Northwestern University

A Puzzle about Desire: The Metaphysics and Epistemology of Wanting

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Philosophy

By

Jared Thomas Peterson

Evanston, Illinois

September 2017

Abstract:

In this dissertation I pose a new puzzle about desire, one grounded in three plausible but jointly inconsistent propositions. According to the standard view in metaphysics, (1) all desires are dispositional states. Epistemologists, though, think that (2) we have privileged access to some of our desires. But (3) it is very difficult to see how we could have privileged access to any dispositional state. In this work, I explore this disconnect between the metaphysics and epistemology of desire. I argue that we need a new metaphysical account of desire—one that captures its phenomenal character. I then argue that it is on the basis of the unique phenomenology of desires—what I call *attraction*—that we come to possess epistemically secure, uniquely first-personal knowledge of our desires. Such privileged access to our desires helps explain how we can posses the type of rational, autonomous agency that is crucial to our understanding of ourselves as persons.

Acknowledgments: Special thanks to my dissertation committee Baron Reed, Sanford Goldberg, Jennifer Lackey, and Brie Gertler, all of whom provided valuable and extensive comments on this work. Thanks as well to the following philosophers for their helpful feedback on this work: Richard Fumerton, Lauren Ashwell, Ruth Chang, Andrew Melnyk, Eric Schwitzgebel, Fabrizio Cariani, Kyla Ebels-Duggan, Michael Glanzberg, Mark Alznauer, Alex Byrne, Michael Smith, Eli Alshanetsky, Matt Duncan, Derek Green, Ezra Cook, Nicholas Leonard, Rebecca Mason, and audience members at the 2016 Pacific APA, the 2016 Society for Exact Philosophy Conference, the 2017 Central APA, and the 2017 Pacific APA conference for their insightful comments concerning issues related to this project.

Table of Contents:

Chapter 1: A Puzzle about Desire

It seems obvious that we know at least some of our desires. I know that I want a drink of water when I'm thirsty, but I also know that I want my friends to be healthy, that I want more ethical treatment of factory farm animals, and that I want Chicago's 'L' to be cleaner and more efficient. How I know that I have these desires is not so obvious. Despite the burgeoning work on self-knowledge—i.e. work on how we know our mental states—there has been little developed discussion concerning the self-knowledge of desires. When philosophers have put forth detailed accounts of the self-knowledge of this attitude, they have been almost entirely silent on what desires are. Jordi Fernandez (2007), Krista Lawlor (2008), Alex Byrne (2012) and Lauren Ashwell (2013) in the only recent full-length treatments of the epistemology of desire offer little by way of a discussion of what these states are.[1] From the other direction, those who have worked on the nature of desire have not been concerned enough, I think, with what the epistemic implications of their work are for their favorite view of desire. Timothy Schroeder (2004), for example, in his recent work on the metaphysical nature of desire is silent on how we know what we want given what he thinks desires are. I think such narrow-minded focus in both the epistemology and metaphysics of desire is problematic. More

[1] Of the philosophers mentioned here, Fernandez (2007) and Ashwell (2013) offer the most developed explanation of what desires are.

precisely, I think this narrow-minded focus has resulted in philosophers missing an important, previously un-discussed puzzle about desire, a puzzle that demonstrates the impact that one's views concerning the metaphysics of desire can have on one's views concerning the self-knowledge of desire, and vice versa.[2] The puzzle I have in mind is as follows:

(1) All desires are dispositional mental states.

(2) We have privileged access to some of our desires

(3) We do not have privileged access to any dispositional mental state.

(1) is taken to be true by a number of philosophers of mind and ethicists. (2) is thought to be true by a number of epistemologists working on the self-knowledge of desire. And (3), as Brie Gertler (2011a) contends in her recent work, is widely embraced by philosophers as well. But (1)-(3) are jointly inconsistent. We have, here, a trilemma. Solving this trilemma is a matter of figuring out which of the three independently plausible propositions above is false.[3]

---

[2] I am unaware of anyone in the literature who has explicitly developed and discussed this puzzle. The closest analogy to it is Paul Boghossian's (1989) puzzle concerning the compatibility of privileged access and content externalism. Briefly put, Boghossian's concern is that insofar as the content of thought is individuated externally, this poses a threat to our privileged access to these states. He argues that this is the case since we cannot have privileged access to relational properties. The puzzle I am interested in is not couched in terms of content but rather in terms of dispositional accounts of desire, analyses that can be coherently coupled with internalist accounts of content.

[3] I should note here that there are "irrelevant ways" one can get out of the puzzle. An irrelevant way to get out of the puzzle would be to establish that we have privileged access to a dispositional state that is not a dispositional desire. This does not resolve the worry though about how we have privileged access to desires since the puzzle can easily be reformulated by replacing (3) with the following proposition: "We do not have privileged access to desires *if desires are dispositional states*." I mean to exclude these irrelevant ways as viable solutions to our puzzle. That being said, as will become clear, I do think we lack privileged access to any dispositional *attitude*. I explore the implications of this claim in Chapter 7.

In this work I argue that the solution to the above puzzle is that (1) is false. That (1) is false, though, does not indicate whether we have privileged access to our desires—that is, whether (2) is true—and if (2) is true, how we achieve this type of access. I take these latter questions to be the larger puzzle about desire. And in the latter part of this work I attempt to solve this larger puzzle as well. I argue that (2) is true by offering a detailed account that explains how we possess privileged access to our desires. I further argue that my solution to our puzzle and my epistemological views concerning desire provide the template for explaining how we know other attitudes as well.

In the remainder of this chapter I offer a more thorough explanation of the puzzle, including an explanation of key concepts it involves. I also explain why a number of philosophers have taken (1)-(3) to be true.

1. Motivating the Puzzle

Recall that the first proposition of our puzzle is that all desires are dispositional states. In his recent work on the nature of desire, Schroeder (2004) notes that a popular way of analyzing desires—what he refers to as a version of the Standard View—runs as follows:

ST: To desire that P is to be disposed to act in ways to bring it about that P.[4]

D.M. Armstrong (1980), Robert Stalnaker (1984), and Michael Smith (1994) embrace this dispositional, action-based account of desire, or something very similar to it. Ashwell

[4] 'ST' stands for "Standard View." Schroeder takes accounts such as this one to be a version of the standard view that desires motivate us to act.

(2014) has more recently acknowledged sympathies with ST as well. As implied above, there are other more complicated formulations of ST, but the majority of these permutations construe all desires as dispositional states as well.

Schroeder goes on to note that the main rival to the standard view is a hedonic theory of desire that sees these states as essentially connected to pleasure. Schroeder implies that the most plausible version of such a view construes all desires as dispositional states. A crude version of such a view would be the following:

HT: To desire that P is to be disposed to tend to feel pleasure if it seems that P, and or displeasure if it seems that -P.

Two of the more popular ways of understanding desires, then, are accounts that view all desires as dispositional states. Let's call the thesis that all desires are dispositional states, *strict dispositionalism about desires* (or *SDD*) and the proponents of such a view *strict dispositionalists*. According to these philosophers, occurrent—i.e., non-dispositional desires—are a myth. Such a view, as Schroeder's work suggests, is widespread.

What, however, does it mean to say that a particular desire is a dispositional state? I follow others—including Eric Schwitzgebel (2002) and Gertler (2011a)—in thinking of dispositional states as states that, at the very least, *entail* counterfactual statements of the following form: if condition C holds, then an agent S will Φ, where Φ can be understood as a mental event, a behavior or an action. Call these mental events, behaviors, or actions *manifestations* of the disposition. Call the event of C occurring the stimulus conditions.

At the very least dispositional properties entail these types of complex counterfactuals. But I am going to remain neutral, for the moment, on whether these properties are nothing over and above the satisfaction of the content of counterfactuals.[5]

On the standard view of desire, then, S has the disposition in question only if given certain conditions C, S acts in ways to bring about the content of that desire. What the conditions in question are will depend on the content of the desire in question. But generally speaking, at least some of the types of conditions that must be in place for the manifestations to occur are conditions that will involve tokening certain means-ends beliefs and other desires. An example will help to explain this. Suppose I desire a glass of Syrah. The stimulus conditions that must be in place in order for me to act in Syrah-seeking ways might be tokening the belief that I can readily obtain this type of wine, or instantiating the belief that drinking Syrah will not kill me, etc., while the desires that might need to be in place include desires to have a pleasurable experience, etc. Needless to say, spelling out the stimulus conditions for even our simplest desires is going to be no easy matter. But at least the broad conditions that must be in place for a dispositional desire to manifest itself should be clear enough.

One might wonder, at this point, why (1) of our puzzle is so widely embraced. The short answer to this question—one that I will return to in chapter 4—is that a number of philosophers of mind are sympathetic with the view that desires are to be understood

[5] Gilbert Ryle (1940) and other philosophers with an empiricist bent have held that, e.g. the solubility of an object O is nothing over and above the fact that were O to be placed in liquid, then O would dissolve. Other so-called *realists* about dispositional properties have held that the solubility of object O is a matter of O possessing a categorical base (e.g. a certain crystalline structure) that makes it the case that were O to be placed in water, O would dissolve. Most parties in this debate, however, agree that dispositional properties *entail* complex counterfactuals—i.e. if X has dispositional property Y, then a certain counterfactual involving stimulus conditions and manifestations is true of X *because* X has Y.

in terms of what these mental states do. And insofar as certain conditions need to be met in order for desires to do what they do, it is natural to adopt the view that (1) is true. Consider: if one thought the role desires play in our mental economy is to motivate action, but one also maintained that desires only motivate action when certain conditions are in place, then one is well on their way towards embracing a strictly dispositionalist account of desires. It is not surprising, then, that (1) is thought by some to be a truism.[6]

The second proposition of our puzzle is the claim that we have privileged access to at least some of our desires. What does it mean to say that one has privileged access to some fact F?[7] I follow others in embracing the following account of privileged access:

PA: An agent S has privileged access to some fact F iff. (a) S knows F in a highly epistemically secure manner and (b) S arrives at this highly epistemically secure knowledge via a uniquely first-personal means.[8]

---

[6] There are also a number of philosophers of mind who think that insofar as a state is a fleeting, short-lived event, that state is not a desire. This is because the latter, according to these philosophers, must be a more stable attitude than that. Admittedly, it is difficult to find this line-of-thought explicitly defended, but I believe that it has motivated a number of philosophers to embrace the view that all desires are dispositional states. I return to this point in chapter 4.

[7] A number of philosophers think we have privileged access to only facts. See Neta (2011) for this claim. If you think we have privileged access to non-facts (e.g. states of affairs, for instance), then you are free to replace 'F' with a variable indicating what type of object you think we have privileged access.

[8] Byrne (2005, 2012) calls (a) privileged access and (b) peculiar access. He, however, notes that a number of philosophers are comfortable construing privileged access as the conjunction of both (a) and (b). I am one of them. Alston (1971) provides a nice overview of various ways of understanding privileged access. I should also mention here that I am sympathetic with a similar, but more complex account of privileged access. I defend that account in my paper, "The Value of Privileged Access." However, because: (i) that account is controversial, (ii) this more complex account of privileged access is very similar to PA, and (iii) because most philosophers, I think, embrace the above account of privileged access, I will for the purposes of this project understand privileged access in the less-controversial way.

Concerning (a), while some philosophers have required epistemic certainty or perfect reliability in order to possess privileged access, I do not think privileged access demands that level of epistemic security. Instead I follow other contemporary epistemologists in understanding the epistemic security needed to have privileged access to be that approaching certainty with respect to internalist accounts of knowledge and that approaching perfect reliability with respect to externalist accounts of knowledge.[9]

On the account of privileged access I favor, it is not enough to merely possess epistemically secure knowledge. One must arrive at this epistemically secure knowledge via a uniquely first-personal means. To say that S knows O in a uniquely first-personal way is to say that S knows this fact by employing a means that cannot, in principle, be used by other agents in order to arrive at this same knowledge.[10] So, for example, I can find out that I desire a certain person's affection by observing my behavior (e.g. I stutter when I'm around her; I go out of my way to get this person's attention, etc.), or I can learn that I have this desire via testimony from a psychologist. But these are ways you can come to learn that I have this desire as well. The manner in which I know that I instantiate this desire must be a manner that you cannot, in principle, also employ to

[9] Admittedly, there is some vagueness with respect to the details of (a). How close to certainty or perfect reliability must an agent's knowledge be in order for her to have privileged access? If fallibilism about knowledge is true, then I want to contend that knowledge of one's own mental states does not suffice for privileged access. See Gertler (2011a), Neta (2011) and Byrne (2005, 2012) for the claim that one can have privileged access without having certain or perfectly reliable knowledge of the fact in question. As I will contend in chapter 3, though, I think some of these philosophers understand the epistemic security required to have privileged access too weakly.

[10] That one's knowledge of one's own mind be arrived at in a uniquely first-personal way can also be understood in terms of directness or immediacy. The thought is that to have privileged access we must come to know our minds in a non-inferentially direct manner. I have refrained from characterizing (b) in this way because some view perception as a highly reliable, non-inferential belief-forming process and I take it we do not want to say that both possessing perceptual knowledge and self-knowledge are a matter of possessing privileged access. There are also some extrospective accounts of self-knowledge, to be discussed in chapter 3, that entail we do have self-knowledge to our minds via a method that is inferential.

arrive at knowledge of this fact about me. PA, I contend, is in keeping with how others in the literature have understood this philosophical term of art, and this reason counts in favor of understanding privileged access in the manner that I have here.

Given PA, it seems at least *prima facie* plausible that we have privileged access to at least some of our desires. The following example will help motivate this:

> **Nevada Desert:**
>
> Carson is driving through a relatively untraveled part of Nevada in the middle of a heat wave when his car breaks down. Unfortunately no one else is around to give Carson a ride and he must walk to get help. As his walk drags on, Carson runs out of water. He is struck with an almost unbearable thirst. He judges that he wants water.

In the above example it seems reasonable to suppose that Carson's knowledge that he wants water is highly epistemically secure. It also seems plausible to suppose that Carson arrives at this highly epistemically secure state via a uniquely-first personal means. Carson, after all, does not have to observe his own water-seeking behavior or wait for someone to come along and tell him he wants water in order to know that he does. This case, and others like it, render it *prima facie* plausible that we possess privileged access to some of our desires.[11]

[11] I stress here that I am only arguing for the *prima facie* plausibility of this claim here. Certainly more needs to be said in defense of (2) before we can embrace the view that this part of the puzzle is true. This is all the more the case since, as will be further discussed in Chapter 5, there are a growing number of philosophers who think that we either lack privileged access to any mental state, or, less radically, that the scope of such access is much more limited than has previously been thought. Carruthers (2005; 2011), Medina (2006), and Schwitzgebel (2008) are three philosophers who argue for this thesis. There are also a number of philosophers who are sympathetic with the view that we have privileged access to some

It also seems reasonable to think, however, that (3) of our puzzle is true; i.e. it seems reasonable to think we do not have privileged access to any dispositional state. Gertler (2011a) has recently defended the view that this is the case. She points out that in order to know that we possess a dispositional state we would have to know that a certain complex counterfactual is true of us. In order to know, for example, that we are disposed to Φ, where Φ-ing is some type of action, we would need to know that we would Φ given certain conditions obtaining. But, according to Gertler, it is reasonable to think that how we would act in a certain situation does not seem to be something that we know via uniquely first-personal means.[12] To see why, consider that determining how I would act in certain circumstances might very well require me to make an inductive inference from behavioral evidence concerning how I have acted in the past to how I would act in similar circumstances. But that would be a process that could be used by third-personal parties to know your mind.[13] Furthermore, and Gertler does not stress this point, it seems that while such a process might afford me with knowledge, it will not afford me with the type of robust knowledge required for privileged access.[14]

---

sensations as well as to some of our occurrent thoughts, but that we lack this privileged access to attitudes such as beliefs and desires. See Pitt (2011).

[12] Gertler (2011a) puts the point this way:

> What we would do or feel under certain circumstances is shaped by a variety of factors. It is plausible that at least some of the factors that would shape our behaviors and feelings in situations we have not yet encountered, are not accessible by an exclusively first-personal method like introspection (73).

[13] One might think that one could know they are disposed to act in certain ways on the basis of "internal cues" or "promptings," which would make the process uniquely first-personal enough. Lawlor (2008) and Cassam (2014) claim that we know many of our desires in this way, and the latter is explicit that the desires he has in mind are standing states. Gertler has also suggested to me in dialogue that one might embrace this position. I address this claim in chapter 3.

[14] As noted in f.n. 10, a number of philosophers have held that privileged access requires having immediate or direct (i.e. non-inferential knowledge). This is in part because the directness afforded by being, e.g.

Gertler, then, maintains that dispositional attitudes are outside the realm of privileged access. She further contends that the majority of philosophers involved in the current scope-of-privileged access debate endorse this claim.[15] If (3) is as uncontroversial as Gertler claims it is, then we are faced with a problem. We have, it seems, three independently plausible propositions all of which cannot be true. The puzzle, then, involves figuring out which proposition (or propositions) in the trilemma is false.

Faced with the above puzzle, I imagine that a number of strict dispositionalists will be inclined to deny (3); in other words, they will be sympathetic with the view that we do have privileged access to some dispositional mental states.[16] Let's call the view that we have privileged access to dispositional mental states, and specifically to dispositional desires, *access dispositionalism* (AD) and those who hold this thesis to be true, *access dispositionalists*. Access dispositionalists are likely to contend, I think, that Gertler has not adequately defended (3). And it should be acknowledged that the latter does not address whether particular theories of self-knowledge are more promising than others at explaining how we have privileged access to dispositional states. Furthermore, Gertler also does not engage in a discussion of what the correct metaphysical account of attitudes such as desires are, something the Access Dispositionalist might, rightly I think, contend needs to be discussed in order to adequately defend (3). So we need, it seems, a lengthier defense of (3) than Gertler provides us with before we are on firm ground in embracing this part of the puzzle. I think such a defense can be given and thus, that

---

directly acquainted with a mental state, is thought to provide epistemically secure knowledge. Bertrand Russell (1912) was an early proponent of this view.

[15] See Gertler (2011a; p. 75-76) for this contention.

[16] Michael Smith, in correspondence, has revealed to me that this is the part of the puzzle that he would reject.

denying this part of the puzzle is untenable. It is the objective of the next two chapters to demonstrate that this is in fact the case.

Before proceeding further, though, it will help to have in place an overview of the chapters to follow.

2. An Overview of the Project:

In chapter 2, I begin my defense of the claim that access dispositionalism is false—i.e., that (3) of our puzzle is true. Building off the work of Gertler (2011a), I discuss introspectional accounts of self-knowledge—i.e. accounts that entail that we know our minds via a process of *looking inward*—and I argue that vis-á-vis any plausible conception of the nature of dispositional desires, such views fail to account for privileged access to these types of states. The two broad introspectional theories of self-knowledge I discuss are the acquaintance approach to self-knowledge and the inner sense theory approach. In the case of the acquaintance approach, I argue that even if some version of this view is true, we cannot be directly acquainted with dispositional mental states. The consequence is that if acquaintance is the only means by which we have privileged access to dispositional mental states, then we do not have privileged access to them.

In the case of the inner sense theory, I argue that one can grant the truth of this account of self-knowledge, and still reasonably maintain that we do not have privileged access to dispositional desires via inner sense. I proceed to offer reasons for thinking that such theories are false, and *a fortiori*, these theories cannot adequately explain privileged

access to dispositional states. My defense of the claim that inner sense theories are false relies on a view about the nature of our epistemic relationship to our minds.

In chapter 3 I consider whether three alternative and somewhat widely accepted views of self-knowledge of our attitudes—transparency accounts, rational agency accounts, and constitutivist accounts of self-knowledge—legitimate a rejection of (3). I argue that they do not. As in the case of inner sense theories, I contend that one can grant the truth of these theories and still demonstrate that they fail to explain why (3) is false. I proceed to argue, though, that the most plausible versions of these theories of self-knowledge face difficult, and in some cases, insurmountable problems. In my discussion in this chapter and in chapter 2 of the prospects of explaining privileged access to dispositional desires, I also criticize a number of metaphysical accounts of desire. The work I do in these chapters, then, provides some ground-clearing criticism of particular theories of self-knowledge and particular theories of desire. This latter work helps pave the way for what I take to be the correct metaphysical account of desire, and the correct account of the self-knowledge of this mental state.

In chapter 4, I offer my solution to our puzzle. I do so by arguing that (1)—the claim that all desires are dispositional states—is false. The argument I offer in defense of the view that this thesis is false—what I call the *Phenomenological Argument*—involves the contention that some desires possess phenomenology, specifically the phenomenology of attraction. I defend this contention by arguing that (1R) desires are the realizers of the causal role of desires, (2R) some realizers of the causal role of desires possess the phenomenology of attraction, and therefore (3R) some desires possess

attraction. Both (1R) and (2R) are controversial claims, and the bulk of the work I do in this chapter involves a defense of these two premises. I argue that we have good reason to think both premises are true and that the above line-of-reasoning is sound. If desires possess the phenomenology of attraction, this spells disaster for strict dispositionalism about desires. This is because no dispositional state, as I argue, possesses phenomenology. In other words, the truth of (3R) entails the falsity of SDD. This chapter's work, then, attempts to put to rest (1) by offering a metaphysical account of desire that views some desires as phenomenal, and hence occurrent, states.

The above solution to our puzzle, however, does not by itself explain how we have privileged access to our desires. In chapter 5 I argue that we do in fact have privileged access to some of our desires—i.e. that (2) of our puzzle is true. The epistemology of desire I offer relies on the previously defended the claim that desires possess phenomenology. More precisely, I argue that we in part come to possess privileged access to our desires via process I call *phenomenal simulation*—a process that involves entertaining a given content and attending to the fact that we are attracted to it. There are a number of criticisms skeptics can raise to such an account. I address some of these criticisms in chapter 6.

In the seventh and final chapter of this work, I argue that my explanation for how we possess privileged access to our desires can be extended to other attitude-types. I do so by defending the view that the state of being curious possesses a unique phenomenology, and that we can know what that we are curious about the particular object of our curiosity via phenomenal simulation. Our privileged access to other

attitudes such as hopes and intentions, I suggest, can be explained in a similar manner. The claims I make about desire, then, provide fertile ground for further research on the self-knowledge of attitudes.

Chapter 2: Access Dispositionalism & Introspectional Accounts of Self-Knowledge

In the previous chapter, I noted that a number of strict dispositionalists about desire are going to maintain that the way to solve our puzzle is to reject (3). In other words, these philosophers will maintain that we do have privileged access to some of our dispositional desires. I called such philosophers *access dispositionalists* and the thesis that we do have

privileged access to some dispositional desires, *access dispositionalism* (AD).[17] As previously mentioned, I think access dispositionalism deserves a more careful investigation than skeptics of this view have given it. In this chapter and the next I engage in this investigation by considering whether the most viable theories of self-knowledge—when coupled with the most plausible theories of desire—can explain our privileged access to these dispositional states. I consider two such theories of self-knowledge in this chapter: acquaintance theories and inner sense theories of self-knowledge.

Both proponents of acquaintance and inner sense hold that self-knowledge is a matter of *looking inward* to know the contents of our minds. So both accounts are what we might call *introspectional accounts of self-knowledge*. I argue that one can grant the truth of these introspectional accounts of self-knowledge and yet still maintain that they fail to explain privileged access to our dispositional states. I also defend the view that regardless of the truth of access dispositionalism, inner sense accounts of self-knowledge are problematic in part because they cannot do justice to the epistemic relationship we bear to our minds.

