of Kansas).

Churchland, Paul (1984). *Matter and Consciousness*, MIT Press

Cicero. *De Divinatione*. A modern English version translated by Bill Thayer is available at http://penelope.uchicago.edu/Thayer/E/Roman/Texts/Cicero/de_Divinatione/home.html

Conee, Earl, and Sider, Ted. (2005) *Riddles of Existence*, Oxford University Press

Cummins, Robert C., (1998). "Reflection on Reflective Equilibrium," In Michael DePaul & William Ramsey (eds.), *Rethinking Intuition*. Rowman & Littlefield

Dietrich, Eric, (2011). "There Is No Progress in Philosophy" in *Essays in Philosophy,* Vol. 12, Iss. 2, Article 9. Available at http://commons.pacificu.edu/eip/vol12/iss2/9

Dennett, Daniel C (1991). *Consciousness Explained*, Boston: Little, Brown

------(1988), "Quining Qualia," in: Marcel, A. & Bisiach, E. (eds.) *Consciousness in Modern Science*, (Oxford University Press).

-----(2003) *Freedom Evolves* (Penguin Books)

DePoe, John M., (2009). "Defeating the Self-defeat Argument for Phenomenal Conservativism", *Philosophical Studies*, doi: 10.1007/s11098-009-9482-7

Descartes, Rene, (1968). *Discourse on Method and the Meditations*, Penguin Books

Eccles, John C. (1994) *How the Self Controls its Brain*, Berlin: Springer-Verlag

Fodor, Jerry (1989). "Making Mind Matter More." *Philosophical Topics*, 17: 59-79

Fodor, Jerry and LePore, Ernest. (1992). *Holism: A Shoppers Guide*. Blackwell.

Foster, John. (1991). *The Immaterial Self: A Defence of the Cartesian Dualist Conception of the Mind*. Routledge, New York.

Frankfurt, Harry. (1969). "Alternate Possibilities and Moral Responsibility", *Journal of Philosophy*, vol. 66, pp. 829-39

Friedman, Thomas. (2010). "Lies, Damn Lies, and Medical Science" in *The Atlantic,* Nov. 2010. Accessed Feb. 2 2012 at http://www.theatlantic.com/magazine/archive/2010/11/lies-damned-lies-and-medical-science/8269/

Fumerton, Richard, (2010). "You can't Trust a Philosopher" in *Disagreement*, Richard Feldman and Ted A. Warfield, Eds. Oxford.

Goetz, Stewart, (2005). "Substance Dualism" in *In Search of the Soul: Four Views of the Mind-Body Problem,* Joel Green and Stuart Palmer, eds. InterVarsity Press

Heller, Mark. (2000). "Hobartian Voluntarism: Grounding a Deontological Conception of Justification", *Pacific Philosophical Quarterly* 81

Holbach, Paul Henri Thiry, baron d' (1770), from *The System of Nature*, reprinted in *Reason and Responsibility, 12e* (2005), Joel Feinberg, Russ Shafer-Landau, eds. (Thompson Wadsworth)

Huemer, Michael. (1999). "The Problem of Memory Knowledge", in *Pacific Philosophical Quarterly* 80

-----"Naturalism and the Problem of Moral Knowledge," *Southern Journal of Philosophy* 38 (2000): 575-97

-----2001. *Skepticism and the Veil of Perception*, Lanaham, MD. Rowman & Littlefield.

-----2005. *Ethical Intuitionism*. New York, Palgrave MacMillian.

-----2006. ""Phenomenal Conservatism and the Internalist Intuition," *American Philosophical Quarterly*, Volume 43, Number 2

-----2007. "Compassionate Phenomenal Conservatism," *Philosophy and Phenomenological Research*, 74.

Hume, David. *Enquiry Concerning Human Understanding*. Reprinted in *Epistemology, Contemporary Readings*, Michael Huemer, ed. (2002), Routledge, New York, p. 47.

-------(1739-40) *A Treatise of Human Nature.* Reprinted by Oxford University Press, (2000)

Ioannidis JPA (2005). "Why most Published Research Findings are False." *PLoS Medicine* 2(8): e124. **DOI:** 10.1371/journal.pmed.0020124

Jackson, Frank (1982), "Epiphenomenal Qualia", *The Philosophical Quarterly* 32: 127-136.