The discussion of introspectional accounts of self-knowledge will help make clear why a number of philosophers concerned with how we know our own attitudes, specifically dispositional attitudes, will think that non-introspectional accounts of self-

[17] One might, of course, maintain that we have privileged access to dispositional desires as well as to other dispositional attitudes. I imagine that anyone who thinks we have privileged access to dispositional desires will also think we can have privileged access to dispositional attitudes that are not desires. We might refer to this stronger thesis as *attitude access dispositionalism*, while referring to the view I have been discussing as *desire access dispositionalism*. My target in this chapter and the next will be *desire access dispositionalism* and for ease of reference I will simply refer to this thesis as *access dispositionalism*.

knowledge stand a better chance of vindicating AD. Non-introspectional accounts of self-knowledge will be taken up in the next chapter.

It will be helpful to begin addressing the plausibility of access dispositionalism by looking at a historically influential approach to self-knowledge that has little chance of helping us make sense of how we possess privileged access to dispositional states. Doing so will help explain why access dispositionalism would have had very few proponents in the past, and why this thesis seems obviously false to a number of contemporary philosophers. Such a discussion will also point up the challenge that faces proponents of alternative theories of self-knowledge who want to defend AD. The introspectional approach to self-knowledge I have in mind is the acquaintance approach to self-knowledge. In what follows I will explain what such an approach is (typically) thought to amount to and work to demonstrate why acquaintance theories are not going to be able to explain our privileged access to dispositional states.

1. Acquaintance Theories of Self-Knowledge & Access Dispositionalism:

A number of philosophers have thought that self-knowledge is special in a way that other types of knowledge are not. Self-knowledge is thought to be unique in that we can be directly acquainted with its objects.[18] To say that we are directly acquainted with an object is to say there is nothing that stands between the knowing agent and the object known. An example will help clarify this. Russell (1912) thought we could possess knowledge of physical objects. He did not, however, think that our knowledge of these

[18] "Object" here is being used neutrally to refer to whatever it is that we are acquainted with. It should not be read as referring to a "concrete particular."

objects is metaphysically direct. According to him, awareness of medium-sized bread box objects is mediated by our visual experiences of these objects.[19] Our visual experiences, on the other hand, are not mediated by anything. Russell holds, then, that we stand in a more direct relationship to our experiences than we do to external objects, and we stand in this relationship via the *sui generis*, non-causal relationship of acquaintance. In other words, acquaintance is an unanalyzable relationship we bear to some of the contents of our minds.

Russell further contends that standing in this metaphysically direct relationship enables one to possess non-inferential knowledge of what one is acquainted with. This is because when one stands in this relationship of acquaintance to, e.g. some experience E, one does not need to infer that one tokens E from any other facts to know this is so. We can contrast this type of epistemic directness with other types of knowledge, e.g. the knowledge that someone else is having a certain experience, which does need to be inferred from other facts—viz. facts about this agent's behavior or mental life—in order to be known.

How do we determine whether we stand in a relationship of direct acquaintance to an object? Russell suggests we determine this by employing Descartes's Method of Doubt. On perhaps the most plausible way of understanding Russell's suggestion, if one can rationally doubt that the object in question exists, then one is not acquainted with the object in question. To return to our previous example, one might, it seems, be unable to rationally doubt that one is having a certain visual experience. It certainly seems

[19] Russell thought of experiences as sense data, but we need not follow him on this point in order to understand the distinction he makes.

reasonable to think, however, that rational doubt is possible concerning whether a given physical object exists. One, after all, might be hallucinating, or the victim of a mischievous neuroscientist, etc. If this is the case, then according to Russell, we should think that we are directly acquainted with experiences but not with physical objects.

Contemporary acquaintance theorists follow the spirit of Russell's view in holding that we bear a metaphysically direct, non-causal relationship of acquaintance to some facts about our minds, and that this relationship is crucial to explaining how we have privileged access to such facts.[20] I am not going to fully explore the details of such accounts in this chapter, although I will return to discuss elements of this approach in chapter 5.[21] What is important for our purposes here is whether the acquaintance approach is going to be promising ways of vindicating AD. The sketch of acquaintance theories I have provided here, I think, suffices to demonstrate why such theories are not promising ways to do so.

Consider again Russell's test for determining whether we are acquainted with an object. Russell asks us to consider whether it is possible to rationally doubt the existence of the object in question. Given that this is the test we should apply for determining whether we are acquainted with an object, then it seems reasonable to think we are not (and probably cannot be)[22] acquainted with our dispositional mental states. This is because we can always, it seems, rationally doubt whether we instantiate such properties.

---

[20] Contemporary acquaintance theorists who subscribe to this thesis, or something very similar to it, include: Fumerton (1995; passim), Fales (1996), Gertler (2001, 2011a, and 2012), Bonjour (2003), Chalmers (2003), Horgan and Kriegel (2008), Horgan (2011), and Balog (2012).

[21] The account of the epistemology of desires I will be similar in spirit to Bonjour's (2003) account.

[22] If, that is, the modal force of this claim is weaker than logical possibility. This is because I do think it's logically possible for us to be acquainted with dispositional states.

That is, we can rationally doubt that we would Φ given that certain conditions obtained. If this is the case, these are not the types of properties that are going to withstand the method of doubt. It is much more reasonable to hold that one can be acquainted with at least some of the manifestations of one's underlying dispositional states. But these manifestations are just that; they are not the dispositional states themselves. This is why, I think, Russell points to states such as sensory experiences—paradigmatic *occurrent* mental states—as the types of states we are acquainted with.[23]

It is not surprising to find, then, that no current acquaintance theorists defend the view that we are acquainted with dispositional properties. All of these theorists implicitly (and in some cases explicitly) endorse the view that are acquainted only with occurrent mental sates.[24] Fumerton (1995), for example, defends the view that we are acquainted with a fact F only when F is, so to speak, *before the mind*. In other words, Fumerton thinks that the facts we are acquainted with are occurrent or non-dispositional in nature. Gertler (2001; 2011a), Chalmers (2003) and others argue that we are acquainted with a first-order state when one's second-order judgment about that first-order state is partly constituted by the first-order state. This type of constitution occurs, according to these philosophers, only in cases in which the first-order state is occurrent.[25] And Bonjour (2003), an acquaintance theorist who is sympathetic with the view that we are acquainted with some of our attitudes, contends that the attitudes we are acquainted with are

[23] As mentioned above, Russell is going to speak of being acquainted with sense data and not visual experiences. This fact does not problematize the point I am making, though.
[24] Or facts about occurrent or conscious mental states.
[25] Some of these theorists, e.g. Gertler, only defend an acquaintance theory of self-knowledge with respect to sensations. This is perhaps because some of these theorists are skeptical that there are occurrent attitudes. This fact alone provides further support for the view that these theorists think we can only be acquainted with occurrent states.

conscious, occurrent states. All of these acquaintance theorists, then, appear to think we are not directly acquainted with dispositional states

AD, I think, does not stand much of a chance if the acquaintance approach to self-knowledge correctly explains our privileged access to our minds.[26] Those who are largely sympathetic with such accounts will, I imagine, find it obvious that we lack such access. There are, however, alternative approaches to self-knowledge that stand a better chance of vindicating AD, including an alternative introspectional approach to self-knowledge—viz. the inner sense approach. I address this approach next.

2a. Inner Sense & Access Dispositionalism—Armstrong & Lycan

According to the inner sense theory of self-knowledge defended by D.M. Armstrong (1968) and William Lycan (1996) the way we achieve self-knowledge is via an observational causal process similar to the type of visual process a number of philosophers think we undergo to arrive at knowledge of the external world. On one standard understanding of the view, this causal process involves becoming aware of a mental state M via attention mechanisms in the brain "scanning" M, an awareness that in turn causes one to introspectively judge that she is in M.[27] To take the inner sense theorist's favorite example, an agent comes to form the belief that she is in pain by first becoming aware of this phenomenal state via a scanning process in the brain.

---

[26] It does not stand much of a chance, that is, if the acquaintance approach is the only approach that can explain such access. Also, I should emphasize here that it doesn't stand much of a chance regardless of what dispositional account of desire one embraces. As long as the property in question is dispositional, it seems, we cannot be acquainted with it.

[27] Some inner sense theorists might hold that the very scanning of a mental state M causes one's belief about M as opposed to first causing awareness of M.

Inner sense theorists typically couple this view of introspection with a version of epistemic externalism in an attempt to explain our self-knowledge.[28] An inner sense theorist, for example, might embrace a crude version of process reliabilism—e.g. the view that a belief is justified iff. it is produced by a reliable belief-forming process—and claim that since the process of introspection is a reliable process an agent is afforded justified beliefs about her mind. Insofar as the inner sense theorist embraces a thoroughgoing externalism, an agent, on this view, need not be aware of the fact (if it is a fact) that the process is reliable, nor does that fact need to be internal to the agent in order for her to have a justified introspective belief[29] of the introspected state. On this way of understanding self-knowledge, then, it is not the relationship of acquaintance—a non-causal relationship—that explains our knowledge of our minds; it is rather the reliability of a causal process involving attention mechanisms or monitoring devices in the brain. And if the knowledge afforded by this process is highly reliable, it might be thought that inner sense affords us with privileged access to our minds.

One might wonder, however, whether the alleged knowledge afforded us via inner sense is uniquely-first personal knowledge. Some philosophers appear sympathetic with

[28] Armstrong (1968) embraces an externalist account of knowledge. The distinction between internalist and externalist accounts of knowledge and justification is a difficult matter to sort out. Two common ways of spelling out this distinction involve a focus on what can be a justifier (with internalists typically holding that only mental states can, and externalists denying this), and/or a focus on accessibility of justifiers (with internalists typically holding that what justifies one's belief is in some sense accessible, and externalists denying this). Fumerton (1995) offers what I take to be the most careful discussion of how to understand the epistemic internalism/externalism divide. He arrives at the conclusion that perhaps the fundamental dispute between internalists and externalists is whether fundamental epistemic concepts can be "naturalized" (i.e. analyzed in terms of "scientifically-respectable" properties).

[29] And perhaps knowledge if the reliability in question is high enough and the belief is true.

the view that it is not.[30] And in their defense, if self-knowledge involves attention mechanisms that causally scan states, it seems that it is in principle possible to "hook up" an agent's scanner to another agent's brain and in doing so enable the former to scan the latter's mental states. Both Armstrong and Lycan embrace the possibility of "rewiring cases."[31] A number of philosophers take this to be a problematic consequence for the inner sense theorist.[32] They might well be right. The important question here, though, is whether rewiring cases cast doubt on the claim that knowledge via inner sense is a uniquely first-personal process.

I want to suggest that the answer is 'no.' Recall again clause (2) of PA. That clause states that in order to possess privileged access, S needs to possess highly epistemically secure knowledge of the fact in question by employing a means that cannot, *in principle*, be used by other agents in order to arrive at this same knowledge. I suggest we understand the *in principle clause* in terms of close possible worlds. In order for one to employ a uniquely first-personal means, it must be the case that in close possible worlds to this one, others are not using that means to know the contents of your mind. As applied to Armstrong and Lycan's view, in order for inner sense to be uniquely first-personal it must be the case that in close possible worlds to this one, no one else is employing an

[30] I think Moran (2001) would contend that any "knowledge" afforded by inner sense is not uniquely-first personal knowledge, but it is not clear whether he is understanding uniquely first-personal in the way I have understood it above. Gertler contends that inner sense theorists might be able to account for uniquely first-personal knowledge in some "attenuated" sense (2011; p. 68).

[31] Armstrong (1968) notes that, "It is perfectly conceivable that we should have direct awareness of the mental states of others" (124), and Lycan (1996) contends that "Neither you nor I could have... functionally direct access to someone else's mental states (*except by some futuristically special rewiring*) (49) [emphasis added].

[32] Gertler (2000) is one such philosopher. She uses the possibility of rewiring cases to argue that the inner sense theory cannot explain the uniquely direct relationship we have to our minds, and should be dismissed because of it.

internal scanner to know your mind.[33] Since rewiring cases are not taking place in close possible worlds to this one, one can, it seems, have uniquely first-personal knowledge via inner sense. Of course, this makes privileged access a contingent affair on the inner sense theorist's view. But I think this is a consequence they would be happy to embrace. If it actually turned out that in this world persons were, with some regularity, scanning each other's brains, then we would lack privileged access via this method. But such a world is (thankfully) not the world we at present inhabit. There does not, then, seem to be anything barring the inner sense theorist from explaining the privileged access we allegedly possess.

We still need to investigate, however, whether such theories, if true, explain our privileged access to dispositional desires, and whether such theories are in fact true. Let's begin with the first question. In addressing it we might first consider whether we can directly perceive dispositional properties via the type of causal process Armstrong and Lycan claim it is most akin to—namely, vision. In the case of vision, it seems, our perceptual systems are not able to directly perceive dispositional properties.[34] We cannot, for example, directly perceive the fragility of a vase, or the elasticity of a cord. We can, one might more reasonably suggest, directly perceive the manifestations of a disposition (e.g. the shattering of the vase, or the stretching of the cord), and from this awareness, we could infer that the object in question has the dispositional property in question. Similarly, it seems that in the type of introspective process envisioned by Armstrong and Lycan, what is directly detected is not a dispositional property, but rather the

[33] Or, more carefully, no one else is employing an inner scanner in the way that you employ an inner scanner to know your own mind.
[34] By "directly detect," I mean "detect without inference."

manifestation of that property. Goldman (2006)—another philosopher sympathetic with the view that self-knowledge involves an introspectional causal process—agrees. He claims that it is hard to see how a dispositional property would be the type of property that "could be directly causally registered in the introspectional process" (248). He suggests in turn that awareness of such dispositional properties would have to be inferred. If Goldman is right about this, and I think he is, what is the consequence of this fact with respect to the tenability of access dispositionalism?

The consequence, I think, is that it is going to be difficult for a proponent of inner sense to explain how we have privileged access to dispositional desires. To see why, first assume that desires are what most philosophers of mind take them to be—i.e. states that dispose us to act in various ways. If the way we know we instantiate such desires[35] is via awareness of the manifestations of these underlying dispositions, then it would seem that what would afford the most reliable knowledge that we instantiate one of these dispositions would be awareness of the manifestations that in part make desires the desires they are. But those manifestations are behaviors.[36] And if we are aware of our desires via inferences from our awareness of how we behave, this process is not going to be a uniquely first-personal process. The process in question would be a third-personal process that others could use to know our minds.

An access dispositionalist might respond by first noting that dispositional desires have multiple manifestations. She might further contend that although what we might call the "essential manifestation" of dispositional desires is observable to others, not all of the

[35] Unless otherwise qualified, when I refer to "desire" in this section, I will be speaking of dispositional desires understood along action-based lines.
[36] I am understanding *actions* here as a subset of *behaviors*.

non-essential manifestations will be accessible to third-parties.[37] She could in turn argue that if one infers on the basis of one or a number of these non-essential manifestations that they have the dispositional desire in question, then one could come to know they instantiate these desires in a way that does not involve awareness of one's behavior. Both Krista Lawlor (2008) and Quassim Cassam (2014) defend the view that we arrive at knowledge of our desires in this way.[38] Lawlor, for instance, contends that we often times do make inferences from what she calls "internal promptings"—i.e. inner, evidential indicators—as a way of coming to know what we want. Lawlor aptly refers to this method of knowing our desires as "casual self-interpretation" (60).[39]Now it seems plausible that in some cases causal self-interpretation, as Lawlor and Cassam suggest, affords us with knowledge of our standing desires. But what is important for us to consider is whether such a process generates uniquely first-personal, highly reliable knowledge of these states. And I think there are good reasons to be skeptical that such a process does afford us with privileged access to our dispositional desires.

First it is not clear causal self-interpretation is a uniquely first-personal means. Lawlor's favorite example of this process in action will help us see why. She has us consider an agent, Katherine, who wonders whether she wants to have another child.

[37] Ostensibly, these non-accessible manifestations will be manifestations of mental phenomena.

[38] Lawlor (2008) focuses most of her discussion on desires. And while she never tells us whether she thinks all desires are dispositional states, the types of desires she focuses on are best understood as dispositional desires. Cassam (2014), on the other hand, contends that it is not just our dispositional desires that we know in this way, but our attitudes, more generally, as well. Cassam appears to be convinced that all attitudes are dispositional states, and he in turn thinks that an inferential process like the one described above is the typical method we use to know what we believe, want, intend, etc. Peter Carruthers (2010) adopts a similar view.

[39] It is reasonable to think that the evidential indicators Lawlor speaks of are, at least in some cases, caused by the underlying disposition. As noted above, Lawlor is silent on issues concerning the metaphysics of desire. It is not surprising, then, that she is silent concerning whether evidential indicators, in most cases, will be what I am calling non-essential manifestations of the standing state. Cassam does not have much to say about this as well.

According to Lawlor, Katherine can arrive at knowledge of this desire by taking note of certain internal promptings. Here is how Lawlor describes the process Katherine undertakes:

> Katherine starts noticing her experiences and thoughts. She catches herself imagining, remembering, and feeling a range of things. Putting away her son's now-too-small clothes, she finds herself lingering over the memory of how a newborn feels in one's arms. She notes an emotion that could be envy when an acquaintance reveals her pregnancy. Such experiences may be enough to prompt Katherine to make a self attribution that sticks. Saying, 'I want another child,' she may feel a sense of ease or settledness (57).

Katherine's self-interpretational process might not end there.[40] But assuming it does, this way of knowing one's desires looks to be a far cry from what we typically think of as a uniquely first-personal way of knowing our minds. After all, it seems like what is doing the bulk of the epistemic work here is the non-first personal process of abduction. To see this more clearly, consider that Katherine could reveal to her partner her recent thoughts, experiences, emotions, etc. that led her to self-attribute the desire for a new child. Katherine's partner could then make the same type of inference Katherine makes from her evidential base. And if all goes right, epistemically speaking, this person would possess the same type of secure knowledge that Katherine possesses. True, Katherine's partner

[40] Lawlor notes that Katherine might attempt to acquire even more evidence before judging that she has the desire. A less thorough agent, of course, might infer that they have the desire in question on the basis of awareness of one internal prompting. More on this below.

does not have uniquely first-personal access to Katherine's evidential base. But one might reasonably think this only serves to indicate that what Katherine has uniquely first-personal access to is not her desire, but the internal promptings that caused her to self-attribute that desire. We have, then, reason to doubt that the causal self-interpretation process is a uniquely first-personal process.[41]

Even if we grant that the causal self-interpretation process, can, in some cases, yield uniquely first-personal knowledge, we might still have doubts that this method of self-knowledge will generate the type of highly reliable knowledge needed to possess privileged access. This is because it is not hard to imagine such reasoning leading one into error. One obvious way this can happen is that we simply fail to correctly interpret our internal promptings. To return to Katherine, consider a circumstance in which Katherine becomes aware of a feeling of envy that comes over her when a friend tells her about her recent pregnancy. On the basis of this awareness Katherine infers that she

[41] One who wants to defend the claim that the abductive process Katherine engages in is uniquely first-personal enough because it involves an inference from an evidential base learned about in a uniquely first-personal manner, owes us an explanation, I think, of why we should not take there to be two different processes at work here—one involving how we come to know our evidential base (a process that might very well be uniquely first-personal), and one involving the reasoning we engage in from that evidential base. I have argued that we should understand Katherine's case as involving two different processes, only one of which is uniquely first-personal, and this account appears to make sense of why most of us, I imagine, would balk at calling the above process uniquely first-personal.

I should also mention that one who wants to claim that insofar as we come to know a piece of evidence in a uniquely first-personal way, this renders the process by which we infer a conclusion from that evidence a uniquely first-personal process, is going to incur some very awkward consequences. Consider an agent who engages in a very lengthy deliberation process involving multiple pieces of evidence, only one of which is learned about via a uniquely first-personal means. This person comes to a conclusion on the basis of such a deliberation, and insofar as she possesses knowledge from such a process, she could be thought to have uniquely first-personal knowledge. But why should the mere fact that this one piece of evidence is learned about in a uniquely first-personal manner, render the entire process uniquely first-personal? Wouldn't it be more reasonable to think that the deliberative process in question is third-personal, and that the one piece of evidence is learned about in a uniquely first-personal way?

wants another child. But this interpretation is mistaken. What Katherine really wants is the experience of being able to tell others she is pregnant.[42]

Now, it is, I grant, more reasonable to think that an inference will be reliable when it is made from a robust evidential base. But it seems obvious that we are not always (and perhaps not typically) in the position of possessing a substantial amount of reliable indicators of our standing desires. This points to another reason why we might think that causal self-interpretation is not going to be a highly reliable process: it might very well be the case that in ordinary circumstances we make inferences from a very paltry number of internal promptings, internal promptings that do not actually support the self-attribution in question. Lawlor discusses cases in which agents self-ascribe a desire on the basis of very little evidence because they are the type of agents who want to have an issue settled. And it is plausible to think such inferences might very well lead to mistaken self-ascriptions. That being said, even if one isn't *that* type of agent, given both our cognitive limitations and the limited amount of time we have to engage in the self-interpretation process, it seems that these types of inferences typically would involve a very limited evidential base, the type of evidential base we might think does not provide us with highly reliable beliefs.

An access dispositionalist with inner sense sympathies might object at this point that focusing solely on cases like Katherine and her desire for a new child is problematic. They might point out that desires like Katherine's are not, in general, the type of desires we should think we can make highly reliable inferences about. We should instead focus

[42] The fact that one wants to be able to tell others she is pregnant might, of course, be consistent with one also wanting a child. In fact, the former might, in certain cases be good evidence of the latter. But one can certainly want to tell others she is pregnant, without actually wanting to be pregnant.

on desires that we are more likely to have privileged access to via inner sense—e.g. desires such as those associated with our bodily sensations like a desire to eat. For example, if one finds themselves aware, via inner sense, of a hunger pang and the thought that "food sounds good at the moment," one could in turn infer that they have the standing desire for food. And this inference, a proponent of this move could suggest, is highly reliable.

I agree that it is more reasonable to think we can, in general, have highly epistemically secure knowledge of these types of desires than the types of desires Lawlor focuses on. But there are a couple points worth noting about this response. First, insofar as the process involved in coming to know that one wants to eat is a matter of causal self-interpretation then the point above about such causal interpretation not being a first-personal process still applies. If this is the case, then regardless of the reliability of the these types of inferences, one still lacks privileged access to the desires one self-attributes on the basis of this reasoning.

Second, this objection raises a more general problem concerning these theories of self-knowledge that has been lurking in the background of our discussion. The problem is a matter of figuring out which process type the process token of coming to know one wants food in the example above falls under. We might wonder whether the process token falls under the process type *causal self-interpretation*, *causal self-interpretation of desires associated with bodily sensations*, *causal self-interpretation of desires associated with bodily sensations of Katherine experienced on June 2nd, 2016*, etc.? We need a principled way to answer this question in order to effectively determine the epistemic

status of Katherine's belief. The issue I am raising here is not a new one. It is an example of a problem that has come to be known as the Generality Problem—i.e. the problem of how to come up with a principled explanation of what process type a given process token falls under. It is a problem that challenges any reliabilist theory of knowledge or justification. And until the access dispositionalist with reliabilist sympathies can provide such an explanation, they should, I think, be wary of claiming we have highly reliable beliefs about our minds via causal-self-interpretation.[43] There are, then, several serious problems facing an inner sense theorist who wants to defend AD, at least if desires are states that dispose us to act.

It seems just as reasonable to think that the same problems outlined above are going to confront access dispositionalists regardless of what strictly dispositional account of desire they adopt. For insofar as we cannot directly detect dispositional states, it seems that an inferential process is going to be involved in coming to know these states, an inferential process that does not appear to be uniquely first-personal. Furthermore, there are going to be similar questions concerning the reliability of such an interpretational process given the types of errors that can and do occur in making such inferences.

For example, if one accepts the view that a desire that P is one that disposes us to belief that P is valuable, then while we might think that one has privileged access to some of their judgments, one would still need to infer on the basis of these judgments that one is disposed to believe that P. And a number of philosophers have cogently argued that it

[43] Or via any means, for that matter. Of course, such a philosopher would not only have to provide us with a solution to the Generality Problem, but would also need to marshal empirical evidence to the conclusion that the means in question leads to highly reliable beliefs. See Conee and Feldman (2004) for a clear explanation of the Generality Problem.

can be the case that one occurrently judge that P, while failing to dispositionally believe that P. I will have more to say about such cases in the next chapter. There I will argue that we have good reason to think such cases are not anomalies. The point here, however, is simply that, insofar as dispositional properties cannot be directly detected, one can be misled about whether they instantiate such properties. Of course, the Generality Problem makes it, at the very least, difficult to determine what the reliability of a given process token will be, but as noted above, this fact is not going to make life easier for the access dispositionalist.

One way of resisting the position I have been developing is to argue that we actually do, in some cases, directly detect dispositional properties. Goldman contends that this is in fact not the case. But one could argue that this is only because we have not paid close enough attention in our discussion to the question of what dispositional properties are. Such a philosopher might in turn argue, as some have, that dispositional properties are identical to categorical bases. So, for example, the dispositional property of solubility is identical to the categorical base of solubility—e.g. the crystalline structure of a sugar cube.[44] If one held such a view, one might in turn hold that we directly detect such properties, and more specifically, directly detect these categorical bases via inner sense.

I think, though, that we can grant the contentious claim that dispositional properties are identical to their categorical bases, and still deny that it is reasonable to

[44] This identity thesis is different than the claim that dispositional properties simply don't exist, or that dispositional expressions don't refer to genuine dispositional properties, but are merely place holders to be filled in by whatever science tells us the causal bases of such properties are. There is a parallel here between reductive physicalist theories and eliminativists in philosophy of mind. The former hold that while mental state types are identical with brain state types, the term "pain" refers to a genuine property. Eliminativists, on the other hand deny this. The metaphysicians I am discussing with respect to dispositions adopt a position similar to reductive physicalists.

think we directly detect such properties via inner sense. In order to see why, it will help to return to the case of visual perception. We do not (and perhaps cannot) directly detect the categorical base of a sugar cube. Even if one is a direct realist, one will think that what we detect are the macro-properties of a sugar cube. We certainly don't have the visual capability of detecting the micro-structure of a sugar cube. And it would seem like the same can be said about our detection of our psychological categorical bases. The properties in these cases, according to most contemporary philosophers, would be complex properties of the brain. But surely we do not (and more contentiously cannot) detect such complex properties of the brain via inner sense. If we could, neuroscience would be much easier than it is. Now Goldman (2006) thinks we do detect neural properties via inner sense. But he would not, I think, claim that we directly detect complex categorical bases of our minds. And for good reason. It would have to be a powerful[45] faculty that afforded us this type of access. So I do not take the above response to be a promising one.

I have been arguing that even if we grant the truth of Armstrong and Lycan's inner sense account of self-knowledge, there are strong reasons to doubt that this theory can effectively support AD. But we might also ask whether Armstrong and Lycan's view of self-knowledge is tenable. There have been a battery of arguments in the literature to the conclusion that it is not.[46] A number of these arguments attempt to draw out a

[45] Some would say "magical."

[46] Perhaps the most famous objection to the inner sense theory is Shoemaker's (1998) Argument from Self-Blindness. In brief, Shoemaker argues that: (i) It's not possible for rational agents to be self-blind (i.e. it's not possible for rational agents with the appropriate conceptual resources to fail to possess self-knowledge), (ii) if the inner sense theory is correct, it is possible for rational agents to be self-blind, (iii) therefore the inner sense theory is flawed.

problematic consequence of the theory. I won't discuss such arguments in detail, but I will register that the most widely discussed of these arguments, I think, are either (a) flawed or (b) dialectically problematic.[47]

An alternative, albeit more arduous approach, is to attack the epistemology that undergirds such theories. There has, for instance, been much written about the alleged fact that the type of brute reliability Armstrong and Lycan rely on cannot do justice to the epistemic relationship we bear to our minds. Those offering such objections motivate the claim that reliability is not sufficient for our possessing knowledge (and/or justified belief) of our mental states. So, for example, we find Peacocke (1998) and Zimmerman (2006) contending that on Armstrong and Lycan's view we are like a reliable clairvoyant with respect to the contents of our minds. On the latter's view, we find ourselves believing propositions about what we believe, desire, hope, or experience, and if all goes right at the sub-personal level, we have justified beliefs (and in some cases highly epistemically secure knowledge) about our states. But the fact that such processes are reliable at the sub-personal level does not suffice for knowledge. We need an explanation of our relationship to our minds that explains how we possess such knowledge at the reasons-giving level.

I am sympathetic with the above claim, although I also recognize that reliabilists have ways they can respond to it. But what I want to argue next is that the type of reliability that underlies Armstrong and Lycan's view is not *necessary* for possessing justified beliefs about the contents of our minds, a claim that at least in the context of self-knowledge, has not yet to be offered. In order to arrive at this conclusion, I want to

[47] I think (a) is true of Shoemaker's Argument from Self-Blindness.

focus on an awkward consequence of embracing the inner sense approach to self-knowledge, a consequence referred to by one of the advocates of the inner sense view, as "misfiring." Here, is Lycan's (1996)gloss on misfiring:

> [A] potentially disturbing implication of the inner sense theory is that in addition to misreporting the character of a first order state, an internal monitor could in principle fire without anything like a proper cause, [thereby] giving a false positive. Thus the inner sense view predicts that it is possible for a person to be unveridically conscious or aware of a sensation that simply does not exist. You might introspect a sharp, severe pain when there is in fact no pain at all (19).[48]

Lycan goes on to note that this is a theoretical possibility he is "happy to live with" (19). A number of philosophers including Karen Neander (1998) and the Lycan of two years later[49] are not entirely happy to live with this result. But I do not want to focus on misfiring as a way to challenge the mechanics behind the inner sense theory. Instead I want to offer a *gedanken*-experiment involving a misfiring scanner that challenges the epistemology that undergirds it. The case in question is as follows:

> **Misfiring Scanner:**
>
> Sophie is the sworn enemy of a mischievous, but skillful neuroscientist named Alvin. One night Alvin sneaks into Sophie's house, anesthetizes Sophie, and alters her internal scanner so that it systemically misfires

[48] Lycan calls this misfiring a "mendacious representation" of a first order pain.
[49] See his (1998).

> generating false positives—misleading representations of pain-like states. Alvin also surgically engineers things so that whenever Sophie's internal scanner misfires, Sophie exhibits pain-like behavior. Next morning Sophie wakes up not knowing what has taken place in the night. Her scanner now with some regularity misfires, and she consistently forms judgments that she is in pain when she is not.[50]

I take it that Sophie's scanner is an unreliable belief forming mechanism with respect to introspective judgments about pain due to Alvin's rewiring operation. However, I also think that when Sophie judges that she is in pain, such judgments are justified. In other words, I think that Sophie has false, but highly justified introspective beliefs. If this is correct, the upshot is that reliability is not necessary for justification.[51]

There are responses a reliabilist can make to the above objection.[52] One might, for instance, argue that in the above case Sophie's scanner is still, in general, a reliable-belief forming mechanism (i.e. introspection for Sophie is in general a reliable belief-forming process), and if that is the case the process in question does lead to justified beliefs in the case of such mendacious representations. Now if one had a solution to the generality problem, discussed earlier, that allowed for this move to be principled and non-*ad hoc*,

---

[50] I think we can allow for the possibility that when Sophie is actually in pain, she still believes that she is in pain.

[51] It is not misleading to think of the the above thought experiment as a self-knowledge version of Stewart Cohen's (1984) "new demon" problem. Cohen's thought experiment involves a scenario in which an agent S is in a Demon/BIV scenario and consequently has perceptual belief-forming mechanisms that are systematically unreliable. S, however, is unaware of this unreliability. Crude reliabilism has the result that such beliefs are unjustified. But many are inclined to say that the BIV victim still has justified perceptual beliefs. If one is inclined to think this, then the upshot is that reliability is not necessary for justification.

[52] Lycan is an advocate of the Higher Order Theory of Perception, the view that to be conscious or aware is just for the subject to be aware of the state through inner sense. Sophie, in the above case, is conscious of the false positive only in the sense that she is made aware of this false positive through inner sense. But Lycan also suggests that one can be phenomenally conscious of the false positive as well.

then I submit, we would indeed have to alter the thought experiment. Sophie's scanner would have to be systematically misfiring with great regularity producing a number of mendacious representations. I think such a scenario poses more problems than the above thought experiment, which is why I refrained from formulating the example in this way. But it seems to me that such problems are not formidable enough so that the same result can be achieved.

There are certainly other responses an inner sense theorist can make to the above objection, and if such objections fail, such a philosopher is certainly free to retreat to a more sophisticated version of reliabilism.[53] But it is reasonable to wonder, given both the objections to both the claim that reliabilism is necessary for justification and to the sufficiency claim, whether even more sophisticated reliabilist and externalist accounts of knowledge and justification are going to adequately capture the epistemic relationship we bear to our minds. I am inclined to think they fail to. My alternative proposal is that such justification is a matter in part, of having strong evidence, evidence in the form of a particular type of phenomenology that I think some desires possess. This proposal will be defended in chapters 5 and 6.

I want to conclude this chapter by looking at a more recent version of the inner sense theory, a version that relies as well on an externalist epistemology—viz. Shaun Nichols and Stephen Stich's (2003) version. In the section that follows I grant for the sake of argument that the type of reliabilist epistemology Nichols and Stich embrace is tenable. I argue, however, that even granting this, their view is flawed.

---

[53] For instance, one might embrace "normal worlds reliabilism."

2b. Inner Sense & Access Dispositionalism Part II—Nichols & Stich

Armstrong and Lycan's view of self-knowledge is arguably the canonical way of understanding what possessing self-knowledge via inner sense amounts to. Shaun Nichols and Stephen Stich (2003) offer a more recent observational account of self-knowledge, one that can reasonably be classified as an inner sense view because of its commitment to the thesis that we arrive at knowledge of our own minds via introspectional mechanisms that cause us to form certain beliefs about the contents of our brains. The Nichols-Stich view is worth discussing here since they are specifically concerned with how we know our attitudes.

Nichols and Stich contend that self-knowledge of our attitudes is in part achieved via a monitoring mechanism that, in the case of our self-knowledge of beliefs, copies representations from our Belief Box, embeds the copies in the representation schema *I believe that* _____, and then places the new representations back in our Belief Box. In other words, our monitoring mechanisms take a representation that P in our Belief Box and produce the representation *I believe that P* as an output. Nichols and Stich think that such a process is involved in how we monitor our other attitudes as well.[54] Such a

[54] Nichols and Stich claim that, "The proposed mechanisms would work in much the same way to produce representations of one's own desires, intentions, and imaginings" (2003; p. 161). I should note here that the latter do not explain how we know our sensations. Goldman (2006) takes this to be a lacuna in their position, and he finds it troubling given that sensations do not appear to have representational content. He notes: "Without an associated propositional content for sensations, it is wholly unclear how they can be handled by [Nichols and Stich's] theory. The theory is a syntactic theory, which says that a monitoring mechanism operates on the syntax of the mental representations monitored. But... [sensations lack] syntax" (238). If Goldman is correct about this, then the Nichols-Stich model for the attitudes would not be applicable for states that lack content. I presume, however, that while Nichols and Stich do not explicitly embrace the view that sensations possess content, they hold that sensations are contentful. Goldman anticipates this response, and he counters it by claiming that not all pains have content (238). I am sympathetic with anti-representationalism with respect to sensations. But I need not defend the view that sensations lack representational content, and that therefore the Nichols and Stich account is incomplete. This is because I think there are more pressing worries we can advance against this view.

mechanism with respect to desires, according to them, "would be trivial to implement" (161). I think, though, that just the opposite is the case. I will now explain why by first turning to an objection to the Nichols-Stich view raised by Goldman (2006), an objection I think can be strengthened.

Goldman notes that on the Nichols-Stich view there is a question concerning how our monitoring mechanisms (MM) are able to determine what attitude to fill into the representation schema—*I Φ that P*—that gets dumped into our belief box. Nichols and Stich would have a ready-to-hand answer to this question, Goldman suggests, if the boxes in question were actual brain boxes. Goldman thinks this is the case because insofar as the boxes were actual brain boxes, MM might simply keep track of the box the initial sentence is taken from. Goldman proceeds to note, however, that brain boxes for Nichols are Stich, are merely a way of talking about functional roles. This is a problem, Goldman thinks, because it is unclear how a brain monitoring mechanism could determine that a piece of syntax has a particular functional role.[55] He in turn notes that Nichols and Stich are entirely silent with respect to this crucial question. Goldman then entertains the possibility that friends of this view could argue that there is a separate monitoring mechanism for each attitude type[56] and that this would explain how our monitoring mechanisms detect the functional roles of the attitudes. But, Goldman contends, such a view would be objectionably unparsimonious. Goldman's criticism of the Nichols-Stich view can be put in the form of a dilemma:

---

[55] Nichols and Stich use talk of boxology as a device for understanding functionality. See (2003; p. 11) for this claim.

[56] And ostensibly for each type of mental state in general insofar as all mental states possess representational content.

(1) Either (a) there is not a multitude of monitoring mechanisms on the Nichols-Stich view, or (b) there are.

(2) If (a), then it is not clear how a single (or limited set) of monitoring mechanisms determines the attitude type in question, and the view is not satisfactory.

(3) If (b), then Nichols and Stich are committed to a proliferation of monitoring mechanisms, and such a view is, because of its unparsimonious, not satisfactory.

(4) Therefore, the Nichols-Stich view is not satisfactory.

I am largely sympathetic with Goldman's objection. But I think some moves Goldman anticipates Nichols and Stich making are more troubling than Goldman makes them out to be. Consider again Goldman's claim that positing individual monitoring mechanisms for each mental state type leads to an unparsimonious view. Goldman, I think, is right about this. It does. But regardless of concerns with parsimony, it strikes me that positing a proliferation mechanisms is not going to resolve Goldman's initial worry about how monitoring mechanisms detect the functional role of the attitude in question on the basis of mere syntactical form. It is one thing to say that each attitude type has an individual monitoring mechanism, but it is an entirely different matter to remove the mystery surrounding how such individual mechanisms detect functionality on the basis of syntax. And it is indeed a mystery. For on a common and I think correct view of the attitudes, we can adopt different attitudes to the same content. But if this is the case, how do we know, merely on the basis of a given content, that we have a certain functionally understood

attitude when that syntactical structure can (partially) constitute any number of different attitudes? Even granting that there is a monitoring mechanism for each attitude, we still need an explanation of how, solely on the basis of syntax, such a monitoring mechanism could determine attitude type. But on the Nichols-Stich view all that we have good reason to think monitoring mechanisms can make us aware of, it seems, is content. In short, then, I do not think that positing a monitoring mechanism for each attitude type makes awareness of attitude type, on this view, any less problematic. And this, I think, is the true moral of Goldman's criticism.[57]

Perhaps my above objection is misplaced and positing an individual monitoring mechanism for each attitude type really does remove the mystery behind our ability to type-identify our mental states via such mechanisms. I want to argue now, however, that even if this is the case, the Nichols-Stich view faces a different problem—viz. that it

[57] A similar problem, I think, confronts another causal-based account of self-knowledge—Dretske's (1994) displaced perception account of self-knowledge. According to Dretske, we are aware of the fact that we are in a first-order mental state via an inference from our awareness of what is represented by that first-order state. For example, I am aware of the fact that I am tokening a mental state that represents a stack of papers via an inference I make from awareness of the stack of papers to having the first-order state. I wouldn't be aware of the stack of papers if it weren't for the fact that I am representing that stack of papers via a first-order mental state with those representational properties. There are a number of criticisms one can offer to such an account of self-knowledge. I won't rehearse them all here. One objection, however, that is relevant to the above discussion is an objection that Dretske is well aware of—viz. that it appears that his account of self-knowledge can (at best) explain our knowledge of the representational properties of our mental states; it cannot explain our knowledge of the *type* of mental state it is. This is because what Dretske-style inferences license are only conclusions about content. But, of course, most of us think we also possess self-knowledge of what mental state type we token. So Dretske's displaced perception account of self-knowledge is, at best, incomplete. Lycan (2003) presses this objection against Dretske, and the latter seriously considers embracing what we might call *mental-state type skepticism*—i.e. the view that we really can't know the type of mental state we token. See Dretske (2011) for a discussion of this type of skepticism. Now the Nichols-Stich account of self-knowledge is unlike Dretske's account in numerous ways. In particular, the former account is an introspectional account, while Dretske's account is *extrospectional*. But I am suggesting that the Nichols-Stich view has a similar problem. Their account, at best, can explain our knowledge of the content of our states, but it cannot explain our knowledge of what mental state types we are in.

makes knowledge of our attitude types far too easy to come by.[58] Consider that if there really is a mechanism for each individual attitude type, and the process is as "trivial to implement" (161) as Nichols and Stich claim it is, then why is it as difficult for us, at least in some cases, to identify what attitude we have towards a particular proposition? Do I *believe* that the next president will have liberal sympathies or do I merely *hope* the next president will? I reflect. I ponder this question. But perhaps I'm just not sure. Why, if we have an individual monitoring mechanism for beliefs and doubts respectively—monitoring mechanisms that enable us to not only recognize the content of our mental states but also the type of attitude we bear to such content—am I not able to easily arrive at such knowledge? If this process is as simple to implement as Nichols and Stich make it out to be, I should seamlessly be able to reliably self-attribute what attitude type I adopt towards the above content. But such is not the case.

I submit, then, that what Nichols and Stich take as a virtue of their theory (the triviality of implementing the self-monitoring process as they describe it) is actually a vice. It renders our awareness of attitude types far too facile. The upshot is this: even if positing a multiplicity of monitoring mechanisms eliminates the mystery behind our knowledge of attitude types, (something I deny above), the Nichols-Stich view is still implausible.

A proponent of the Nichols-Stich view might object at this point that all I have demonstrated, if I have demonstrated anything, is that we need an explanation of why the process is more difficult to implement than it appears to be. In response, given what

[58] Interestingly enough, I think a similar type of objection can be leveled against a radically different theory of self-knowledge—viz. Bonjour's acquaintance-like epistemology of attitudes. I will have much more to say about this in chapter V.

Nichols and Stich have told us about the self-monitoring process with respect to attitudes, and given that they do seem pushed into embracing the unparsimonious view that Goldman's dilemma saddles them with, it is not at all clear what that explanation would be. One can, of course, say that monitoring mechanisms, like any other evolutionary advantageous faculty, in rare occasions, can be unreliable. Granted. But that move would not eliminate the worry. The process would still be, it seems, much more successful at identifying our attitude types than we have reason to think we in fact are at such a task. In short, then, we need a principled explanation for why the process of attitude-type identification is, at times, much more difficult than the Nichols-Stich line makes it out to be. Of course, even if we had such an explanation, this would still not remove the mystery behind how a process that operates on syntax alone can identify the functional nature of our attitudes.

In this section I have argued that the Nichols-Stich view of self-knowledge is problematic. It is problematic because, while it might be able to explain how we have knowledge of the content of attitudes, it cannot explain in an acceptable manner how we have knowledge of the attitude we adopt towards that content. I think, then, that both the canonical versions of the inner sense theory as well as more recent attempts to explain our self-knowledge as an introspectional causal process founder. It is fair to say, however, that the accounts of Armstrong, Lycan and Nichols and Stich, and for that matter direct acquaintance accounts, are (currently) minority views, especially as accounts of how we know our attitudes. An increasingly more popular line-of-thought has it that we do not know our attitudes via an introspectional process, whether causal or

not. Those who deny the introspectional accounts offered above typically embrace one of the following broad approaches to the epistemology of attitudes: (i) transparency accounts, (ii) agential rationalist accounts, or (iii) constitutivist accounts. I will turn to (i)-(iii) in the next chapter, with an eye to whether these types of theories are any more promising with respect to AD. A number of philosophers will, I imagine, think they are; at the very least, they will think that (i)-(iii) are more likely to be able to explain our privileged access to dispositional states than the types of views of self-knowledge I have discussed in this chapter. I will, however, argue this is in fact not the case.

Chapter 3: Access Dispositionalism and Non-Introspectional Accounts of Self-Knowledge

We saw in the last chapter that introspectional approaches to self-knowledge such as acquaintance and inner sense theories of self-knowledge have difficulty making sense of how we possess privileged access to dispositional states. In this chapter, I examine whether a number of approaches to self-knowledge that might be thought to stand a better chance of vindicating AD are able to do so. In particular, I discuss whether (i)

transparency accounts, (ii) agential rationalist accounts of self-knowledge, and (iii) constitutivist accounts of self-knowledge can make sense of how we have privileged access to dispositional desires. What (i)-(iii) have in common is that proponents of these theories deny that we know our attitudes, at least in the typical case,[59] via a process of introspection. I argue that (i)-(iii) are unable to vindicate AD. Insofar as this is the case, this provides us with more evidence that AD is false. I further argue (i)-(iii) face daunting problems as theories of self-knowledge in general.

I begin with a discussion of one currently popular way of understanding our self-knowledge of attitudes—viz. transparency theories of self-knowledge.

1. Transparency Accounts of Self-Knowledge & the Epistemology of Desire:

A number of epistemologists are convinced that we know attitudes like belief and desire not through introspectional observation, but by examining features of the world these states are about. These philosophers follow Gareth Evans (1982) in holding that we know whether we believe, for example, that the next president will be a Democrat, by considering whether the next president will in fact be a Democrat. Moran (2001) defends a version of this view by invoking what he calls *the deliberative stance*. According to Moran, we know whether we believe that P by considering whether we *ought* to believe that P; this question in turn requires us to consider what reasons we have for believing this proposition. Moran explains what a purely extrospectional or transparent method of self-knowledge involves as follows:

[59] Some of these philosophers, as we will see, allow that we can know our attitudes via introspectional observation, but they do not view such introspectional means as the typical, or most significant, way in which we come to possess knowledge of these states.

> A statement of one's belief about X is said to obey the Transparency Condition when the statement is made by considerations of the facts about X itself, and not by either an 'inward glance' or by observation of one's own behavior (2001; p. 101).

Consideration of the facts about *X itself* in turn involves a focus on the reasons for holding X. Knowing one's beliefs, for Moran, then, is a matter of having one's reasons for that belief determine that belief. The method is thought to be transparent insofar as one directs one's attention not towards the belief itself, but towards one's reasons to believe.

Moran further suggests that it is not just our beliefs we come to know by adopting the deliberative stance, but our desires as well. His contention is that in determining whether we desire, e.g., a career in medicine, we consider whether we ought to desire that career, which in turn leads us to consider whether a career in medicine has features that make it worth desiring. Moran holds, then, that we come to know that we desire P by directing our attention to the reasons we have to desire P, reasons concerning the value of P. It is worth pursuing whether Moran's transparency account of the self-knowledge of desires makes plausible that we have privileged access to our dispositional desires.

1a. Moran's Account of Self-Knowledge & Access Dispositionalism

On Moran's view of self-knowledge, what is unique about knowledge of our own beliefs, desires, and other attitudes is not that we have a means of inwardly detecting these types of states, a means that no one else can use that leads to epistemically secure

knowledge; it is rather that the judgment-sensitive nature of these attitudes—i.e. the fact that attitudes such as beliefs and desires are capable of being evaluated and revised—enables us to arrive at knowledge of these types of states by considering the (normative) reasons there are for having the state in question. On the standard reading of Moran's view, when we engage in the deliberative stance and come to recognize that our reasons favor P, we in turn avow that P. This avowal, or declarative utterance made in light of one's reasons, enables us to somehow know that P. One might wonder, however, how engaging in the deliberative process enables us to acquire knowledge of the attitude in question. How, in other words, does making up your mind about whether to believe that P, make it the case that you know that you believe that P?