Kahneman, Daniel (2011) *Thinking, Fast and Slow*, (Farrar, Straus, and Ciroux)

Kim, Jagewon (2005), *Physicalism or Something Near Enough*. (Princeton University Press)

Kornblith, Hilary. (1999). "Distrusting Reason." *Midwest Studies in Philosophy* 23:181-96

Kravitz RL, Franks P, Feldman MD, Gerrity M, Byrne C, et al. (2010) Editorial Peer Reviewers' Recommendations at a General Medical Journal: Are They Reliable and Do Editors Care? PLoS ONE 5(4): e10072. doi:10.1371/journal.pone.0010072

Krugman, Paul. New York Times. July 1, 2010. Retrieved July 3, 2010 from http://www.nytimes.com/2010/07/02/opinion/02krugman.html?_r=2&ref=opinion

Leibniz, G. W. *Discourse on Metaphysics, Correspondence with Arnauld, Monadology*. Quotes from 1997 ed., translation © 1902, Open Court Publishing.

Lewis, David. (1966). "An Argument for the Identity Theory," *Journal of Philosophy*, 63: 17–25.

Libet, Benjamin. (1985). "Unconscious cerebral initiative and the role of conscious will in voluntary action." *Behavioral and Brain Sciences*, 8:529-566.

Lund, David (2005), *The Conscious Self: The Immaterial Center of Subjective States*. (Humanity Books)

Lycan, William (2009). "Giving Dualism its Due," Australasian Journal of Philosophy 87, pp.551-63

Mackie, J. L. (1977). *Ethics: Inventing Right and Wrong*. Penguin Books.

Malinowski, Bronislaw and Ellis, Havelock (1929). *The Sexual Life of Savages in Northwest Melanesia*. Available on Google Books at http://books.google.com/books?id=f0CGuj7r1O8C&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

Markie, Peter. (2006). "Epistemically Appropriate Perceptual Belief," in *Nous*, 40:1

McGinn, Colin (1989), "Can we solve the Mind-Body Problem?", *Mind*, vol. xcviii, no. 391, pp. 349-366.

Merricks, Trenton. (2001). *Objects and Persons*. Oxford.

Mill, John Stuart. *Utilitarianism*. Page citations from *Ethics: History, Theory, and Contemporary Issues, 4e*, Steven M. Cahn and Peter Markie, eds. (2009), Oxford University Press, New York

Moore, G. E. (1903) *Principia Ethica* Cambridge University Press, Cambridge:

---- (1939) "Proof of an External World," *Proceedings of the British Academy* 25. Quotes from excerpt reprinted in *Epistemology, Contemporary Readings*, Michael Huemer, ed. (2002), Routledge, New York, p. 47

Nagel, Thomas (1974) "What is it like to be a bat?", *Philosophical Review*, pp. 435-50 (repr. in *Mortal Questions*).

------ (1979)*Mortal Questions*, (New York: Cambridge University Press); page reference is to a reprinted portion in *Reason and Responsibility, Twelfth Edition, Instructors Edition*, Joel Feinberg & Russ Shaffer-Landau, eds.(2005), (Wadsworth).

Newton, Isaac, *Philosophiae Naturalis Principia Mathematica* ("*Mathematical Principles of Natural Philosophy*"), London, 1687; Cambridge, 1713; London, 1726.

Ney, Alyssa (2008). "Physicalism as an Attitude," *Philosophical Studies*, 138, 1-18

Nichols, Shaun, and Knobe, Joshua (2007) "Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions*", Nous* 41:4, 663–685

O'Connor, Timothy (2000) *Persons and Causes: The Metaphysics of Free Will* (Oxford University Press)

Papineau, David (2002) *Thinking About Consciousness.* (Oxford University Press)

Paul, L. A. (2012). "Metaphysics as Modeling: The Handmaiden's Tale" *Philosophical Studies*, 160, 1-29

Plantinga, Alvin (2007) "Materialism and Christian Belief" in *Persons: Human and Divine*, Peter van Inwagen, ed. (Oxford University Press)

----- (2011)

Plato. *Phaedo*. Quotations from translation by Benjamin Jowett available at http://classics.mit.edu/Plato/phaedo.html