This question is pressing for the transparency theorist since, as Gertler (2011a) points out, in most typical cases of propositional knowledge, being a knower requires being sensitive to or controlled by that which is known. If I know, for example, that there are lights on in a certain room it is because I am sensitive to the fact that I know—viz. the lights being on in the room. On Moran's view, however, the reverse is the case; the facts known (i.e. facts about one's mental states) are sensitive to or controlled by the states of the knower, and this sensitivity somehow enables us to know these facts. So it is natural to wonder whether Moran's position explains how we know our own attitudes as opposed to explaining how we normatively shape such states. It is not surprising, then, that we find Lucy O'Brien (2003) and Sydney Shoemaker (2003) puzzling over how attention to the reasons we have for a particular attitude enables us to know the attitude in question.

In response to O'Brien and Shoemaker's concern, Moran (2003) offers a transcendental argument to the conclusion that we have the right to assume that the deliberative stance yields self-knowledge. While Moran does not explicitly spell out the details of such an argument his defense of this claim appears to be as follows:

(1a) We, qua rational agents, engage in critical reasoning about practical and epistemic matters.

(2a) Engaging in critical reasoning requires that we conceive of ourselves as agents whose attitudes are determined by our reasons.

(3a) We cannot verify that our attitudes are determined by our reasons since this would involve considering our attitudes in isolation of our reasons.

(4a) Insofar as exercising rational agency (i.e. being rational agents who engage in critical reasoning) requires operating as if our attitudes are determined by our reasons, and it is impossible to verify that they are, we have the right to assume that our attitudes are based on our reasons.

(5a) If (4a) we have the right to assume that we achieve self-knowledge of our attitudes by attending to our reasons for holding those attitudes.

(6a) Therefore, we have the right to assume that we achieve self-knowledge of our attitudes by attending to our reasons for holding those attitudes.[60]

Achieving self-knowledge by attending to our reasons, not to the attitudes themselves, is achieving self-knowledge via the deliberative process. We have, here, then, an argument

[60] This interpretation of Moran's defense of the claim that we have the right to assume that the deliberative stance yields self-knowledge closely follows Gertler's (2011a; p. 189) interpretation.

to the conclusion that we have the right to think the deliberative process affords us with self-knowledge.

But how successful is Moran's attempt to ground our self-knowledge in rational agency? I think it is not successful. In particular, I hold that premise (2a) of Moran's argument—the claim that engaging in critical reasoning requires that we conceive of ourselves as agents whose attitudes are determined by our reasons—is false. This is because, as Christopher Peacocke (1998) has argued, one can assess evidence and in turn base one's beliefs on the appropriate evidence, at the sub-personal level. Peacocke uses the following example to motivate this claim:

> Suppose you come home and see that no car is parked in the driveway. You infer that your spouse is not home yet... Later, you may suddenly remember that your spouse mentioned in the morning that the brakes of the car were faulty, and wonder whether she may have taken the car for repair. At this point, you suspend your original belief that she is not home yet. For you come to realize that the absence of the car is not necessarily good evidence that she is not home. If the car is being repaired, she would have returned by public transport Then finally you may reach the belief that she is home after all, given your next thought that she would not have taken any risks with faulty brakes (276).

Nothing in the line-of-reasoning that one goes through in the above example involves, it seems, conceiving of oneself as an agent whose attitudes are determined by one's reasons. Such awareness never enters into the picture. And yet, because such reasoning

involves the assessing of evidence and the basing of one's beliefs on that evidence, such reasoning appears to be critical in nature.[61] If this is correct, then we can engage in critical reasoning without conceiving of ourselves as agents whose attitudes are determined by our reasons. Another way of putting this point is that there is conceptual space for the possibility of agents whose critical reasoning takes places without conceiving of themselves as agents whose attitudes are determined by their reasons. Premise (2a) of Moran's transcendental argument is suspect.[62]

The above point would also seem to call into doubt (4a). Insofar as critical reasoning can take place sub-personally, we need not think that in order to exercise our rational agency we must operate as if our attitudes are determined by our reasons. We can instead hold that such rational agency does not necessarily involve the type of intellectualized procedure than Moran's argument depends on.

Although I have expressed doubts about the soundness of Moran's transcendental argument, I now want to argue that even if it is sound, such an argument does not help out the access dispositionalist. This is the case because the argument's conclusion is merely that we have the right to assume that we possess self-*knowledge* via the deliberative process; the argument does not get us to the conclusion that we achieve privileged access to our attitudes. The latter would require possessing highly epistemically secure knowledge of these types of states. Nor does it seem that there an

[61] Cassam (2014, p. 41-42) cites the above passage in defense of the view that critical reasoning need not require self-knowledge.

[62] Moran might object that the type of reasoning Peacocke's agent engages in is not *genuine* critical reasoning. I am not entirely sure what *genuine* critical reasoning would be. But note that if Moran adopts a more controversial account of critical reasoning that involves more than assessing and basing one's attitudes on one's reasons, then the plausibility of premise (1a) is called into doubt.

easy adjustment we can make to Moran's line-of-reasoning that will enable him to draw this stronger conclusion. We, after all, wouldn't need to have highly epistemically secure knowledge of our attitudes in order engage in the type of critical reasoning Moran thinks we engage in. Knowledge alone, it seems, would suffice. The upshot is that even if an access dispositionalist were to embrace Moran's transcendental defense of the deliberative process such an argument on its own would not enable her to adequately defend AD. We would need an additional reason to think Moran's approach to self-knowledge explains how we have highly epistemically secure knowledge.

I take the Peacocke-inspired objection to Moran's transcendental defense, as well as the fact that this argument does not help us arrive at the conclusion that we have privileged access to our mental states to suffice for putting to rest Moran's approach to self-knowledge as one that helps vindicate AD. But I want to mention next that there is yet another concern for any access dispositionalist who wants to claim that Moran's account enables us to understand how we possess privileged access to our *dispositional* attitudes. The worry is that it appears that the only attitudes that are known via the deliberative process, on Moran's view, are occurrent attitudes. Consider: in order to know one's attitudes via transparency, according to Moran, one must avow the attitude in question; i.e., one must actively commit oneself to the attitude in light of one's reasons. Insofar as avowals are (mental) actions that involve such a commitment, it would seem that the avowal that, e.g. P or I believe that P would not be a dispositional belief, but rather an occurrent judgment or belief that P. The same can be said for other attitudes as well. So, as Gertler (2011a) points out, it appears that the types of attitudes we can avow,

and hence know via the deliberative process, are not dispositional attitudes. If this is in fact the case, then Moran's approach to self-knowledge does not explain how we know our dispositional desires after all.[63]

One might contend that when we come to, e.g. occurrently believe that P, we also come to dispositionally believe that P. As I will argue in more detail below, this is not always true. In certain cases our occurrent beliefs and our dispositional beliefs come apart. But even setting aside this concern, what is important to note is that even if one's occurrent beliefs tend to match up with one's dispositional beliefs, on Moran's theory of self-knowledge one would not know one's dispositional beliefs via the transparency process. We would instead come to know our standing attitudes via inference from knowledge of our occurrent beliefs. The upshot is this: if the above line of reasoning is correct, then the transcendental argument Moran offers in defense of transparency, is only applicable to the self-knowledge of occurrent attitudes. So the rationalist underpinnings of this approach to self-knowledge even if sound would not vindicate access dispositionalism. I submit, then, that a fan of AD should look elsewhere for an account of self-knowledge that can legitimate their thesis.

1b, Byrne's Transparency Account of Self-Knowledge

Byrne (2012) in more explicit fashion than Moran, defends the view that a transparency account of self-knowledge can explain how we possess privileged access to our desires.

[63] This is a significant conclusion to arrive at given that some rationalists with transparency leanings, including Moran, appear to think that a purely transparent process generates knowledge of one' s dispositional states. Moran (2001), for example, writes that his project involves explaining self-knowledge regarding "various standing attitudes of the person such as beliefs, emotional attitudes, and intentions" (9).

But the former does not think the epistemic basis of knowledge of our attitudes is grounded in our rational agency. He instead claims that we have privileged access to our desires via the following rule:

DES If P is desirable, believe that you desire that P.

Byrne's thought is that a judgment that P is desirable[64] can provide the basis from which to infer that one desires that P. Insofar as there is an intimate connection between what we judge to be desirable and what we desire, DES might be a reliable enough rule to afford us with knowledge of our desires. And if that is the case, directing our attention to P can in turn generate knowledge that we want P.

It is important to note, however, that Byrne does not think following DES is what explains how we have *privileged access* to our desires, where following DES involves it actually being the case that P is desirable. He suggests instead that merely *trying* to follow DES generates reliable beliefs about our desires, the type of reliability that affords us with privileged access to these states.[65] As he contends, "Since one's desires tend to line up with one's beliefs about the desirability of the options, *whether or not those beliefs are actually true*, DES is *strongly* practically self-verifying" (178, emphasis added). In this context, being *strongly practically self-verifying* means that if one merely

[64] Byrne appears to think that what it is for x to be desirable is for x to have the type of qualities that tend to cause us to want x— properties such as being pleasant, agreeable, delectable, and goodly. See (2012; p. 76) for this suggestion.
[65] See (2012; p. 178) for this contention.

tries to follow DES, then, one's beliefs about what one wants are likely to be true.[66] The fact that DES is strongly practically verifying, i.e. that merely trying to follow DES leads to reliably produced true beliefs, according to Byrne, is what explains our privileged access to our desires.

Access dispositionalists might think Byrne's transparency approach to the self-knowledge of desires stands a better chance of vindicating their view than Moran's approach. Whether it can will in part depend on Byrne being right in thinking that: (i) our desires tend to line up with our beliefs about the desirability of the options, and (ii) this alignment allows for highly reliable beliefs about our desires.[67] Call the conjunction of (i) and (ii) the Desire-Desirability Reliability Thesis (or DRT). If DRT is false, then, trivially, trying to follow DES will not be *strongly practically self-verifying* and Byrne's approach will not vindicate AD. I think DRT is false. I will defend this claim by examining how DRT fares if desires are what most Strict Dispositionalists think they are—viz. states that generate action.[68]

1c. Access Dispositionalism, Byrne's Transparency Rule & the Standard Account of Desire

[66] Byrne notes that trying to follow an epistemic rule involves believing that the consequent of that rule is true *because* one believes that the antecedent obtains, regardless of whether the antecedent actually does obtain. See his (2012; p. 171) for this claim.

[67] *Belief* here is being understood broadly to include both (occurrent) judgments and dispositional beliefs. I should also mention here that for the purposes of this discussion, I am going to bracket any worries one might have about the externalist epistemology that Byrne appears to be embracing. See ch. 2 for a discussion of externalism/reliabilism vis-á-vis self-knowledge.

[68] I will also have a number of things to say below concerning how DRT fares if desires are understood along non-action based lines.

Recall that the most popular theory of desires construes them as dispositional states that dispose us to act to bring about the state of affairs in question. If this is what desires are then whether DRT is true will depend on whether our beliefs about what things are desirable are a reliable guide to how we are disposed to act.[69] In examining this question we might first ask (i) whether mistaken beliefs about what is desirable can provide a reliable guide to what we are disposed to do, and (ii) how common are mistakes about what things are desirable. These questions are important since, as discussed above, Byrne thinks DES explains our privileged access to our desires even if we make mistaken judgments about what things are desirable. But if DES is not reliable when we do make mistaken judgments about what things are desirable, and such mistakes are not uncommon, then DRT is false. So let us first consider whether DES is, in general, a reliable rule when our judgments about what things are desirable are false.

Consider the following mundane example: Sandy judges that pursuing a career in medicine is desirable. However, Sandy believes this only because he has been influenced by others in making this judgment, and his judgment is mistaken. Furthermore, Sandy's behavior suggests that he is not disposed to act in ways to bring it about that he has such a career (e.g., he puts off enrolling in pre-med courses; bringing up his future career in medicine puts him in a foul mood, etc.). Now if Sandy were to attempt to conform to DES, he would infer that he desires a career in medicine. But it seems clear that, insofar as desires are dispositions to act in certain ways, Sandy does not desire this. Cases in which (a) an agent S mistakenly believes that P is desirable, (b) infers that she desires

[69] In what follows, unless otherwise noted, when I use the term *desire* this should be understood as shorthand for what strict-dispositionalist action-based theorists of desire such as Stalnaker and Smith think desires are.

that P, but (c) is not disposed to act in ways to bring it about that P do not seem difficult to multiply. Given this, it seems reasonable to think that in a wide-range of cases, when one mistakenly judges that P is desirable, this is not going to lead to a true self-ascription.[70]

The above point, as I earlier implied, is not a problem for Byrne's view if we are, by and large, reliable judges with respect to what things are desirable. But, it must next be determined, whether we are such reliable judgers. Admittedly, when one focuses on more mundane cases, e.g. judging whether a glass of water is desirable, it seems like we would be quite reliable. However, when one focuses on desires that are not so closely tied to our bodily sensations, i.e. the type of desires Moran and other transparency theorists appear to be most concerned with, it is less clear that we are reliable determiners of desirability. There is, in fact, evidence in the social psychology literature that these types of judgments are often times based on the most accessible, easy to verbalize, and salient features or reasons, where these features or reasons might very well not be representative of all the features or reasons that bear on one's judgment.

Timothy Wilson and Dolores Kraft (1993), for example, describe an agent, Lucy, who attempts to determine whether she will ever marry her current partner. She comes to the conclusion that she won't marry this person on the basis of the fact that when she considers this question, what comes to mind is the dim future earning potential of her

[70] I think this is also the case if desires are dispositions to have certain pleasurable experiences. It might not be the case if desires are dispositions to judge that something is desirable or valuable; however, examples below will call into question this particular strictly dispositional account of desire.

partner.[71] However, if she were to consider the entire body of evidence in support of a future marriage, she would arrive at a different, and correct, conclusion about this possibility. It does not seem unreasonable to think that something similar might occur, and occur with some regularity, in the case of our beliefs about what things are desirable. The types of judgments are, after all, similar. And given the cognitive effort it would take to bring to mind all the considerations that might need to be laid bare to make an accurate desirability-judgment, it does not seem difficult to conclude that the type of evidence we would typically consider to make these judgments is the type of evidence Wilson and Kraft think we rely on.

It also bears noting that the above discussion of the ways in which we form beliefs about what is desirable implies that we often carefully reflect on what objects we find desirable. But why should we think this? If we really do attempt to conform to DES to know most of our desires, then, insofar as this knowledge is rich and widespread (and I imagine Byrne thinks it is), we would need to be making these types of inferences with some frequency. And it would be cognitively tasking to consider all the relevant features and/or reasons in making these desirability-judgments. It seems reasonable to think, then, that a large number of these types of judgments would not be the product of a meticulous process of gathering and evaluating evidence. To suggest otherwise would be to over-intellectualize the cognitive lives of agents. If people are not carefully considering all the relevant features that might need to be considered in order to arrive at a correct assessment of whether x is desirable, then this would seem to provide further support for

[71] This case is discussed in Gertler (2011b) in the context of her criticism of Byrne's theory of the self-knowledge of belief. It is being used here for different purposes.

the view that a number of these judgments might very well be mistaken. And if this is the case, then we might not be as reliable at determining what is and is not desirable as we would need to be in order for DES to be strongly practically self-verifying.

I now want to argue, however, that even if, in most cases, we are able to judge correctly that P is desirable, conforming to DES will not be a reliable enough rule to afford us with privileged access to our desires. The reason is that there appear to be a number of cases in which we recognize that P is desirable and yet are not disposed to act in ways to bring it about that P. Consider one of Byrne's own examples:

> Lying on the sofa, wallowing in my own misery, I know that going for a bike ride by the river is a desirable option. The sun is shining, the birds are twittering... these facts are easy for me to know, and my torpor does not prevent me from knowing them. If I concluded that I want to go cycling, I would be wrong. If I really did want to go, why am I still lying on this sofa? (177).

In the above example, going cycling is a desirable option. But it seems reasonable to suggest that Byrne does not want to go cycling. After all, as he would be the first to admit, he does not appear to be disposed in the least to get up and go cycling. This case of accidie, or lethargic apathy, is one in which desirability and desiring come apart. If such cases are not rare, this calls into question DRT.[72]

[72] It deserves mention here that the type of accidie cases Byrne discusses appear to call into doubt crude judgment-based accounts of desire—viz. the view that to desire that P is to judge or be disposed to judge that P. As Byrne's example illustrates, it seems like desiring that P can come apart from judging or being disposed to judge that P is desirable. Byrne, for example, judges that going cycling is desirable, but it

Byrne, however, has a ready-to-hand response to this objection. He claims that in cases like the one above we are not inclined to slavishly follow DES and misattribute a desire to ourselves. Rather, the above case points up that DES needs to be supplemented with a defeater. Byrne suggests the following:

> DES DEFEATER: In cases in which S knows that φ-ing is a desirable option and considers the question whether she wants to φ, S will not follow DES and conclude she wants to φ if she believes: (a) that she intends to ψ, (b) that ψ-ing is incompatible with φ-ing, and (c) that ψ-ing is neither desirable nor better overall than φ-ing (2012; p. 182).

As applied to the above example, DES DEFEATER entails that insofar as Byrne believes that he intends to sit on his couch, believes that sitting on the couch is incompatible with cycling, believes that sitting on the couch is not desirable, and knows that cycling is desirable, then he will conclude that he does not want to go cycling. DES, when supplemented with DES DEFEATER, it might be argued, can help us explain why Byrne's approach can explain our epistemically secure knowledge of our desires. This is because, if DES DEFEATER is true, it explains why Byrne wouldn't mistakenly judge that he desires to cycle in the above case, or in cases like it where desirability and desiring come apart. It would thereby ward off the charge that following DES is, in some cases, unreliable.

---

seems intuitively plausible to think he does not desire to go cycling. If this is correct, then such crude versions of judgment-based accounts of desire are untenable.

Ashwell (2013), though, notes that DES DEFEATER *alone* will not help Byrne avoid the above objection. As she points out, not all cases in which desirability and desiring come apart are going to be cases in which (a)-(c) are satisfied. In particular, desirability and desiring can also come apart in cases in which what one intends to do (i.e. Ψ) is judged, like that which one does not intend to do (i.e. Φ), to be desirable as well. For example, there could be cases in which Byrne judges that sitting on the couch *is in fact* desirable. But he would still not be able to reliably infer that, because cycling is desirable, he wants to go cycling. Byrne's defeater rule is silent on these types of cases, and therefore, DES even when coupled with DES DEFEATER is not going to render a reliable judgment in these circumstances.

Perhaps the above objection only reveals that Byrne needs to offer additional defeaters. But even granting this, I think DES DEFEATER does not solve the problem generated by the fact that desirability and desiring come apart. This is because DES DEFEATER, as I will now argue, is false. There are cases in which conditions (a)-(c) are met and yet it also seems reasonable to think that we would still conclude we want to φ on the basis of the fact that φ-ing is a desirable option. Here is one such example:

> I know that it's desirable to stay in tonight and read Borges's *Labyrinths*. I also believe that I intend to meet a business associate, Sam, for a night on the town. I believe that meeting up with Sam is incompatible with staying in and reading Borges. But because Sam is a lout, I further believe that a night on the town with this person is going to be undesirable.

In the above case I know that staying in and reading is desirable. But I also believe that: (a) I intend to meet Sam, (b) meeting Sam is incompatible with staying in and reading, and (c) meeting Sam is undesirable. In such a case, then, the conditions of Byrne's defeater rule are met. But why think I wouldn't conclude that I want to stay in and take a second look at "The Garden of Forking Paths"? I imagine I would conclude just this on the basis of my knowledge of the desirability of staying in and reading, and be upset by the fact that I had to meet up with Sam.

Now one might object that if I truly desired to stay in and read then I would do so. But this is false. One need not act on a desire one has. More specifically, if one is an action-based dispositionalist, one can make sense of how I do have such a desire. I am disposed to stay in and read tonight and I'll do so insofar as certain conditions are met. But in the above case those conditions aren't met. So out the door I go. It is not true, then, that in all cases in which the conditions of DES DEFEATER are met agents will reason as Byrne thinks they will. So DES DEFEATER is false. And it is not clear how to easily fix the problem.[73] Insofar as DES DEFEATER is untenable, Byrne cannot appeal to it in order to safeguard his account from cases in which desirability and desiring come apart.

Taken together, the above problems indicate that Byrne's account is susceptible to the charge that desirability and desiring come apart in a way that poses a serious problem

[73] One might think Byrne should just stipulate that in *most* cases in which (a)-(c) are met, one will not follow DES and conclude that they want to φ. This emendation, however, would be problematic. First, insofar as Byrne made this move, he would be conceding that there are cases in addition to the types of cases Ashwell mentions in which desirability and desiring come apart. Such cases, in sum, would, it seems, be enough to call into doubt DRT. Second, the types of cases that cause problems for Byrne's defeater do not seem uncommon. Given this, it is not clear that in *most* cases in which Byrne's conditions are met, one will respond in the way Byrne thinks they will.

for his account. Even if we are reliable judges with respect to what things are desirable, we still have good reason to believe that following DES is not going afford us with the type of epistemically secure knowledge needed to possess privileged access.

1d. Access Dispositionalism, Transparency, & Value-Based Accounts of Desire:

At this point, it might be thought that transparency accounts of self-knowledge have a much better chance of vindicating AD if desires are not states that dispose us to act, but rather states that are intimately connected up with value. More specifically, one might think that transparency theories of desire like Byrne's account are most plausible when coupled with the view that desires are—what philosophers such as Dennis Stampe (1987) and Graham Oddie (2005) think they are—viz. perceptions or appearances of goodness. After all, if desires are perceptions of goodness then it would seem more reasonable to think that our judgments about what things are desirable would track what things we desire.

The main problem, however, with such a suggestion in the context of this dialectic is that the most well-developed value-based accounts of desire in the literature are accounts that construe desires as occurrent, non-dispositional states. For example, Oddie's (2005) view is that desires are appearances of goodness. These appearances are—like the perceptual experiences they are analogous to—occurrent mental events. Such a view of desire would be rejected by a strict dispositionalist. Now the latter could, of course, propose a strictly dispositional value-based account of desire. But such an account does not seem to have much going for it. This is because if one is convinced that

appearances of goodness are intimately connected up with desiring, why hold that it is only being disposed to have such appearances as opposed to the appearances themselves[74] that are desires?

If one insists, however, that all desires really are being disposed to token appearances of goodness, there is a further problem. The problem is that it is not clear whether, on such a view, desires are mental states at all. Indeed, one might reasonably think such properties are non-mental dispositions to token mental states. This seems like a plausible suggestion when one considers that appearances of goodness are supposed to be akin to perceptual experiences, and we do not typically think that we have dispositional perceptual states. It seems more plausible to think of these dispositions as non-mental dispositions to token perceptual experiences.[75] Something similar, I think, should be said concerning dispositions of appearances of value. Insofar as this is correct, it is not clear how an SDD-ist can plausibly adopt a strictly dispositional value-based account of desire of the stripe embraced by Oddie and Stampe and hold that desires are mental states at all. At the very least, then, the burden is on a strict dispositionalist to come up with a plausible, strictly dispositional value-based account of desire. I am skeptical that such an account is in the offing.

There is, however, an even simpler way to insure that there is an intimate connection between (occurrently) believing that P is desirable and desiring that P. Such a

---

[74] Or at least embrace the view that appearances of value *and* being disposed to have such appearances are desires.

[75] The point I am making here is influenced by Robert Audi's (1994) discussion of the distinction between dispositional beliefs and dispositions to believe. I am arguing that dispositions to token appearances of goodness are more like dispositions to believe in that the latter are not, as Audi convincingly argues, mental states at all.

way involves pairing Byrne's account with the view that desires just are (occurrent) beliefs of value. Of course, no access dispositionalist will accept this account of desire since it would run up against their claim that all desires are dispositional states.[76] A strict dispositionalist could, however, embrace the view that desires are dispositional beliefs that Φ is desirable. And such a view might be thought to be a nice complement to Byrne's epistemology of desire. After all, it seems that when one tries to follow DES, one forms the judgment, as opposed to the dispositional belief, that P is desirable. This judgment, it might reasonably be thought, will line up with what one dispositionally believes to be desirable. And insofar as that happens with great regularity, one can in turn hold that trying to follow DES will be highly reliable after all.

The proposed view, however, is only as good as the underlying metaphysical view of desire. Is it at all plausible to think that desires just are dispositional beliefs that P is desirable? David Lewis (1988) has argued that results from decision theory demonstrate that desires cannot be beliefs. Lewis's argument hinges on the claim that, given the decision theory calculus, there are crucial differences between rational belief revision and rational desire revision. Lewis's controversial argument aside,[77] it seems clear to me that one can want to do P without being disposed in the least to thinking that P is desirable. To take a mundane example, an alcoholic can strongly desire a glass of wine, but not believe that such a drink is desirable. One might object that such an agent must see something good in having the drink or he wouldn't want it. But it seems perfectly

[76] It would also render Byrne's account circular. Byrne quite clearly thinks that one's judgments and beliefs about desirability can come apart from one's desires, so he is not a fan of crude belief-based accounts of desires.

[77] Price (1989) and others criticize Lewis's argument. Lewis (1992) in turn responds to some of these criticisms.

reasonable to suppose the alcoholic accurately believes there is nothing at all good in having that drink, and rather believes that everything about bringing about such a state of affairs is undesirable. In other words, it seems reasonable to hold that we can think that the object of our desire is undesirable *tout court* when we desire it.[78] Byrne, I think, would agree, which is why he is sympathetic with the view that desiring and believing that something is desirable can come apart.[79] If this is the case, then desiring that P cannot be a matter of dispositionally believing that P is desirable.

I now want to argue, however, that even if desires that P are dispositions to believe that P is desirable, this still would not insure a tight enough connection between desires and beliefs about desirability to render DRT true. The reason is that our occurrent judgments might not, and in a number of cases, will not track our dispositional beliefs. Cassam (2014) and Gertler (2011b) both defend this claim in their recent evaluation of Byrne's account of the self-knowledge of belief. The latter notes that phenomena like belief-perseverance—the fact that some beliefs continue to be held in the face of countervailing evidence for those beliefs—make it reasonable to believe that our judgments do not always coincide with our dispositional beliefs. Consider one such example. William might judge that it is desirable that a greater number of women should assume managerial roles. William, however, has deep-seated sexist prejudices that continue to influence his behavior (e.g. his hiring practices), and this behavior indicates a more stable, long-standing belief that it really isn't desirable for more women to assume managerial roles. It seems reasonable to think, in a case like this, that William's

[78] Michael Stocker (1979) defends this view at length.
[79] Byrne has mentioned to me in conversation that he is not sympathetic with belief-based accounts of desire.

evidentially based judgment comes apart from what he, unfortunately, dispositionally believes.

Evidence from social psychology suggests that belief-perseverance is not a rare phenomenon. Indeed, the alleged ubiquity of phenomena like implicit bias appears to indicate how prevalent it is to continue to dispositionally believe propositions that one lacks good evidence for, and in some cases, disconfirming evidence against. An agent, for example, might have evidence that systemic racism exists and good evidence in favor of the view that such racism can in part be combatted by affirmative action policies. She might further judge that implementing such policies is, all things considered, desirable. However, because of her implicit bias, this agent's behavior might reveal that she does not dispositionally believe that implementing such policies is desirable (e.g. she finds herself not supporting candidates who plump for such policies; she is less than eager to defend such policies in discussions with others, etc.) The seeming pervasiveness of such cases calls into doubt the claim that what we judge to be desirable will reliably line up with what we in fact dispositionally desire.[80] So even if an access dispositionalist were to couple Byrne's approach to self-knowledge with belief-based accounts of desire, such a view, I submit, would fail to vindicate DRT.

1e. Alternative Accounts of Desire, Transparency & Access Dispositionalism:

[80] Cassam (2014; Ch. 10) offers a detailed discussion of how implicit bias poses problems for transparency accounts of self-knowledge.

There are certainly more strictly dispositional accounts of desire than the accounts I have surveyed here. One might, for example, embrace a strictly dispositional pleasure-based account of desire of the type discussed in chapter 1. I think, however, that a number of the same problems an access dispositionalist faces if they couple Byrne's epistemology with an action-based or belief-based account of desire are applicable to these accounts as well. To take just one such problem, if desires are intimately connected up with the experience of pleasure, as some have thought, it is going to be difficult for the strict dispositionalist to come up with an account that views desires as mental states. Dispositions to have certain pleasurable experiences, I think, are not mental states proper, but rather dispositions to token mental states—viz. experiences of pleasure. But insofar as access dispositionalists think that desires are mental states, the above strictly dispositional account of desire would not be available to them. In short, then, the strict dispositionalist is going to be tasked with the challenge of coming up with a strict dispositionalist pleasure-based account of desire in which desires are in fact mental states.

Setting the above issue aside, pleasure-based accounts of desire have never struck me as very plausible. The reason is that it seems entirely possible for one to desire that P without either experiencing pleasure at the thought of P, or being disposed to experience pleasure when one entertains such a thought. I might, for example, desire to shovel snow out of my driveway in order to keep a promise I made to my neighbors. I can token this desire, it seems, without being disposed to take pleasure in the thought of shoveling of my driveway. If this is in fact the case, then a desire that P cannot be dispositions to experience pleasure at the thought of P.

One might object here that I am surely disposed to take pleasure in the thought of keeping my promise to my neighbors. But this need not be the case. Perhaps I regret making the promise in the first place. Or perhaps I'm a severe deontologist who finds no pleasure at all in fulfilling what I take to be my moral obligations, but nevertheless desires to fulfill such obligations. In such a case, it would seem, desiring and being disposed to have certain pleasurable experiences come apart. Of course it seems reasonable to think that in a wide-range of cases we do in fact take pleasure in the thought of satisfying a desire. But this, as the above examples work to motivate, need not always be the case.

In this section I have argued that Byrne's Transparency Account when coupled with the most plausible accounts of desire fails to render DRT true. Byrne's specific externalist version of transparency offers, I think, a more promising transparency defense of access dispositionalism than Moran's account. But even it falls short of providing support for this thesis. In the next section I will investigate an approach to the self-knowledge of attitudes that shares affinities with Moran's view—viz. Burge's agential rationalist position.

2. Burge's Agential Rationalism and Access Dispositionalism:

Burge (1996), like Moran, thinks that what is epistemically unique about self-knowledge is that this type of knowledge is not grounded in our introspectional grasp of evidence, the general reliability of our internal monitors, or the truth-conduciveness of certain rule-following behavior. It is instead grounded in the agential relationship we stand in with

respect to these attitudes. Unlike Moran, however, Burge does not claim that we achieve self-knowledge via transparency. In fact, Burge is relatively non-committal with respect to the details of how such self-knowledge is achieved.[81] In this section I am going to briefly explain Burge's approach to the self-knowledge of attitudes, as well as his defense of this approach. I proceed to raise objections to this defense. But my main contention will be that regardless of whether Burge's argument in defense of his theory of self-knowledge—a theory I will refer to as *agential rationalism*—is sound, it fails to vindicate access dispositionalism. My reasons for thinking this are similar to those offered above for why Moran's transcendental argument does not vindicate AD—viz. that even if Burge is correct, and our knowledge of our own attitudes is grounded in rational agency, this merely entails that we have *knowledge* as opposed to *privileged access* of our attitudes.

Before proceeding further, it will help to get clear on a fundamental notion in Burge's epistemology of self-knowledge, the notion of "entitlement." On one standard reading of how Burge understands this notion, entitlement involves both a reliability component and a permissibility component.[82] Concerning the former, Burge claims that one is only entitled to beliefs or judgments that are "well positioned to indicate the truth in normal circumstances" (184). Being "well positioned to indicate the truth" can be understood in terms of reliability. Reliability, then, is a necessary condition for entitlement. But Burge thinks that reliability is not sufficient for entitlement. One's belief must also be within one's epistemic right. To say that S's belief that P is within her *epistemic right* is to say that it is permissible for S to believe that P. For Burge,

[81] I refer to Burge's view as a non-introspectional account of self-knowledge because Burge is not committed to the claim that self-knowledge is achieved via an introspectional process.
[82] See Gertler (2011a) for this way of understanding entitlement.

permissibility is the normative component of entitlement. Burge thinks that when it comes to the self-knowledge of our attitudes, accounts such as the inner sense approach discussed earlier, can explain the reliability component of entitlement. But such accounts cannot explain why such beliefs are permissible for us. Burge's defense of this claim relies on his view of what it is to be a critical thinker.

Burge holds that critical reasoning involves basing one's attitudes on reasons that justify them. Such a process, for Burge, involves immediately embracing or abandoning a given attitude when one comes to recognize that one's reasons dictate that one should embrace or abandon the attitude. So, for example, if one recognizes that her evidence for believing her partner has been faithful is weak, she will, qua critical reasoner, immediately abandon this belief. On the other hand, if that same agent were to recognize that her evidence for the proposition in question is weak, but fail to abandon her belief, that agent would not be responding to the evidence as a Burgean critical reasoner does. She would fail to be operating as such because her reasons do not play the type of direct role they play in the former case. On Burge's view, then, in order to be a critical reasoner one's attitudes must be appropriately sensitive to one's reasons, where being appropriately sensitive to one's reasons is a matter of those reasons exerting a type of direct normative pressure on such states. It is this type of direct immediacy that Burge thinks observational accounts of self-knowledge fail to explain. By *observational accounts* of self-knowledge, Burge appears to have in mind not only inner sense accounts of self-knowledge, but also acquaintance accounts and transparency accounts in which the epistemology of self-

knowledge is explained purely in terms of reliability or another externalist account (e.g. Byrne's view).[83]

Why does Burge think that observational accounts of self-knowledge cannot account for the type of critical reasoning he thinks we engage in? The answer appears to be that Burge holds that an observational apprehension of a mental state does not involve grasping or apprehending that mental state as a state that is: (i) revisable, (ii) one we are responsible for, and (iii) one that is sensitive to our reasons.[84] According to Burge, if observational accounts were true, (i)-(iii) is something we could only do *after* observing the attitude, not by observing the attitude. Burge, then, appears to think that the type of self-knowledge afforded by observational self-knowledge yields a type of limited knowledge of our attitudes that is not in keeping with how critical reasoners grasp their attitudes. The idea, I take it, is that if we did not grasp our attitudes as having (i)-(iii), then such attitudes could not play the type of direct, immediate role they play in our critical reasoning. Consider: if we did not directly grasp that a particular belief of ours is a revisable, judgment-sensitive commitment that we are responsible for, then our

[83] It bears noting here that Burge is operating with a very restricted notion of critical reasoning. It seems reasonable to think that insofar as an agent merely evaluates her reasons for or against the adoption of a particular attitude, that agent is engaged in a type of critical reasoning regardless of whether she ends up basing her attitude on justifying reasons. Furthermore, I take it that we want to leave open the possibility that agents can engage in a type of critical reasoning despite the fact that their recognition that they should have a particular attitude fails to lead them to adopt such an attitude. We would also, it seems, want to allow for the possibility of critical reasoners who are simply bad critical reasoners. I have in mind here agents who simply fail to recognize what their evidence dictates or who reason in flawed manner with respect to the evidence they do have. So Burgean critical reasoning is really one particular *type* of critical reasoning. The fact that Burge is working with this more restrictive notion of critical reasoning, it might be argued, is not a problem for his view, since his objection to observational theories of self-knowledge can accommodate this fact. All Burge needs to argue, it might be suggested, is that there is a type of critical reasoning we do engage in and empiricist theories of self-knowledge cannot make sense of it. I will not evaluate this line-of-reasoning here since I think there are more pressing worries for Burge's view.

[84] Boyle (2009) and Moran (2001) hold a similar view. My understanding of how to answer the above question has been influenced by Gertler (2011a; p. 171-177).

recognition that we did not have evidence in favor of such a belief, could not immediately lead to our abandoning our belief. We would need to see that belief as being a judgment-sensitive commitment of ours that we can revise in order for the belief to be immediately revised in light of our awareness of this fact.

Let's call self-knowledge that involves grasping our attitudes as possessing (i)-(iii)—*critical self-knowledge*. It is this type of knowledge Burge thinks empiricists cannot account for.[85]

Burge goes on to argue that what is needed in order to possess critical self-knowledge is entitlement to one's self-ascriptions of one's attitudes, and in particular, that one have the epistemic right to make such self-ascriptions. It is the permissibility component, according to Burge that observational theories of self-knowledge leave out. On his view, it is the fact that we are rational agents who are obligated to meet particular rational norms that makes self-attributions of our attitudes permissible for us. We have the (epistemic) right to make such self-ascriptions because without doing so we couldn't meet particular rational norms (e.g. that our beliefs are supported by good evidence) that we are obligated to meet. Part of the epistemic basis, then, of our self-attributions of our attitudes is grounded in a normative component—viz. the epistemic right we have to make such attributions, a right we only have because of the agential relationship we stand in with respect to our attitudes. Empiricist accounts of self-knowledge fail to make sense

[85] Burge appears to acknowledge that we can, in particular cases, know some of our attitudes (or facts about those attitudes) via observational means. But such knowledge would not be *critical self-knowledge*. It is interesting to note, though, that if Burge thinks that entitlement is a necessary condition for all types of knowledge (or at least all types of self-knowledge) and entitlement involves a normative dimension, then it is not clear to me whether he can embrace the view that empiricist views of self-knowledge do afford us with knowledge of our minds in certain cases. This is because a number of these accounts do not build in a normative dimension to knowledge.

of this agential relationship because they are unable to account for the fact that at least in some cases we grasp our attitudes as having the properties needed in order to be Burgean critical reasoners.

A reasonable way of summarizing Burge's argument in defense of agential rationalism is as follows:

(1BUR) Qua rational agents, we have obligations to satisfy certain rational norms.

(2BUR) In order to satisfy some of these rational norms we need to be critical reasoners.

(3BUR) We cannot be critical reasoners without knowing our attitudes in a non-empirical manner.

(4BUR) Therefore, we can only meet our obligations qua rational agents, if we can know our attitudes in a non-empirical manner.

(5BUR) We are only obligated to do that which we can do.

(6BUR) Therefore, as rational agents we can know our attitudes in a non-empirical manner.[86]

I think the crucial premise in this argument is (3BUR).[87] I noted above that Burge appears to think this premise is true because he holds that empirical accounts of self-knowledge fail to explain how an agent can grasp an attitude as a judgment-sensitive revisable

[86] Burge does not explain his defense of agential rationalism in as clear a manner as we might hope. However, I take the above interpretation, which comes from Gertler's (2016) interpretation, to be an accurate representation of Burge's view.

[87] There is a lengthy literature on whether ought does imply can. Frankfurt cases have convinced a number of philosophers that one is obligated to perform some actions even if one cannot do otherwise. I am going to bracket worries about this premise here because I think there are more pressing worries with Burge's argument. Such a discussion would also take us too far afield.

commitment. An opponent of Burge's argument, however, might think there is an equally compelling non-rationalist account of what happens when we grasp our attitudes that is consistent with our *never* immediately grasping these states as having these properties. The alternative explanation is that we immediately grasp a belief as a belief and then after the fact we implicitly[88] and effortlessly attribute these properties to that attitude. One who held this view could claim that once we attribute these properties to the attitude we then proceed to critically reason in a self-reflective way about it. Given this alternative empiricist-friendly explanation, one might wonder why we should prefer the rationalist explanation to this account.

One possible answer to the above question is that we cannot grasp an attitude as the attitude it is unless we apprehend that it is a judgment-sensitive, revisable commitment. If this is the case, we couldn't grasp the fact that we have a certain belief and then, after the fact, implicitly attribute these properties to it because we couldn't grasp that the attitude in question is a belief at all without first grasping that it has these properties. But Burge does not appear to think this. He instead holds that we can have what we might call observational knowledge of our attitudes—i.e. knowledge of our attitudes via empirical means. What we cannot have is the type of critical knowledge via such means that he thinks we have. Insofar as this is the case, then, Burge appears to allow that we can achieve knowledge of our attitudes without grasping our attitudes as

[88] At the sub-personal level, that is.

judgment-sensitive revisable commitments.[89] I am not convinced, then, that Burge can successfully rule out the above response to his argument.

An alternative explanation of why (4BUR) is false is that we actually can be aware of our attitudes as revisable commitments via empirical means. One might, for example, argue that when we are acquainted with our attitudes qua the attitudes they are, we are (or at least can be) acquainted with the features of those attitudes, including properties such being judgment-sensitive, revisable, etc. Burge, it might be contended, has underestimated the type of introspective access we can possess to our attitudes, access that enables us to read off the attitudes the features he thinks they have. Such a response, I think, has promise, but it will also involve a fair amount of toil. One needs to offer an explanation of how we are able to grasp these attitudes as attitudes with these particular features. I will, for the moment, leave it open whether such an account is in the offing.

Regardless of whether the above responses to Burge's argument succeed, I now want to argue that Burge's defense of agential rationalism is of little use to the access dispositionalist. This is because the above argument does not get Burge to the conclusion that we have privileged access to our attitudes via non-empirical means. And this is what the argument needs to arrive at if it is to be of use to the access dispositionalist. A rationalist sympathetic with access dispositionalism needs to alter the above argument if she is to arrive at the conclusion that we have privileged access to our attitudes via non-

[89] A rationalist might think that, *pace* Burge, we actually can't come to apprehend that a given state is the type of state it is without (at least) grasping that the state possesses (a)-(c). Such a view seems implausible given that there might very well be other conditions that enable us to grasp e.g. a belief as a belief. Such properties might include phenomenal properties of the attitude that suffice for making the state in question the type of state it is.

empirical means. One way to do this that is in keeping with the spirit of Burge's argument is simply to change (3BUR) to read:

(3BUR*) We cannot be critical reasoners without possessing privileged access to our attitudes in a non-empirical manner.

One can then alter the other premises of the argument in order to arrive at the claim that we can possess privileged access to our attitudes via non-empirical means. The problem is that (3BUR*) seems false. One can be a critical reasoner in Burge's sense if one possesses reliable knowledge of one's attitudes in a non-empirical manner without the reliability being high enough to afford one with privileged access. In defense of this claim, note that in the original version of the argument Burge maintains that some rational norms apply to us as critical reasoners that would not apply to us if we failed to possess knowledge of our attitudes. But it is not at all clear why having highly reliable knowledge would be necessary in order for such norms to apply to us. Reliable knowledge, alone, would seem to suffice. If that is the case, then (3BUR*) is false.

The point I am making here is similar to the point I made above about Moran's transcendental defense of the deliberative method. It seems entirely possible for us to be critical reasoners without having epistemically secure knowledge of our mental states. Insofar as being a critical reasoner requires possessing knowledge of one's attitudes via non-empirical means, one can, it seems, be a critical reasoner without possessing the type

of epistemically secure knowledge required for privileged access. I think, then, that Burge's way of defending agential rationalism, even if sound, fails to vindicate AD.

But even setting aside the above privileged access/knowledge problem, there is an additional problem one who wants to defend access dispositionalism via a defense of Burge's argument faces. The problem is this: Burge's argument does not enable an access dispositionalist to claim that we have knowledge of our *dispositional* attitudes via non-empirical means. After all, Burge, it seems, would (and should) allow that we do not need to know *all* of our attitudes via non-empirical means in order to meet certain rational norms and/or be critical reasoners. It might very well suffice for us to meet such norms and be such reasoners if we know our occurrent attitudes in a non-empirical way. And if that's the case, the access dispositionalist is not entitled to infer, on the basis of Burge's argument, that we know our dispositional attitudes in a non-empirical way.[90]

Of course, a strict dispositionalist could argue that all attitudes are dispositional attitudes, and therefore we know at least some of our dispositional attitudes in a non-empirical way. But this highly controversial metaphysical claim is not part of Burge's view, and needs independent support. I argue in the next chapter that such a claim is false.

The point I am making is this: Burge's argument *alone* does not get one to the conclusion that we know our *dispositional* attitudes in a non-empirical way. We need an independent argument, one that Burge is not committed to, in order to arrive at this stronger conclusion. So, even setting aside the above privileged access/knowledge worry,

[90] It is in part difficult to determine whether on Burge's approach to self-knowledge we have knowledge of occurrent states, dispositional states or both via non-empirical means because Burge does not offer us an account of what means we do use to know such states.

a worry I think suffices to demonstrate that Burge's argument and rationalist view in general is of little help to the access dispositionalist, there is also the problem that Burge's argument alone does not demonstrate that we have even non-empirical *knowledge* of our *dispositional* states. Burge's agential rationalism, and in particular his argument in defense of it, then does not serve to help the access dispositionalist's cause.

I will conclude this chapter by investigating Shoemaker's (1996) constitutivist view of self-knowledge, another non-introspectional account of self-knowledge that attempts to ground self-knowledge, at least in part, in our rationality.

4. Shoemaker's Constitutivism & Access Dispositionalism:

Shoemaker (1994) defends the view that knowledge of our own desires is a matter of their being a constitutive relationship between first-order desires and second-order judgments about those desires. In other words, Shoemaker thinks that second-order states are in part constituted by first order states. He, however, does not think such a relationship obtains in virtue of one having the appropriate introspectional grasp of one's mental states as acquaintance theorists believe; instead he holds that this relationship obtains solely in virtue of an agent having the right conceptual capacities and the appropriate amount of rationality. Shoemaker, then, denies that what makes self-knowledge unique is that such knowledge is arrived at via a certain type of introspectional process.[91]

[91] Shoemaker attempts to establish the need for a theory of self-knowledge in which one's first order state bears a constitutive relationship to one's second order judgment about that state, by criticizing theories of self-knowledge in which such a relationship is not constitutive. His primary target is the inner sense theory of self-knowledge (what Shoemaker refers to as *the broad perceptual model of self-knowledge*) discussed

Perhaps the most tenable argument in defense of Shoemaker's constitutivist view is one that relies on the claim that a particular version of functionalism is true. Shoemaker, qua committed functionalist, thinks that mental states are to be understood in terms of their causal roles. Pain, for instance, can be defined in terms of certain inputs (e.g. tissue damage) and certain outputs (e.g. wincing, moaning, the desire to have the pain cease, etc.) Shoemaker calls that which actually performs this causal role, the *core realization* of the state in question. So, for example, insofar as the causal role of pain involves being caused by tissue damage and in turn causing pain behavior, and what plays this causal role in humans is c-fiber excitation, then the core realization of pain for humans is c-fiber excitation.