Popper, Karl and Eccles, John C. (1977). *The Self and its Brain: An Argument for Interactionism,* Berlin, Springer-Verlag

Reid, Thomas (1788) *Essays on the Intellectual Powers of Man*. Page number taken from 2002 edition, Derek R. Brookes, ed. (Penn State University Press)

----- (1813) *Essays on the Active Powers of the Human Mind.* Page citations from 1969 edition (Cambridge, MA: M.I.T Press)

-----(1764) *An Inquiry into the Human Mind on the Principles of Common Sense*. Page citations from 1997 edition, Derek R. Brookes, ed. (Penn State University Press)

Ryan, Sharon. (2003). "Doxastic Compatibilism and the Ethics of Belief", *Philosophical Studies* 114

Ryle, Gilbert. (1949). *The Concept of Mind*, London: Hutchinson. Page references are to the 2000 republication, London: Penguin Books.

Saul, Jennifer, "Unconscious Influences and Women in Philosophy" forthcoming in *Women in Philosophy: What needs to Change?*, Fiona Jenkins and Katrina Hutchinson, Eds.

Searle, John (1984) *Minds, Brains, and Science* (Harvard University Press)

Soon, et. al (2008). "Unconscious Determinants of Free Decisions in the Human Brain," Nature Neuroscience 11, 543 - 545 (2008)

Steup, Matthias. (2004). "Internalist Reliabilism", in *Philosophical Issues*, 14, *Epistemology*

Stove, David, (1991). "What is Wrong with Our Thoughts? A Neo-Positivist Credo" in *The Plato Cult and Other Philosophical Follies*, Basil Blackwell, Inc.

Sturgeon, Nicholas, (1985). "Moral Explanations," in *Morality, Reason, and Truth*, David Copp and David Zimmerman (eds.), Totowa, NJ: Rowman and Allanheld, pp. 49–78.

Suess, Dr. (1961) *The Sneetches and Other Stories*. Random House, New York

Thomasson, Amie (2006) "Debates about the Ontology of Art: What are we doing Here?", *Philosophy Compass,* 1: 245-55

Van Inwagen, Peter (1983) *An Essay on Free Will* (Oxford University Press)

Wolfe, Thomas, (2002). *Hooking Up*. Farrar, Straus, and Groux, New York

Yudkowski, Eliezer. "An Intuitive Explanation of Bayes' Theorem." Retrieved from http://yudkowsky.net/rational/bayes/ Retrieved most recently on March 6th, 2013

Curriculum Vitae

Matthew Skene
Syracuse University

Contact Information

8330 E. Quincy Ave., I-307
Denver, CO 80237

(303) 731-9779
mskene@syr.edu

Education

B.A.	University of Colorado, Boulder	2002 (magma cum laude)
Ph.D.	Syracuse University	Ph.D. defense August 20, 2013

Dissertation: *Putting the Ghost Back in the Machine: A Defense of Common Sense Dualism*

Supervisor: Robert Van Gulick

Abstract: Gilbert Ryle once ridiculed substance dualism, describing it as the view that we are a "ghost in the machine." Since that time, substance dualism has found few defenders, and a presumption toward naturalism has dominated philosophical inquiry. Here, I offer an unapologetic defense this unfashionable view of the self. To do so, I first explain why philosophy should endorse a shift in method away from naturalism and toward common sense philosophy. I then show how, from within that approach, substance dualism is far better supported than its competitors. My defense of the common sense method rests heavily on an account of justification and of evidence defended by Michael Huemer, and results in the view that we should give strong presumptive weight to common sense. I proceed to argue that a commitment to active, immaterial minds is built into the conception of human nature used in everyday life, and therefore qualifies as a tenet of common sense. This fact gives a strong presumption to dualism. I finish by considering objections to dualism, and showing that they not only fail to overcome this strong presumption, but would fail to overcome even a moderate conservatism with regard to beliefs about the nature of the mind.

Areas of Specialization

Epistemology, Philosophy of Mind

Areas of Competence

Meta-Ethics, Thomas Reid, Introductory Logic, Ethics

Publications

"Seemings and the Possibility of Epistemic Justification" (2013) *Philosophical Studies,* vol. 163, no. 2, 539-559, DOI 10.1007/s11098-011-9830-2

Works in Progress/Under Review

"The Causal Argument Against Physicalism"- I argue for the following claims: (1) that agent causation is an important aspect of our pre-theoretical understanding of mental causation, (2) that agent causation is incompatible with the commitments of physicalism, and (3) that this poses a far more serious problem for physicalists than has been traditionally realized.