According to Shoemaker, the notion of a core realization, at least in rational agents, can perform two causal roles. It can play the causal role of pain and it can also play the causal role of the belief that one is pain. Here is a summary of Shoemaker's explanation for how this occurs:

> C-fibers firing in rational humans typically cause a desire for the pain to cease. Part of the causal profile, then, of the core realization of pain is this defining effect. Furthermore, a belief that one is in pain, typically causes in rational agents the desire for the pain to cease. So believing that one is in pain is to be understood in part in terms of the desire for the pain to cease. Given this, Shoemaker proposes that the effects that (in part) explain what it is to believe that one is in pain will be caused

---

in the last chapter. As I mentioned in the last chapter, Shoemaker offers the Argument from Self-Blindness against the inner sense theory (and any account of self-knowledge that posits the relationship between an introspected state and an introspective state as contingent).

> by the core-realization of pain. So the core realization of pain performs two causal roles—the causal role of being in pain and the causal role of believing that one is in pain.[92]

On Shoemaker's view, then, it is not that pain causes a belief that one is in pain; it is rather that given the dual causal role that the core realization of pain plays, pain constitutes—i.e., is the same mental event—as the belief that one is in pain.[93]

Shoemaker takes this detour into philosophy of mind in order to defend a startling conclusion about our access to our minds—viz. that for any rational agent S, if S is in pain, then, necessarily, S has a true belief that she is in pain. In other words, having the first order mental state guarantees having a belief that one is in that mental state, and since the first order state is the truth-maker for the belief in question, S will always have a true belief about what state she is in.[94]

It is important to note here, however, that even if this is correct, Shoemaker still has work to do to explain how we have self-knowledge of these states and not merely true beliefs.[95] One way to explain how we possess self-knowledge in a constitutivist friendly way is to claim that pain states can constitute one's reasons for believing that one is in pain. On such a view, one would have the justification certain internalists think is

---

[92] This summary of Shoemaker's view follows closely Gertler's (2011a) summary.

[93] The "mental event" here will be, on Shoemaker's view, a brain event. This does not mean that Shoemaker holds that pain and a belief that one is in pain are the same *type* of mental state. On his view, they aren't. According to him, they are different types of states because of the different type of causal role they play.

[94] Shoemaker puts the point this way: "All you have to add to the available first-order belief in order to get the second-order belief is the appropriate degree of intelligence, etc. It is not that adding this pushes the creature into a new distinct state, distinct from any it was in before, which is the core realization of believing that it has this belief. It is rather that adding this enables the core realization of the first-order belief to play a more encompassing causal role, one that makes it the core realization of the second-order belief as well as the core realization of the first order belief (Shoemaker 1994; 228-229).

[95] A number of philosophers have pointed this out—most notably Peacocke (2002).

necessary for knowledge. An alternative approach would involve embracing reliabilism or perhaps a view in which a safety condition needs to be met for possessing knowledge. If one opts for a safety condition—e.g. S would believe P only if P were true—then insofar as such a condition holds, Shoemaker can arrive at the view that all fully rational agents know they are in pain when they are.[96]

Shoemaker implies that a similar theory of self-knowledge can be offered with respect to desire. He, however, does not work out the details of such a theory. I think it is worth trying to do so, though, especially since a number of strict dispositionalists are functionalists who are likely going to think constitutivism offers us a tenable way of explaining how we possess privileged access to dispositional states.[97]

We can attempt to work out these details by considering whether what Shoemaker claims with respect to the epistemology of pain/sensations is applicable to the epistemology of desires. When we consider this question, however, we are immediately confronted with a problem. The problem is that it is not obvious how the core realization of desire (ostensibly a brain state of some kind), can play the dual causal role Shoemaker thinks the core realization of pain plays. The typical effects of desire, as we have seen, are thought to be actions. But actions, as Shoemaker himself holds, do not necessarily bring along with it a belief that one has the particular desire that motivates that action. Indeed, there are a number of examples of actions we perform on the basis of desires

[96] This is not to say that this is the epistemic position that Shoemaker holds. It is only a suggestion for what the story might be. Given Shoemaker's commitments, it is not an unreasonable to think he would be sympathetic with an externalist account such as this.

[97] I will explain why a number of strict dispositionalists are sympathetic with functionalism, and in particular a specific version of functionalism known as role functionalism, in the next chapter.

without having beliefs about the fact that we desire that which we are trying to acquire.[98] In defense of this view, consider that we often times come to learn that we have a desire through exchanges with others (e.g. my friend made me realize that what I really wanted to be was a literature major, not a pre-med student), or by observing our own behavior (e.g. I didn't know I wanted to be with that person until I kept blushing around him). In such cases, the agent already has the desire in question, but she is simply unaware of that fact and has to rely on other methods to discover it. Now, if there were a constitutive relationship between desires and beliefs about desires, then in these cases we would already believe that we have these desires.

Furthermore, there are numerous cases in which we are simply confused about whether we have a desire although we in fact have it. "Do I really want to buy this dress? I'm not sure," says the ambivalent shopper, even though she truly does want to buy it. Or, less trivially (perhaps), consider the troubled law student who frets over whether she should drop out of law school. As a result of her state of confusion she doesn't believe she wants to. However, it is clear to those around her that she truly does want to leave. These types of cases do not strike me as atypical.

Further support for the view that the core realization of desire does not play a dual causal role, can be marshaled by noting that there appear to be a number of cases where we believe that we don't have a particular desire when we in fact do. To take a mundane

[98] Nor, does it seem, would any non-essential manifestation of an underlying dispositional desire also entail awareness or beliefs about tokening a particular desire. The experience of pleasure at the thought of a particular state of affairs being realized, for example, would not necessarily entail that one believes or is aware of the fact that they desire that particular state of affairs (although, as we have seen, in some cases this might be good evidence that one has the desire in question.) I should also mention here that we might, of course, need to token beliefs (e.g. means-ends beliefs about how to accomplish bringing about a certain state of affairs) in order to act. In fact I think we do need such states. But such beliefs are not beliefs about the fact that we have a particular desire.

example, a grad student might sincerely believe she does not want a tenure track position at a research institution, while fretting over whether UCLA received her job application. Insofar as such an agent in fact has the desire in question, she would, according to the constitutivist position I am considering, believe she desires what she believes she doesn't desire. If such cases are common, then the constitutivist is faced with the problem of having to attribute contradictory beliefs to these agents.

Now it might very well be possible to have contradictory beliefs. But if these types of cases are common enough, we would have a much larger number of contradictory beliefs than it seems reasonable to think we have. On such a view, we would suffer from severe and somewhat extensive cognitive dissonance. It would seem more reasonable to think that in such cases we simply don't believe we have the desire in question. The moral to be drawn here, I think, is that it seems reasonable to think that the core realization of desires does not play the dual-role it needs to play in order for there to be a constitutive relationship between the first-order state in question and the second-order state.

An access dispositionalist with constitutivist sympathies might respond that all or most of the cases discussed above are cases of isolated failures of rationality, failures that would not occur for rational agents. But such a suggestion seems problematic. After all, a number of these cases are cases in which agents are unaware of their desires through no (obvious) fault of their own. These types of cases, then, do not appear to be cases of irrationality. Furthermore it is mistaken to suggest that such cases are limited or in some sense aberrant; on the contrary, such cases seem to be somewhat widespread. Insofar as

this is the case, then one who adopts the above position would be committed to the view that most agents, at least with respect to their desires, are irrational. But this view, for some of the reasons adumbrated above, is misguided. It is also a position that most rationalists would not warmly embrace.

The upshot of this discussion is that it is reasonable to think that even if the epistemology Shoemaker offers with respect to pain is tenable, it is implausible with respect to desire.

Conclusion:

In this chapter, and in the previous chapter, I argued that the most well-developed, viable theories of self-knowledge fail to vindicate AD. Certainly there are other theories of self-knowledge in addition to the theories I have addressed. But given the plausibility of (3) as well as the fact that the theories countenanced in the previous two chapters are unable to explain how we have privileged access to dispositional states, I think we are on good ground in taking (3) to be true. If (3) is true, then strict dispositionalists face an uncomfortable choice: either abandon SDD or deny that we possess privileged access to dispositional states. Faced with such a choice, I think most SDD-ists would reject (2) of our puzzle—the claim that we have privileged access to some of our desires. I think a better solution to our puzzle is to deny (1)—the thesis that all desires are dispositional states. In the next chapter I will present an argument to the conclusion that this thesis is in fact false, one that relies on the contention that some desires possess phenomenology. I turn to this task next.

Chapter 4: Strict Dispositionalism about Desires and the Phenomenology of Wanting

In the previous two chapters, I argued that there is good reason to think that proposition (3) of our puzzle is true. In particular, I argued that given the most widely accepted accounts of the self-knowledge of desires, it is reasonable to think we lack privileged access to dispositional desires. Confronted with the choice between denying (1) of our puzzle—the claim that SDD is true, or denying (2)—the claim that we have privileged access to some of our desires, I imagine most SDD-ists will deny the latter. In this chapter, however, I argue that strict dispositionalism about desire is false.

I defend the claim that SDD is false via an argument I call *the Phenomenological Argument* (PA). PA can be broadly formulated as follows:

(1PA) Some desires possess (a non-sensory) phenomenology;

(2PA) If some desires possess (a non-sensory) phenomenology, then not all desires are dispositional states;

(3PA) Therefore, not all desires are dispositional states.[99]

The argument is clearly valid. Its soundness depends on the truth of the first two premises. The most controversial premise of the argument is (1PA), the claim that some desires possess a unique phenomenology, a felt quality that is experienced by the desiring agent.[100] I will thus spend the bulk of the time in this chapter defending (1PA). (2PA) will receive attention after I have defended the first premise.

Insofar as PA is in fact sound, we will have found our way out of the puzzle. As importantly, though, if my reasons for denying SDD are to the mark, we will have also found the resources to develop an alternative, non-strictly dispositional account of desire. This alternative account of desire, I believe, is the key to explaining how we achieve privileged access to our desires; i.e. how (2) of our puzzle is true.

---

[99] As will become clear below, I hold that if some desires possess *any type* of phenomenology, then SDD is false. However, since I think that desires possess a non-sensory phenomenology, a claim that is an important part of my view of desire, I include the above parenthetical qualification in the formulation of my argument.

[100] Philosophers concerned with phenomenology sometimes use the locution "what's-it-like" to explain what phenomenology and in particular phenomenal states are. I prefer the phrase, "felt quality" as I think it is less vague than talk of what's-it-likeness.

Given the involved nature of this chapter, it will help to provide an outline of it. In part 1, I offer a detailed characterization of what it means to say that an attitude *possesses* non-sensory phenomenology. In part 2, I argue that the most popular arguments offered in defense of the existence of such non-sensory attitudinal phenomenology (NSAP), only get proponents of such arguments to the conclusion that non-sensory phenomenology exists. In other words, these arguments fail to get one to the conclusion that it is the attitude *itself* that possesses phenomenology as opposed to some non-attitudinal state associated with that attitude. As I contend, in order for these arguments to successfully arrive at the conclusion that NSAP exists, a position about the nature of attitudes must first be ruled out—namely a specific version of a view known as role functionalism. Role functionalists hold that attitudes are second-order properties of having some first-order property that plays a causal role. With respect to the popular arguments offered by proponents of NSAP, role functionalists can claim that (a) what possesses the phenomenology are first-order non-attitudinal properties, not second-order attitudinal properties, and (b) these latter properties, the attitudes themselves, lack phenomenology. In order to defend NSAP, then, we need an argument that eliminates this role functionalist position.

In part 3 I offer that argument. The argument I defend relies on embracing a view that stands in marked contrast to role functionalism—viz. realizer functionalism. Realizer functionalists hold that attitudes are not second-order properties but rather first-order realizer properties that play a particular causal role. I embrace this view of the nature of attitudes, and in turn argue that with respect to desires, states that realize the causal role

of desiring possess non-sensory phenomenology. If I am correct about this, the above line-of-reasoning serves to vindicate premise (1PA)—the claim that some desires possess a unique, non-sensory phenomenology. The argument I offer in defense of (1PA), however, is only as plausible as the realizer functionalist position that undergirds it.

In part 4, then, I proceed to explain why we should be realizer functionalists, and in particular why we should be realizer functionalists as opposed to role functionalists. I do so by discussing the major reason role functionalism is embraced over realizer functionalism, and in turn argue that such a reason is not a compelling one. In doing so, I complete my defense of (1PA).

In part 5, I defend (2PA), the claim that if some desires possess phenomenology, not all desires are dispositional states. I do so by focusing on what it means to say that a mental state is an occurrent state as opposed to dispositional state. Understanding this distinction will help explain why, insofar as some desires possess phenomenology, such states are not dispositional states. This section completes my defense of the Phenomenological Argument and, I believe, puts to rest strict dispositionalism about desires.

In part 6 of the paper, I provide a richer explanation of what desires are. In particular, I offer an explanation of why the phenomenology of attraction is the unique phenomenology of desiring, and why all occurrent desires possess phenomenology. I also explain what the relationship is between occurrent desires and dispositional desires, and in particular, why the latter should be countenanced as genuine desires. The claims I defend in this section, and specifically the contention that desires possess a unique

phenomenology, bear significantly on my views concerning how we possess privileged access to these states.

I: Non-Sensory Attitudinal Phenomenology Defined

The first premise of the Phenomenological Argument is that some desires possess (a unique, non-sensory) phenomenology. But what does it mean to say that attitudes such as desires possess non-sensory phenomenology? I will answer this question by first explaining what non-sensory phenomenology is. By *non-sensory phenomenology*, I mean phenomenology that is not identical with or reducible to any of the following types of experiences:

(a) perceptual experiences (e.g. olfactory experiences)

(b) the experiences of bodily sensations (e.g. the experience of hunger pangs)

(c) the experiences of imagistic imagery of a non-linguistic sort (e.g. the experience of thinking of one's distant friend)

(d) the experiences of linguistic imagery (e.g. the experience of thinking 'I'm tired' in words)

Non-sensory phenomenology, then, is phenomenology not encompassed by (a)-(d).[101] I understand *non-sensory phenomenology* this way in large part because a number of those

[101] Non-sensory phenomenology also sometimes goes by the name "cognitive phenomenology."

who deny the existence of NSAP, philosophers I will refer to as *NSAP conservatives*, are sympathetic with this construal of it.[102]

By *attitudinal phenomenology*, I mean phenomenology that at the very least is *possessed* by the attitude in question. To say that a token attitude A *possesses* phenomenology of type P is to say that A instantiates P. Any proponent of the existence of NSAP, philosophers I will refer to as *NSAP liberals*, who embraces *attitudinal phenomenology* must at least accept the view that it is the attitudes themselves that instantiate certain phenomenal properties. Consider: if a particular phenomenology were merely *associated* with an attitude, as opposed to being possessed by it, then it would not, strictly speaking, be the attitude that instantiated the phenomenology but the state associated with it that did. Such *associated* phenomenology, I take it, would not be deserving of the name *attitudinal phenomenology*.

Now most NSAP liberals hold that the relationship between phenomenology and the attitude that has it, is stronger than one of mere possession. The majority of these liberals appear to embrace the view that at the very least the phenomenology that certain attitudes (allegedly) possess *suffices* for making those attitudes the type of attitudes they are.[103] Most of these philosophers also hold that certain attitude types like desire have a characteristic, distinctive phenomenology that is associated with *only* attitudes of that type. To use David Pitt's (2004) term, these theorists think some attitudes have a

[102] Tye and Wright (2011)—from which the above quartet is largely adapted—are two such conservatives. Prinz (2011) is another conservative who embraces this characterization of non-sensory phenomenology.
[103] Horgan and Tienson (2002) hold this stronger view. Christopher Shields (2011) and Kriegel (2015) appear to think the same.

proprietary phenomenology, which individuates attitude types and enables us to type-identify attitudes on the basis of their phenomenology.

I am sympathetic with the view that desires and other attitudes possess a proprietary phenomenology, and in section VI of this chapter, I will defend that view. Here, however, what is important to note is that one can embrace the view that NSAP exists without embracing the view that some attitudes have a *proprietary* phenomenology.

NSAP liberals have offered a number of arguments in defense of the existence of non-sensory/cognitive phenomenology. At least some of these philosophers believe that the arguments they offer in defense of the existence of the latter also enables them to arrive at the conclusion that non-sensory *attitudinal* phenomenology exists.[104] But, with respect to two of the more popular defenses of cognitive phenomenology, we will see that this is in fact not the case. Explaining why these arguments do not get the NSAP liberal to the conclusion that NSAP exists will also help illuminate what needs to be done in order to defend the first premise of the Phenomenological Argument.

II: How *Not* to Defend Non-Sensory Attitudinal Phenomenology

The two most common types of arguments NSAP liberals have offered in defense of the existence of cognitive phenomenology—types of arguments they appear to think allow them to arrive at the conclusion that some attitudes possess non-sensory phenomenology—are (a) contrast arguments and (b) partial-zombie arguments. Neither

[104] Horgan (2011), for instance, thinks that the arguments he offers in defense of cognitive phenomenology justify the claim that NSAP exists.

type of argument, even if sound, gets the liberal to the conclusion that NSAP exists. Consider (a) first. So-called contrast arguments involve the presentation of two scenarios that are thought to be identical in terms of sensory phenomenology, but nevertheless involve a difference in phenomenology. Strawson (1994), for example, has us consider a case in which two subjects allegedly have the same auditory experience of certain sounds, but one subject understands the sounds as words and the other does not. It is then claimed that there is a phenomenological difference between them. Insofar as the sensory experiences of the two persons are identical, so the argument goes, the phenomenological difference must be a matter of non-sensory properties.[105]

Conservatives have responded to the above case by contending that the phenomenal difference between the agent with understanding and the agent who lacks it can be accounted for in terms of a difference in sensory phenomenology.[106] But setting aside the issue of whether this conservative move is tenable, what has not been made explicit in the literature is that even if the phenomenology in question is non-sensory in nature, such contrast arguments, *on their own*, do not get liberals to the conclusion that the attitudes in question possess phenomenology. To see why, it will be helpful to work with a type of mental state that is more obviously an attitude.

Christopher Shields (2011) has recently argued that curiosity[107] is a cognitive attitude that possesses phenomenology. Shields could (although he doesn't) attempt to argue for the existence of the NSAP of curiosity by embracing a contrast argument in which we are asked to consider the contrast between two agents—S and R—both of

[105] Siewert (1998; p. 275) defends a similar type of contrast case argument.
[106] Carruthers (2011), Prinz (2011) and Tye and Wright (2011) offer this response to contrast-style cases.
[107] By *curiosity*, here, I mean *the state of being curious*.

whom are presented with a question only S is curious about. One could then follow Strawson's lead in contending that there is a phenomenal difference between S and R, where this phenomenal difference is not a matter of a difference in sensory phenomenology. But note here that even if there is a non-sensory phenomenal difference between our two agents, this would not *entail* that it is the attitude of curiosity that possesses phenomenology. This is so because one can be a role functionalist about this attitude type.

Role functionalists, as noted above, hold that mental states, including attitudes, are second-order states of having a first-order state that plays the causal role of the attitude in question. The first-order state, according to the role functionalist, need not be (and in many cases will not be) an attitude proper. So while a role functionalist can accept that there is a phenomenal difference between the two cases, the phenomenal difference, it can be contended, is a matter of S tokening a realizer state that plays the causal role of curiosity; it is not a difference in phenomenology at the second-order level. S's realizer state, it can then be contended, is not an attitude proper, but rather that which *realizes* or *fills* the causal role of an attitude type. Given this, a role functionalist can embrace the view that curiosity itself (i.e. the second-order property) is not a phenomenal state. Hence, contrast arguments with respect to attitudes like curiosity do not get one to the conclusion that it is the attitude itself that instantiates a non-sensory phenomenal property. In order to arrive at that conclusion, one would need to deny the above role functionalist move by either defending the view that: (i) role functionalism about the attitudes is false, or (ii) if role functionalism is true, second-order role states that are the

attitudes are states that possess phenomenology.[108] Neither (i) nor (ii) is easy to defend.[109] The point, then, is that contrast arguments fail, *on their own*, to enable NSAP liberals to defend their view. More work would need to be done.[110]

Consider (b) next. Partial-zombie arguments involve the claim that there are (metaphysically) possible beings that lack sensory phenomenology and yet still possess a phenomenal life. The possibility of such beings, it is claimed, is supposed to provide us with (at least) a defeasible reason to think there must be some states—typically attitudes—that possess non-sensory phenomenology. Kriegel (2015) offers an example of such an argument. He has us imagine a creature, Zoe, who lacks sensory experiences, but who nevertheless has certain phenomenological experiences such as those involved in realizing the truth of mathematical facts. Kriegel thinks that since Zoe lacks all sensory phenomenology, the phenomenal experiences she tokens must be non-sensory in nature. Horgan (2011) offers a similar type of argument with respect to different cognitive and conative states—e.g., certain types of beliefs and desires. He takes the conclusion of his discussion of partial-zombies to be that these attitudes instantiate non-sensory phenomenal properties.

---

[108] As opposed to the realizer states being the only states that possess phenomenology.

[109] As I will discuss below, I think (i) can be defended. However, defending this claim will take some work. I think that (ii) is false. But in this chapter, I won't defend this claim at great length given that I argue that role functionalism is false.

[110] It might be objected that contrast arguments were never intended to get liberals to the conclusion that NSAP exists. But if *understanding* and states like it are cognitive attitudes, then this objection, I think, misses its mark. Proponents of these arguments such as Strawson (1994) and Siewert (1998) do take these arguments to enable them to arrive at the conclusion that such states possess non-sensory phenomenology. But even if the types of states referenced in these arguments are not attitudes proper, the important thing to note is that contrast arguments do not suffice for getting the NSAP liberal to the conclusion that NSAP exists. At the very least then, certain arguments NSAP liberals have offered in defense of the non-sensory phenomenology non-attitudinal cognitive states are thought to possess, cannot be adopted for more commonplace attitudes.

As with contrast cases, conservatives have ways they can push back against partial-zombie arguments.[111] But even if these arguments provide liberals with a defeasible reason to think non-sensory phenomenology exists, they do not permit us to conclude that NSAP exists. This is because a conservative might hold that the phenomenal states that Zoe allegedly tokens are realizers of second-order states, where these realizers are not attitudes proper. They might further contend that the second-order role states lack phenomenology. The NSAP liberal would thus need an additional argument to arrive at the conclusion that the particular states that possess phenomenology in partial-zombie cases are in fact attitudes.

The point I am making here mirrors the point I made above concerning contrast arguments. The role functionalist can accept that in partial-zombie cases there is a phenomenal difference of a non-sensory nature, but deny that this phenomenal difference indicates that it is the attitudes themselves that *possess* phenomenology. The upshot, then, is that partial-zombie arguments at best get the liberal only to the conclusion that we have a defeasible reason to think that some sort of non-sensory phenomenology exists. Similar points, I think, can be made about other liberal defenses of NSAP.[112]

---

[111] A conservative could, for instance, argue that our phenomenal lives would mirror Zoe's phenomenal life if we too lacked sensory phenomenology. The reason, it could be argued, that we don't realize this is because we aren't aware of the various ways our sensory phenomenology actually constitutes the phenomenology Kriegel thinks Zoe lacks.

[112] For example, Shield's (2011) parity-arguments. Briefly put, Shields argues that if one holds that mental states such as sensations are phenomenal states, then one should think that certain attitudes possess non-sensory phenomenology. His strategy is to point out that some features that sensory states possess are also features that cognitive states possess, and that positing NSAP is the only way to explain this similarity. I lack the space here to address Shield's parity-arguments in detail. I do, however, want to note that I think such arguments do not get one to the conclusion that NSAP exists for similar reasons to the ones adduced above with respect to contrast arguments and partial-zombie arguments. I also worry that regardless of this problem, Shields arguments will be seen as question-begging by conservatives.

What has gone wrong here for the NSAP liberal is that the main arguments they offer involve an implicit transition from the contention that (i) non-sensory phenomenology exists to the claim that (ii) non-sensory *attitudinal* phenomenology exists. But, as I have argued, (i) does not entail (ii). Nor do these types of arguments give us any reason for thinking that insofar as (i) is true, (ii) must be true as well. They fail to do so because contrast arguments and partial-zombie arguments do not provide us with good reason for thinking that either: (i) role functionalism about the attitudes is false, or (ii) if role functionalism is true, the second-order role states that are attitudes possess phenomenology.[113] And this is precisely what the liberal needs to defend in order to avoid the role functionalist move I have been discussing. NSAP liberals, then, need a defense of their view that does not focus merely on demonstrating the existence of non-sensory phenomenology, but also gets them to the conclusion that such phenomenology is actually possessed by some attitudes. In the next section, I offer such an argument. I do so by looking at a view in the literature on the phenomenology of desire I am largely sympathetic with—namely, Ruth Chang's (2008) position.

III: A Defense of the NSAP of Desire

Chang (2008) has recently contended that some desires have a characteristic type of phenomenology, which she describes as an experienced attraction to the desired object.

[113] It might be suggested that proponents of contrast arguments and partial-zombie arguments have independently motivated views about the nature of attitudes that are being assumed to be true when they offer such arguments. If these independently motivated views are true, then these arguments really would get liberals to the conclusion that NSAP exists. But note that one who offers this suggestion is conceding the main point of this section—viz. that contrast arguments and partial-zombie arguments alone don't get one to the conclusion that NSAP exists. One must also rely on controversial positions about the nature of attitudes as well. In the next section of the paper I make these controversial positions about the nature of attitudes NSAP liberals need to rely on explicit.

Acknowledging certain conceptual and linguistic limitations, I think what Chang calls attraction can most aptly be characterized as the experience of being *drawn* to the desired object. Such a *draw* can be faint in cases of weaker or mundane desires to the point that it can be tokened without an agent taking notice of this fact; or it can take center stage in one's mental life in cases of passionate, extremely intense desires. One can experience such an attraction because one *sees* the good in a certain state of affairs and comes to want it to obtain on that basis, or because one is simply caused by a certain bodily sensation to be attracted to the object. This type of phenomenology is familiar, I think, to anyone who has ever longed to see a loved one, craved a glass of wine, or yearned to see their child do well at a sporting event.

Other theorists, including NSAP conservatives, acknowledge the existence of such experiences. Tye (2015), for example, contends that there is something it's like to be drawn to an object[114] and that this experience is intimately connected up with desiring.[115]

I think that attraction is a genuine phenomenon. But I also hold the more controversial view that attraction is a type of non-sensory phenomenology that *some* desires *possess*.[116] Given this, I recognize the burden of needing to defend both the claim that (a) attraction is non-sensory in nature as well as the claim that (b) some desires

[114] "Object" here is being used in a neutral way to refer to that which we desire. A number of philosophers of mind think that what we desire is the realization of certain states of affairs; see Smith (1994) for this common view. Recently, however, some have argued that what we desire are, in some cases, concrete particulars; see Montague (2007) for this alternative position. I need not weigh in on this debate here.

[115] Tye writes: "Often when we strongly desire something, we experience a feeling of being 'pulled' or 'tugged.'" The type of experience Tye references here, I think, is akin to what Chang calls *attraction*. Tye goes on to suggest that this type of experience is reducible to a set of sensory experiences. I will address this typical conservative move below.

[116] I emphasize the word "some" because I do not think that *all* desires possess phenomenology.

possess phenomenology.[117] Contrast arguments and partial-zombie arguments, as I showed above, might establish (a). But they do not demonstrate that (b) is true. How should NSAP liberals go about defending (b)? I think we can get a good idea of how to answer this question by returning to a discussion of functionalism vis-á-vis NSAP liberalism.

I noted above that an NSAP conservative can reason as follows: (i) role functionalism is true and hence all attitudes are second-order states, and (ii) these second-order states lack phenomenology. But there is an alternative view, realizer functionalism, that makes it much more plausible to hold that it is the attitudes themselves that possess phenomenology. As noted above, realizer functionalists hold that an attitude is not a second-order state of having a first-order state that realizes a particular causal role C, but rather is the first-order state that realizes C. This position has been championed by D.M. Armstrong (1968) and David Lewis (1980), among others, in defense of reductive physicalism—the view that types of mental states are reducible to types of brain states.[118]

Reductive physicalism aside, if one embraces realizer functionalism then insofar as the state which performs the causal role of the attitude in question possesses phenomenology, it would follow that the attitude itself possesses phenomenology. An adequate defense of realizer functionalism, then would not only block the above NSAP conservative move, but it would also open up the possibility that NSAP exists. While I

[117] Chang does not provide a defense of (a) and (b). This is perhaps because her primary concern is not with the NSAP liberal/conservative debate, but with the debate over whether desires can provide reasons for action. As will be noted below, Chang holds, *pace* Scanlon (1998), Raz (2002), and Parfit (2004), that some desires do provide reasons for action.

[118] That being said, realizer functionalists like Armstrong and Lewis have typically not been sympathetic with the view that NSAP exists.

will defend realizer functionalism at length later, I now want to argue that we have good reason to believe that, in some cases, states that realize the causal role of desiring are states that possess the phenomenology of attraction. In order to defend this contention, I first need to explain what I take to be the causal role of desire.

On the view I am sympathetic with, desires are states that tend to be caused by certain mental states (e.g. perceptual experiences, beliefs, and other desires)[119] and in turn tend to generate action. For example, the desire for a glass of Syrah might be caused by the perceptual experience of a bottle of Syrah. Such a desire would in turn tend to cause, when certain conditions obtain,[120] one to act in Syrah-seeking ways. The causal role of desire, then, on my view, is action-based. The view that desires are intimately connected up with action, as discussed in chapter 1, is arguably the standard view of desire among philosophers of mind.

There are, however, objections that have been leveled against the action-based approach to desire. A critic could, for instance, argue that the above account of the causal role of desire does not distinguish desire from other types of mental states. It fails to do so, according to the objector, because there are other states that generate action besides desire. I lack the space to countenance all the types of states that might be thought to play the causal role I am claiming desires play. But I do want to go some way towards responding to this worry by considering a popular candidate for a type of mental states

[119] Admittedly, it is not easy to make more specific what types of perceptual experiences, beliefs, and other desires typically cause desires, although some would be sympathetic with the view that, e.g., the types of beliefs that tend to cause desires are beliefs about the goodness of states of affairs obtaining.
[120] Such conditions would include having the appropriate beliefs and lacking stronger conflicting desires.

that can be thought to generate action besides desire—viz. belief, and in particular, belief that one has a particular duty or moral obligation.

G.F. Schueler (1995) among a host of other ethicists, claims that the belief or judgment that one has a particular duty can generate action in the absence of desire. In defense of this claim, Schueler uses the example of being motivated to go to a PTA meeting on a cold winter night because he views it as his duty to attend. Schueler contends that it cannot be a desire that gets him out the door since he does not want to go to the meeting. What moves him in this scenario, is rather, according to Schueler, the judgment that he has a duty to attend.

Schueler's example and others like it, however, can plausibly be explained as a conflict of desires with one stronger desire winning out. Indeed, Schueler can reasonably be thought to want to perform his duty. Wanting to perform his duty, in turn, gets accorded more weight in his decision calculus than wanting to remain inside. This explanation seems preferable to Schueler's explanation, given that it seems reasonable to think that if he did not *desire* to perform his duty, he wouldn't leave the comforts of his home. Schueler-style cases, then, fail to show that beliefs/judgments that one has a particular duty can perform the same causal role as desire.

Now I think it is more reasonable to hold that pro-attitudes such as intentions, hopes, and wishes generate action. If that is the case, then the above functional characterization of desires is too broad. One way, though, of maintaining the view that the above characterization is fine as it stands is to embrace the view that intentions, hopes, and wishes, are complex mental states involving beliefs and desires. So, for

example, we might understand intention as a complex state involving desiring that Φ and believing that one will Φ. Similarly, we might construe hope as a complex state involving desiring that Φ and believing that it is at least possible (albeit, perhaps, improbable) that Φ obtains.[121] And we might hold that wishes are a conjunctive, complex mental state involving wanting Φ to obtain and believing there is little or nothing one can do to bring about Φ. What makes these attitudes pro-attitudes is that they are involve an action-based component (i.e. the desire) that partly constitutes the complex state. Desire, then, on this view, is the fundamental pro-attitude, the type of attitude that makes intentions, hopes, and wishes partly motivational states.

One reason one might embrace this view is that there does seem to be both a cognitive-belief like component to these states as well as an action-based component, and what distinguishes such states is the type of cognitive stance we adopt towards the proposition in question. Wishes, it might be thought, appear to differ from hopes in terms of the type of beliefs one has about the object of such states. And hopes differ from intentions in that the latter does seem to involve a belief that one will bring about the object of one's attitude. So these pro-attitudes, it can be argued, do not call into doubt the above functional characterization of desire.[122]

---

[121] We might also need to add to this analysis that one believes that it is it not entirely up to her whether Φ obtains. I leave open the possibility that there are other refinements that would need to be made to the above analysis, but the general idea should be clear enough.

[122] An alternative approach to these other pro-attitudes, one that I will not explore here, but one that I think demands further attention, is to argue that there is indeed a "motivational component" to these states that is not the result of the attitude being partially constituted by a desire. This, however, does not call into doubt the spirit of the above functionalist account of desire. What it demonstrates is that we need a more precise characterization of the functional of all of these pro-attitudes, one that clarifies in what sense all of these states are action-based or motivational in nature, but also why they are all unique (i.e. non-complex) states. As I implied above, I will not pursue this option here, but I do think it merits further attention.

While the above discussion worked to fend off the charge that there are other types of mental states besides desire that generate action, a less pressing, but still serious objection to the above functional characterization of desire is that some desires fail to generate action. If this is in fact the case, then it will be pointed out that the causal role of desire is not action-based. Schroeder (2004) has recently argued that there are cases of desiring where the agent is not in the least bit motivated to act. He has us imagine an Ancient Greek mathematician who allegedly desires that π be a rational number despite the fact that (a) the mathematician believes there is nothing she can do to bring about this state of affairs and (b) there is in fact nothing she can do. In such a case, Schroeder contends, the agent will not be motivated at all to bring about the state of affairs in question. If Schroeder is right about this, then desires are not sufficient for generating action, and hence, the above functional characterization of desire cannot be right.

I am not, however, convinced that Schroeder's example, and others like it, call into doubt the above view. In the π-case, I think a plausible explanation of why the agent does not act to bring about the state of affairs in question is because the agent does not have the appropriate beliefs. Indeed, if the agent were to believe there was something she could do to bring about the object of her desire, e.g. making sacrifices to the gods on Olympus, then, all things considered, I think she would do this. This reveals that it is not implausible to think the desire the mathematician has generates action after all. Like all states of wanting, the desire the mathematician has would generate the appropriate action given the appropriate beliefs. I see in these cases, then, no reason to doubt that desires are essentially motivational states. There are, of course, other objections that have been

offered to the above functional characterization of desire. But given the intuitive plausibility of the Humean view of desire and the fact that two major objections fail to call it into question, I think we are on fairly good ground in thinking that this characterization is to the mark.

Insofar as the causal role of desire is action-based, what reason do we have to think that states with attraction generate action? To answer this question, I suggest we turn to a passage from Theodore Dreiser's *Sister Carrie*. In the passage, Dreiser describes his protagonist, Carrie's, numerous desires to own certain material goods upon visiting a Chicago department store for the first time. Here is Dreiser's memorable description:

> Carrie passed along the busy aisles, much affected by the remarkable display of trinkets, dress goods, stationery, and jewelry. Each separate counter was a show piece of dazzling interest and attraction. She could not help feeling the claim of each trinket and valuable upon her personally, and yet she did not stop. There was nothing there... which she did not long to own. The dainty slippers and stockings, the delicately fritted skirts and petticoats, the laces, ribbons, hair-combs, and purses, all touched her with individual desire (77).

Dreiser goes on to speak of Carrie "feeling the drag of desire" (77). I take it that what Dreiser imagines Carrie experiencing here is a paradigmatic example of what I have been calling attraction. Carrie experiences an attraction to the state of affairs in which she owns these items; she is, in some obvious sense, drawn to the objects of her desire and she is well aware of this experience. Let us change the story slightly and stipulate that

Carrie has enough money to purchase the products in question. What explains why she purchases these items? The best explanation, I submit, is that "feeling the drag of desire" does. In other words, attraction moves her to act when coupled with the right beliefs. After all, if one is drawn to having a certain state of affairs brought about, then *ceteris peribus*,[123] it seems reasonable to think one will act in ways to bring this state of affairs about. Cases like Carrie and the department store, I submit are ubiquitous. It seems, then, that we are on good ground holding that in certain cases, what realizes or fills the causal role of desire (i.e. the states that actually move us to act) are states with the phenomenology of attraction.

It seems, then, that in some cases states that possess attraction realize the causal role of desire. And, if that's the case, then insofar as realizer functionalism is true, some desires possess phenomenology.

Now conservatives can grant the above line-of-reasoning and still deny the existence of NSAP. They can do so by arguing that the phenomenology of attraction is sensory phenomenology. The NSAP liberal, then, needs an additional defense of the claim that attraction is non-sensory in nature. I noted above that contrast arguments and partial-zombie arguments might provide defeasible reasons for thinking that non-sensory phenomenology exists. The importance of such arguments, I think, is that they dialectically force the conservative into having to account for the phenomenological difference between, e.g. an agent who experiences attraction and an agent who doesn't by appealing to sensory experience. Prinz (2011) and Tye (2015) attempt to do just this. The

[123] This clause is needed because if the agent does not have the appropriate beliefs or has other conflicting, stronger desires, then the agent will not act in ways to bring about the content of her desire.

former claims that desires might be tokened along with emotional-based experiences like anxiety, apprehension, or anticipation.[124] He's right. They might be. But a reasonable response to Prinz's suggestion is to point out that attraction can be experienced in the absence of such emotion-based experiences. It seems possible, for instance, to experience the drag of desire without anticipating obtaining the object, perhaps because one does not believe the object can be obtained. Furthermore, it seems possible to experience attraction without feeling any type of anxiety, apprehension, or even delight.[125] If attraction can be experienced without experiencing these emotion-based experiences, then, the former cannot be reduced to the latter.

A more plausible suggestion would be that some desires possess certain sensations such as hunger pangs, which are part of the above quartet of sensory experiences mentioned above. Alvarez (2008) appears to defend the view that *bodily appetites*—i.e., desires we have in virtue of possessing certain bodily needs—possess sensations such as the experience of being hungry. Now I am not sympathetic with account of desire. But even granting that Alvarez's view is correct, her position is compatible with some desires—viz., desires that are not bodily appetites—possessing attraction. Alvarez's account, then, does not call into doubt the existence of NSAP.

---

[124] Prinz (2011) claims that, "Cognitive desires may be accompanied by emotions. If I want it to be the case that my candidate wins, I will feel nervous anticipation, and the thought of victory will instill delight, while the thought of defeat will usher in waves of despair. On experiencing any of these fluctuating feelings, I may report that I desire a victory. *There is no one feeling of desire*, but rather a family of anticipatory emotions" (190) [emphasis added]. I take it that Prinz might claim that some of the emotions that compose this family can explain what I am calling attraction.

[125] I take it as obvious that we *can* experience the drag of desire without also experiencing dread, anxiety, or apprehension. This also seems possible with respect to delight. Consider an alcoholic who experiences the draw of having a drink but who does not experience anything close to delight at this thought. It is of course possible for a conservative to dig in her heels here and contend that there must be some set of sensory experiences the combination of which feels like what I am calling attraction. But such a move seems like a last resort that only one with die-hard conservative sympathies would find plausible.

I have offered, in this section, a defense of the NSAP of desires. That defense can be summarized as follows: (i) desires are the states that realize the causal role of desiring, and (ii) the states that realize the causal role of desiring possess non-sensory phenomenology; (iii) therefore, desires possess non-sensory phenomenology. Such a defense, of course, is only as plausible as the realizer functionalist position that supports it. And I have yet to offer reasons why we should think such a view is true. In the next section, I defend this version of functionalism. I do so by investigating what I take to be the most promising alternative position to this view—viz., role functionalism—and in turn, arguing that the main motivation for embracing it over realizer functionalism is not a compelling one. If such a defense is successful, then we are on good ground in taking premise (1a) of the Phenomenological Argument—the claim that desires possess (a non-sensory) phenomenology—to be true.

IV: Multiple Realizability, Two Versions of Functionalism, & the NSAP of Desire

The main reason a number of philosophers opt for role functionalism over realizer functionalism is because the former view, unlike the latter, is thought to be able to account for the multiple realizability of psychological kinds. To say that a psychological kind M is multiply realizable is to say that M can be instantiated in a number of physically diverse organisms. So, for example, the psychological kind *pain* is thought to be multiply realizable insofar as a wide-range of organisms, from humans to octopuses to newts, can be in pain. If one embraces role functionalism, it seems, one can make sense of the multiple realizability of this psychological kind. After all, if being in pain is being

in a second-order state of having a first-order state that plays the causal role of pain, then as long as an organism tokens a type of state that plays the causal role of pain, that organism can be thought to be in pain.

A realizer functionalist, on the other hand, must say, it seems, that the human and the octopus do not token the same psychological kind. This is because what it is to be in pain according to the realizer functionalist, is to be in a first-order state that plays the causal role of the mental state in question. Insofar as humans and octopuses are not in the same first-order state, these organisms cannot be thought to both be in pain. And it seems reasonable to think these two types of organisms won't be in the same first-order state given their physiological/neurological differences. Realizer functionalism it seems, cannot make sense of the multiple realizability of pain. It is this objection more than any other that has led a number of philosophers to reject realizer functionalism in favor of role functionalism.[126]

The first thing to note, however, about the multiple realizability objection vis-á-vis the NSAP liberal-conservative debate is that it targets only liberals with reductive physicalist sympathies. The liberal who embraces property dualism, for instance, can accept that the realizers of the causal role of desire are a heterogeneous lot.[127] This can include realizers that possess, e.g. attraction, as well as states that lack this type of phenomenology. The property dualist can allow that what realizes the causal role of

[126] The same objection applies, *mutatis mutandis*, to attitudes as opposed to sensations. Putnam (1967) was the first to defend a version of the multiple realizability objection. It has since become one of the most widely discussed arguments in philosophy of mind. For a survey of the literature on multiple realizability, see Funkhouser (2006) and Bickle (2013).

[127] Property dualism, while certainly not a popular view, has had its share of recent advocates including David Chalmers (1995), Brie Gertler (2008) and Richard Fumerton (2014).

desire might be a motley mix of states because she is unsympathetic with the mental state-brain state identity thesis that multiple realizability poses a challenge to.[128]

One might object, however, that there is still a problem for the dualist who defends the existence of NSAP in the way I have suggested—namely that she still needs to provide an account of what unifies the psychological kind in question.[129] The dualist, though, has a few options available to her with respect to this question. One option is to contend that phenomenology unifies the kind in question. On this view, what makes the set of realizers desires is that they all possess a particular type of phenomenal character—viz. attraction. Such a thesis however, is certainly bold.[130] An alternative response to the unification challenge is that what unifies the set of realizers is causal role. So while on this view, the set of all desires is a heterogeneous lot, what unifies the lot is the fact that all of them are states that produce action. A dualist can maintain that the causal role of desire is what plays this unifying role without committing herself to role functionalism about the attitudes.[131] If such a move is tenable, then the dualist has an alternative, non-phenomenologically based response she can offer to the unification question.

Such a response to the multiple realizability objection is not available to the reductive physicalist. And I certainly don't want the way I have defended NSAP to be held hostage to the truth of dualism. So how might a reductive physicalist with NSAP

---

[128] Such a dualist, of course, will have to hold that these immaterial states have causal efficacy.

[129] A number of philosophers have posed a similar question to reductive physicalists who embrace domain-specific or local reductions of psychological states. See Kornblith and Pereboom (1989). I will have more to say about domain-specific reductions below.

[130] Although I imagine some NSAP liberals who are property dualists might be sympathetic with this position.

[131] This is the case because such a dualist does not hold that desires are second-order properties of having some first-order property that plays the causal role of desire. Kim (1992), qua reductive physicalist, offers a similar response to the question of what unifies the heterogeneous lot of physical realizers of desire. I will have more to say about Kim's defense of reductive physicalism below.

sympathies respond to this objection? I suggest they do so in the same way that a number of reductive physicalists who lack NSAP sympathies have responded to it—viz., by defending domain-specific reductions of psychological kinds.[132] On such a view, psychological states are relativized to sets of organisms, viz., organisms with the appropriate neurological kinds. Lewis's (1980) species-relative reductivist program provides one model for how such reductions can be carried out. Here is an example: perhaps a certain neuronal kind realizes a particular type of desire in humans (e.g. D-excitation), while a different type of neuronal kind realizes desires in octopuses (e.g. OD-excitation). If that's the case, then qua realizer functionalists, we can posit D-excitation as a desire-that-Φ for humans, and OD-excitation as a desire-that-Φ for octopuses. Relativizing mental kinds to species might be problematic insofar as there can perhaps be a heterogeneous lot of intra-species neural states that play the causal role of a single psychological state. But if there is such a problem, as Horgan (1993) and others have held, it can potentially be handled by relativizing the latter to more specific sets of organisms.

The above response to multiple realizability concerns will likely not appease the committed role functionalist. The latter will object that it has the consequence that humans and octopuses cannot token the same psychological kind—e.g., desire—since desires for humans are of a different kind than desires for octopuses. I think, though, that this consequence role functionalists find so objectionable is not all that problematic when we investigate the matter more carefully. Indeed, insofar as the cognitive/neural

[132] Armstrong (1968; passim), Lewis (1980), Kim (1992), Polger (2002) and a host of others have defended reductive physicalism in such a manner.

architecture of these two species is in fact remarkably different, it seems reasonable to hold that human desire and octopus desire really are two different psychological kinds. The reason, I submit, that we might think this is not the case is because the causal role these states play for both species is the same. But when one takes into account the radically different nature of what is playing the causal role of these states, the above objection loses much of its force. In short, then, I deny that the type of multiple realizability role functionalists think we need to make sense of is a phenomenon that needs to be made sense of; or, more carefully, needs to be made sense of in the way that the role functionalist thinks it does.[133] So, there are, I think, reasonable responses the physicalist can make to multiple realizability concerns.

I have attempted to undercut the primary reason for being a role functionalist by arguing that multiple realizability concerns do not tip the scale in favor of role functionalism over realizer functionalism. My own view is that role functionalists would prefer to be realizer functionalists if not for multiple realizability issues. The reason is that there are advantages to being a realizer functionalist as opposed to being a role functionalist. For example, realizer functionalists appear to have a much easier time accounting for the causal efficacy of mental states. Kim (2006), McLaughlin (2006) and others have argued that the second-order properties role functionalists posit as mental states have no role to play in causal explanations. So, for instance, it has been argued that one can provide an entirely sufficient causal explanation of why an agent grimaces when she stubs her toe by simply positing the realizer state. There is no explanatory need to posit an additional second-order state because there is simply no causal work for this

[133] See Polger (2002) for a similar, more developed response to this worry.

second-order property to do. And it does seem reasonable to think that what is doing the causal work here is not the second-order state of having a property that plays the particular causal role of pain, but rather that which realizes this causal role. Exclusion arguments like these, if sound, would indicate that the role functionalist is committed to radical epiphenomenalism. Insofar as radical epiphenomenalism is false, this would constitute a *reductio* of role functionalism. Although there are responses in the literature to this objection, the realizer functionalist appears, at the very least, to be on firmer ground with respect to the causal efficacy of mental states than the role functionalist.