"Philosophical Error and the Economics of Belief Formation"- I look at some recent work in medical meta-research, in applied economics, and in cognitive psychology and show how they provide good reason to believe that academic philosophy is likely subject to systematic biases. I explain why these biases are most likely to be found in presuppositions of philosophical method, and argue that this problem requires us to re-examine our philosophical methods in order to counter-act this influence.

"Rich, Varied, and Parsimonious"- I argue for the view that best reasons we have for endorsing theoretical and ontological simplicity cuts against the most common uses of these considerations in philosophy. Specifically, I claim that the most persuasive basis for simplicity rests in an anti-speculative philosophical approach that privileges foundational beliefs over theoretical ones. I then argue that most standard uses of the principle, especially those advocating error theoretical or eliminativist positions, run contrary to this foundation by requesting that we privilege theoretical and often speculative views about cognitive error over more direct and reliable exercises of our intellectual capacities.

"Minimally Decent Lives and Social Justice"- I provide a clear account of the notion of a "minimally decent life" and explain its role in political reasoning. The account I offer grows out of commitments of common sense morality, and defines a minimally decent life as a life we are not morally permitted to engage in rights violations in order to avoid. I then use this notion to arbitrate a dispute between advocates of social justice and libertarian minimal-state theorists. Specifically, I argue that this conception provides a clear reason for libertarians to care about social welfare as an influential political concern, but does not justify most of the traditional forms of government intervention advocated by social justice proponents.

Teaching Experience

Syracuse University

Teaching Assistant

Introduction to Logic	Fall 2003, Spring 2004, 2005
Human Nature	Fall 2004

Sole Instructor

Theories of Knowledge and Reality (Epistemology, Existence of God, Free Will, Phil. Mind)
(8 terms, 13 total classes) Fall 2005-Summer 2008

Community College of Aurora

Adjunct Instructor

Introduction to Philosophy	Spring 2010-Fall 2013
Ethics	Fall 2010, 2013; Spring 2011, 2013

Platt College

Adjunct Instructor

Critical Thinking	Spring 2010- Fall 2013
Philosophy and Society	Winter 2012

University of Colorado at Denver:

Adjunct Instructor for the CU Succeeds Program (Class taught at D'Evelyn High School)

Introduction to Philosophy	Fall 2011

Academic Experience

Conference Organizer, 2007 Graduate Conference in Philosophy, Syracuse University.

Presentations

1. "Allowing for Unjustified Deliberative Beliefs", commentary on Elizabeth Palmer's "Deliberative Belief Formation as an Action", April 8, 2005, Graduate Conference in Philosophy, Syracuse University.
2. "Questioning the need for Usefulness and for Non-Circular Justification", commentary on Deke Gould's "Rule Circular Justification and Deduction", ABD Series Presentation, March 30, 2007, Syracuse University.
3. "Seemings and the Possibility of Epistemic Justification", ABD Series Presentation, September 4, 2007, Syracuse University.
4. "The Causal Argument Against Physicalism", ABD Series Presentation, April 2, 2009, Syracuse University.

References

Professor Robert Van Gulick
Department of Philosophy
Syracuse University
Syracuse, NY 13244-1170
(315)-443-5828
rnvangul@syr.edu

Professor Andre Gallois
Department of Philosophy
Syracuse University
Syracuse, NY 13244-1170
(315)-443-5825
agallois@syr.edu

Professor Kris McDaniel
Department of Philosophy

Professor Michael Huemer
Department of Philosophy

Syracuse University
Syracuse, NY 13244-1170
(315)-443-5818
krmcdani@syr.edu
Prof. Mark Heller (teaching reference)
Department of Philosophy
Syracuse University
Syracuse, NY 13244-1170
(315)-448-5822
heller@syr.edu

University of Colorado
Boulder, CO 80309-0232
303-492-7276
owl232@earthlink.net
David Spiegel (teaching reference)
Community College of Aurora
16000 E. CentreTech Parkway
Aurora, CO 80011
(303)-340-7562