If the primary reason for being a role functionalist is not compelling, and there are advantages to embracing realizer functionalism over role functionalism, then that provides support for embracing the former over the latter. Given the virtues of being a realizer functionalist, and the fact that the alleged main vice of being a realizer functionalist can be successfully handled, I think we are on fairly firm ground in accepting realizer over role functionalism. But why think that we should either be realizer or role functionalists?

The answer to this question, I think, is that conceptually-speaking, attitudes, unlike sensations, are to be understood in terms of their functional roles, *not*, for instance, their phenomenology. Consider: if it were in fact the case that states that generate action lack phenomenology, it would still be reasonable to think of these states as desires.[134] This indicates, I think, that *desire* is a functional concept. And given this, I contend that the two best options we have concerning the ontological nature of these states are role

[134] I do not, however, think the same can be said about sensations. But I do not need to defend that claim here.

and realizer functionalism. Insofar as this is the case and it is also true that there are strong reasons to think that we should be realizer functionalists as opposed to role functionalists, we have good reason to think we should be realizer functionalists about the attitudes.

The above completes my defense of (1PA). I will next explain why if (1a) is true, not all desires are dispositional states; i.e., I will explain why (2PA) is true of the Phenomenological Argument is true.

V. Phenomenal States, Dispositional States, and Occurrency—a Defense of (2PA)

In order to explain why dispositional states are not phenomenal states I first need to take a detour into metaphysics and explain more carefully the distinction between occurrent mental states and dispositional mental states. Occurrent mental states, on the view I am sympathetic with, are mental events. These types of states, in other words, are mental occurrences. Tim Crane (2001), who is sympathetic with this way of understanding the distinction between occurrent states and dispositional states, buttresses this distinction by noting that mental occurrences have genuine duration whereas dispositional states do not. What Crane means is that occurrent states are mental occurrences that can typically be measured, more or less precisely, by the clock. Dispositional states, on the other hand, are not occurrences the duration of which can be measured more or less by the clock. Now if occurrent states are events proper and dispositional states are non-eventful long-standing states, then occurrent states are not dispositional states.[135]

[135] I know of no philosopher who takes occurrent mental states to be dispositional states.

Given this distinction between occurrent mental states and dispositional mental states, we are now in a position to see why (2PA) is true. This is because phenomenally conscious states, states that *feel* a particular way to the one tokening the state, are paradigmatic examples of mental events. They are mental episodes that have a noticeable, determinate duration. The experience of a sharp pain, for instance, is an episodic mental occurrence. The phenomenally conscious experience of hearing the sound of thunder is a mental event that lasts for a measurable amount of time. It is not surprising, then, that we find some philosophers simply defining occurrent states as states with phenomenology.[136] I have refrained from doing so. But I do think we have good reason to hold that every mental state with phenomenology is an occurrent state.

Support for premise (2a) is bolstered by the fact that no one in the literature currently defends the view that dispositional states have phenomenal character and for good reason. There is simply nothing it is like to instantiate a particular dispositional state. They are not experienced. They are not felt.[137] But if that is the case and dispositional states are not occurrent states, as argued for above, then this provides

[136] Thomas Senor (1993) suggests this when he claims that, "A belief at t is occurrent iff. the belief is conscious at t" (461). If the type of consciousness Sensor has in mind here is phenomenal consciousness, then at least with respect to beliefs Senor takes phenomenal consciousness to be both necessary and sufficient for occurrency.

[137] This appears to be a point of agreement among those participating in the current scope-of-phenomenal states-debate (i.e. the debate over what types of mental states have phenomenology.) In surveying this debate, Bayne and Montague (2011) note that one point of common ground among all parties to the debate is that "dispositional or unconscious states have no phenomenological character" (11). They go on to note that reference to the phenomenology of cognitive attitudes should be taken to refer to occurrent tokens of such states only.

further support for the view that if a token mental state has phenomenology, the state is an occurrent state.[138]

Now the manifestations of dispositional states might have a qualitative character. But this is not to say that dispositional states themselves possess phenomenology. Consider: an agent might dispositionally believe there is no solution to the Liar's Paradox. A manifestation of this dispositional belief might be the agent's judgment that the Liar's Paradox has no solution. Some have held that a judgment of this sort has a phenomenal character.[139] But even if the judgment possesses phenomenology, the dispositional belief from which it arose lacks phenomenology. It lacks phenomenology in the same way that the dispositional property of fragility lacks the property of sharpness that a manifestation of fragility—e.g. broken glass—possesses. I think it is reasonable to hold, then, that all dispositional states lack phenomenology, and that hence, (2PA) is true. And if that is the case, we have here an argument to the conclusion that not all desires are dispositional states. (1) of our puzzle is false.

So far in this chapter I worked to put to rest SDD by arguing that desires possess a non-sensory phenomenology. There is more to be said, however, concerning the nature of desire. In particular, in the next section, I defend more specific details about the metaphysics of this state, details that play a role in helping us explain how we have privileged access to these attitudes. I turn to that task next.

---

[138] This line of reasoning would not provide conclusive evidence for this claim since there might be other types of states that are neither occurrent nor dispositional that possess phenomenology. That being said, I don't think there are any non-occurrent states that possess phenomenology.

[139] Searle (1992; passim) and Silins (2013) are two philosophers who are sympathetic with this view.

Part VI: A Richer Account of the Nature of Desire

In this section I work to fill out my account of the nature of desire by explaining: (i) why attraction is the unique phenomenology of desiring, (ii) why all occurrent desires possess attraction, and (iii) why we should countenance dispositional desires in our ontology, as well as what the relationship is between occurrent desires and dispositional desires. I begin with (i) first.

I argued above that some desires possess non-sensory phenomenology. If I am correct about this, then such phenomenology is not identical to or reducible to the type of phenomenology possessed by sensations. It would follow, then, that the only other types of states that could possess attraction would be attitudinal states that are not desires. I think these latter states lack attraction as well. In defense of this view, it is not out of place, it seems, to make a phenomenological appeal. The type of phenomenology that other attitudes seem to possess does not feel like attraction. Consider: what occurrently judging that P feels like, for instance, is different than what occurrently desiring that P feels like. And it certainly seems like the phenomenology of judging a proposition is nothing like attraction.[140]

A more plausible case could be made for there being pro-attitudes in addition to desire that possess attraction. However, one strong reason for thinking that, e.g. intending has the phenomenology of attraction is that such states are the conjunction of two different intentional states—viz. a particular belief (e.g. the belief that one will do the

[140] One might think that judging that P just is being attracted to the truth of P. But this isn't the case. One can be attracted to the truth of a proposition, and yet not judge that the proposition is true.

thing in question), and a particular desire to do that thing.[141] But if this is the case, then the reason intentions have the phenomenology of attraction is because desires have the phenomenology of attraction. Such a view would not call into doubt the claim that desires have a unique phenomenology.

The above discussion indicates that a proponent of the view that pro-attitudes in addition to desire possess the phenomenology of attraction would need to find a case of a pro-attitude that is: (a) not the conjunction of a desire and another belief, and (b) possesses the phenomenology of attraction. I am compelled to think, on phenomenological grounds alone, that such attitudes do not exist.

If one wants an argument for the above claim, though, we can offer a variant of the types of partial-zombie cases offered by Horgan (2011) and Kriegel (2015), discussed above. We can, that is, construct scenarios in which a being lacks any sensory phenomenology *and* lacks the experience of attraction, but nevertheless possesses pro-attitudes such as hopes and intentions. Such creatures, it seems reasonable to think, would still have a phenomenological life. There would still be something it was like for this creature to hope that P were the case, or intend to do P. If this is the case, the phenomenology in question couldn't be a matter of their tokening states that possessed attraction. Such partial-zombie cases, it might be thought, give us a defeasible reason for thinking that hopes and other pro-attitudes possess phenomenology that is not reducible to attraction and sensory phenomenology.

[141] As I will discuss in more detail in chapter VI, some philosophers (e.g. Searle, (1983)) appear to think that all attitudes that are not beliefs or desires are either a particular type of belief or desire or a conjunction of these two states. Schroeder (2004) offers a nice discussion of the view that pro-attitudes such as hopes and intentions are a conjunction of a desire and a belief. He rejects this view.

Admittedly, it might be thought that such partial-zombie scenarios are not conceivable. One reason for thinking this, though, is that hopes, intentions and other pro-attitudes really are partially constituted by desires. So one simply couldn't token an occurrent intention without also tokening an occurrent desire, a desire that on my view possesses attraction. But, of course, if that is one's reason for denying the conceivability of such creatures this does not effect my claim that desires possess a proprietary phenomenology, the type of phenomenology that is unique to desiring.[142] The burden, then, is on one who thinks that pro-attitudes such as hopes and intentions are not reducible to desires, and yet nevertheless possess attraction.

I not only hold that all occurrent desires possess the *proprietary* phenomenology of attraction; I also hold that all occurrent desires possess attraction. I am sympathetic with this view because I believe that all occurrent desires are occurrent desires in virtue of playing the causal role of generating action. And I am not sure what generates action if it is not the felt quality of attraction.

In order to see the force of this reasoning, consider the view that some occurrent desires lack phenomenology but are nevertheless occurrent desires in virtue of being state/event conscious states (i.e. in virtue of our being aware of such states). Insofar as desires really do perform the causal role of generating action, and such states lack the phenomenology of attraction, what is it that enables these states to play the causal role of desiring? I am not entirely sure there is a good answer to this question. On my view of

[142] One might think that even if desires are the only states that possess attraction, it might be possible for some desires to possess a phenomenology that is not attraction. We would need, though, a principled reason for thinking that such desires exist, especially given that I have argued that desires are intimately connected up with action and attraction is the felt quality that generates action. I am not sure what those reasons would be.

occurrent desires, we have a good answer to the question: "how do desires generate action?" The answer is: via attraction. Those, then, who want to deny that all occurrent desires possess attraction, and yet want to maintain that the casual role of desire is action-based, owe us an explanation of how desires realize such a role. I am skeptical than a reasonable explanation is forthcoming. [143]

I have defended the view that *all occurrent desires* are states with the unique phenomenology of attraction. But I do not think that *all desires* are states with attraction. This is because I think there are dispositional desires, and these dispositional desires, as argued for above, lack phenomenology. I think dispositional desires exist because it strikes me as reasonable to suppose that, for example, a mother desires that her children are healthy even when she is not occurrently experiencing an attraction to that content. It seems reasonable to suppose, that is, that we have desires that exist despite the fact that such states are not occurrently being experienced. But if we want to countenance dispositional desires in our ontology, we need an explanation of what makes these states desires.

One explanation for what makes these states desires follows Searle's (1998), Kriegel's (2011), and Smithies' (2013) explanation for what makes dispositional states in general the types of states they are. Broadly speaking, these philosophers hold that these standing states are the types of states they are because of their relationship to states that possess phenomenology. What makes standing desires, desires, on this approach is that

[143] Note, though, that while I hold that all occurrent desires possess the proprietary phenomenology of attraction, I don't embrace the view that we have privileged access to or even knowledge of all our occurrent desires. From the mere fact that all occurrent desires possess a unique phenomenology, absolutely nothing follows about our access to such states. As will be discussed in greater detail in chapter V, I think it is possible to token occurrent desires without knowing that we do.

the essential manifestations of these types of states are occurrent desires that possess attraction. In other words, because the essential manifestations of dispositional desires are states that possess the proprietary phenomenology of desiring, these states are desires as well. Dispositional desires, on this view, are phenomenologically-individuated in a derivative way. So while there are desires that lack phenomenology, they are desires because their essential manifestations are token states that possess the unique phenomenology of desiring.

This phenomenologically-based explanation, however, is not one I want to embrace. This is because, as I argued above, we should type-individuate occurrent desires in terms of their casual role, not their phenomenology. If I am correct about this, we need an explanation in terms of causal role that explains why dispositional desires are the states they are.

One proposal is that dispositional desires realize the same causal role as occurrent desires. But, I am not entirely sympathetic with this suggestion. This is because I am not convinced that dispositional desires do generate action. Above, I suggested that what "gets the limbs in motion" or compels us to act is the phenomenology of attraction. But dispositional desires, as defended in section VI above, lack phenomenology. So insofar as dispositional desires generate action, we need a non-phenomenological explanation for how this occurs. And I am not entirely sure what that explanation would be.[144]

An alternative explanation of what makes dispositional desires the type of attitudes they are, one I am more sympathetic with, involves taking a leaf from the Searle

[144] This is one reason why I think action-based, strict dispositionalists about desire offer an explanatorily deficient account of desire.

(1992), Kriegel (2011), et al. view discussed above. My thought is that we can hold that dispositional desires are desires in virtue of the fact that their essential manifestations are states that realize the causal role of desiring. On this view, dispositional desires are desires in what we might call a causally-derivative way; they are desires in virtue of their intimate relationship to desires that generate action. If one embraces this view, desires strictly speaking would be states that play the causal role of desiring *or* states the essential manifestation of which play the causal role of desiring. This proposal is, admittedly, in need of further development. But I think it might very well be the best way to countenance the existence of standing desires.

Certainly more could be said about the metaphysics of desire than I have said here. However, the basic outline of the view of desire I embrace is hopefully, at this point, clearer.

Conclusion:

In this section I have argued that SDD is false because some desires possess attraction. I in turn defended the view that all occurrent desires are states that possess attraction and dispositional desires are desires because of their relationship to these occurrent desires. It is left to discuss, given this theory of desire, whether we possess privileged access to these states. I argue in the next chapter that we do possess privileged access to some of our desires, a defense that relies on the fact that some desires possess the proprietary phenomenology of attraction.

Chapter 5: The Epistemology of Desire Part 1

In the previous chapter I argued that some desires possess a unique phenomenology that only desires possess—viz. the phenomenology of attraction. I argued that because some desires possess attraction, then, insofar as dispositional states lack phenomenology, not all desires are dispositional states. We were thus able to reject proposition (1) of our puzzle. But the falsity of (1) does not entail that (2) of our puzzle is true; in other words, the fact that not all desires are dispositional states does not entail that we actually possess

privileged access to some of our desires. The focus of this chapter is on whether we do possess privileged access to some of our desires.

We have already seen that the following theories of self-knowledge—(a) inner sense, (b) transparency (c) agential rationalism, and (d) constitutivism—not only face daunting challenges when it comes to explaining how we possess privileged access to some of our desires; they also have problems as accounts of self-knowledge in general. Given this, it is wise, I think, to consider an alterative view of self-knowledge. The view I have in mind shares affinities with accounts I referred to in chapter 2 as "acquaintance accounts of self-knowledge." As discussed in chapter 2, those who embrace the acquaintance approach think we have direct (i.e. non-inferential), introspective access to some of our experiences, access that is not the result of a causal process. Furthermore such acquaintance theorists hold that our knowledge (or justified belief) of these experiences is not grounded in the reliability of the means by which we know these states. Nor is it grounded in the rational agency we might have with respect to our attitudes. It is rather grounded at least in part in introspective evidence. Acquaintance theorists, then, offer an alternative, competing view to (a)-(d) above.[145]

Why, though, should we consider embracing an acquaintance approach of self-knowledge with respect to attitudes? The short answer to this question is the one implicitly offered in chapter 2: the access we have to some of our attitudes, and in particular our occurrent attitudes, seems direct in a way that our access to other events are

[145] As noted in chapter 3, perhaps the most plausible way of understanding Shoemaker's constitutivist view is by coupling it with a reliabilist epistemology. Shoemaker's view also relies on a claim about the rational capacity of agents that acquaintance theorists do not rely on (although acquaintance theorists are going to claim that at least some degree of rationality is required in order for agents to possess knowledge by acquaintance.)

not. Consider: my awareness of a my neighbor waving at me seems to be metaphysically indirect. I know (if I know at all) about this occurrence via my visual experience of my neighbor. But my occurrent desire, a mental event, for instance, to wave back seems to be unmediated in that I don't need to be aware of anything else in order to be aware of that desire; nothing, in other words, stands between the event and my awareness of it. It is this direct relationship of one's mental events that, according to the acquaintance theorist, helps ground the type of privileged access we think we have to at least some of our occurrent mental states.

In this chapter I am going to develop a version of this approach with respect to desire, an approach I believe has the resources to explain how (2) above is true. My explanation for how we possess privileged access to some of our desires involves a process I call *phenomenal simulation* (PS). Phenomenal simulation involves entertaining a given mental representation and attending to the fact that we have a particular experience to it. PS, at least in some cases, involves the creation of a desire, a process of creation that puts us in a prime position to have privileged access to these states. Such a means, though, must be carried out properly in order for it to yield this type of knowledge. I offer a detailed version of just how this process can afford us with privileged access below.

Before proceeding further, I want to stress that I am not interested in defending the view that we have certain or infallible knowledge of our minds. I am only interested in defending the view that we have uniquely first-personal, *highly epistemically secure* knowledge of our minds, where the latter does not require that we possess justification so

strong that it precludes the possibility of error.[146] That is just to say that I am only interested in defending the view that we possess privileged access to some of our desires, not that we possess certain knowledge of these states. Furthermore, while I am sympathetic with the view that we can possess direct (i.e. non-inferential) knowledge of facts about our minds, I am not interested in defending the view that we possess non-inferential knowledge of such facts. In other words, I am not interested in defending a version of foundationalism concerning the contents of our minds. Of course, if we possess only inferential knowledge of our minds, this might lead some to think that the epistemic security of such knowledge will fall short of privileged access. In some cases that will be the case. But not, I think, in all cases.[147]

I. An Introduction to Phenomenal Simulation

I want to begin developing my account of the epistemology of desire by explaining a process that, while not necessary for the possession of such knowledge, is nevertheless a means that, if conformed to appropriately, puts agents in a prime position to determine what they want. I call the process in question *phenomenal simulation.*

Phenomenal simulation occurs when an agent entertains a given mental representation[148] and focuses on whether she is attracted to it. Undertaking this process, in the case in which one is attracted to a particular content, is, on my view, just to create

[146] Bonjour (2010) refers to such reasons or evidence as "conclusive reasons."
[147] Consider: the issue with Byrne's account of self-knowledge was not so much that it involved inference, but that the inferences weren't reliable enough to afford us with privileged access.
[148] I will also refer to these types of mental representations as "content," although this might be slightly misleading.

an occurrent desire.[149] Attention to the process of creating this desire puts us in a prime position to access the output of this process—the desire itself. An example will help illuminate this means: an agent might reflect on whether they want to order an eggplant soufflé. In doing so they might entertain the mental representation—*I order an eggplant soufflé*—and come to experience an attraction to that representation. To be attracted to ordering an eggplant soufflé just is on my view what it is to desire that one order an eggplant soufflé. The process of phenomenal simulation, then, enables an agent to create a desire (or attend to the fact that no such desire is created).[150] And the proposal is that such a process puts one in a prime position to possess privileged access to the fact that they have or fail to have a particular desire.

I call this process "phenomenal *simulation*" because it involves a cognitive, first-personal exercise in which one becomes aware of a phenomenal experience they have[151] to a content that represents what might or might not obtain. But it is important to note that the awareness I think one who engages in phenomenal simulation has is not awareness of how one would behave, react, or feel in a given situation. That is to say, the awareness is

[149] I speak here of "creating a desire" because occurrent desires, as discussed in chapter 4, are momentary events. The same *token* occurrent desire cannot be instantiated multiple times, although the same *type* of occurrent desire can be. When an occurrent desire ceases to exist that is the terminus of that desire. So the process of phenomenal simulation really is the process of, at least in some cases, creating a new occurrent desire.

[150] I think we sometimes experience a *felt revulsion* to a particular content. And I am sympathetic with the view that we sometimes become aware of the fact that we lack a desire by attending to an experience of revulsion we have towards a particular mental representation. While I will not explore this way of coming to know that we *don't* desire a particular object, I think such a suggestion deserves further development than I can give it here.

[151] Or aware of the fact that they lack a particular experience.

not of a complex counterfactual. What one becomes aware of, on my view, is rather an attraction to an object; it is awareness of a desire.[152]

Phenomenal simulation, shares some similarity with the first-person simulation method cognitive psychologists such as Wilson and Dunn (1999) and Schultheiss and Burstein (1999) and philosophers such as Gordon (1986) and Goldman (1989) are interested in. The process they focus on involves a subject imagining a future scenario and then attending to how that scenario would make them feel, in order to determine whether they have an unconscious motive or a dispositional attitude. Such an exercise, according to Wilson and Dunn, can provide the simulator with good evidence that she has a particular dispositional or unconscious attitude. While such a process is similar to phenomenal simulation, if I am correct, the result of undertaking the latter plays a crucial role in providing us with a more direct awareness of our phenomenal experience to a particular content, awareness, that is, of an attitude.

Phenomenal simulation also shares affinities with Moran's (2001) agentialist version of transparency. As previously discussed, Moran thinks we come to know our attitudes via a process of authorship, a process that involves actively reflecting on/deliberating about what we ought to believe, desire, etc., and ends with a commitment to an attitude in the form of an avowal. On my view, a person puts herself in position to be *aware* of her attitudes, and in particular her desires, by undertaking the cognitive process of creating a desire, a process that it involves actively creating (or not creating) a desire. But it does not involve the type of authorship of an attitude Moran thinks is necessary for critical self-knowledge. This is because phenomenal simulation does not

[152] A more cautious way of putting this point, is that it is awareness of that which is a desire.

require the type of deliberation about what we ought to desire that Moran's account relies on.

The above is not to say, however, that the type of deliberation Moran thinks is central to self-knowledge plays no role at all in my epistemology of desire. On the contrary, when we have before their mind a given content we might well consider the value of having that content actualized. We might consider the reasons for or against desiring that particular object. We might, in short, consider what we ought to desire. Such deliberation could in turn, generate an attraction to the content in question. Consider the following example: I might be puzzled about whether I actually desire to go for a run. After considering the options for and against going, I might determine that I have good reason to go running. Such a conclusion might very well lead me to be attracted to the state of affairs in which I run. Then again, it might not. But the latter is not a mark against phenomenal simulation. As I defended in chapter 3, we (at least) occasionally don't desire that which we ought to desire.

The point of the above discussion, then, is that deliberation can have an impact on whether we experience an attraction to a particular content, and in some cases such deliberation will lead us to experience an attraction to that which it would be prudent for us to desire. Such agency does not, however, provide the epistemic basis for how we know what we want. If it did, creating a desire via phenomenal simulation would entail that we know or are justified in believing that the product we create exists. This, however, is not my view.[153] Given, then, that creating a desire via phenomenal simulation

[153] If it were, my position would be a defense of the view that we can possess so-called "maker's knowledge" of the contents of our minds. On one understanding of maker's knowledge, the mere fact that

does not entail that we know that which we create, difficult questions remain concerning how we acquire privileged access on the basis of it. In particular one might wonder how we become aware of the output of such states in a way that insures our beliefs about them will be epistemically secure. One might also wonder how we can conceptualize the output of PS in such a way that insures that we classify such states accurately. In the next section, I address these questions.

III. Phenomenal Simulation and Privileged Access

We can begin to to see how engaging in PS can afford us with privileged access by considering how we can come to be aware of the output of PS in such a way that insures our beliefs about the latter will be epistemically secure. Such details need to be addressed since it is commonly held that (a) phenomenal reality can outstrip our introspective, discriminative capacities, and (b) the way a mental property appears to us can incline us to believe false things about it.[154] On the view I am sympathetic with we can, in part, avoid these problems by insuring that the way attraction and the mental representation *epistemically appears* to us is the way these two components of the state in question actually are. The way mental properties epistemically appear to us is a matter of what such properties incline us to believe.[155] If the way a state epistemically appears to us is

---

we make/create x, provides us with epistemically secure knowledge of x. Such a view would certainly make my position simpler. But I can't figure out how the mere fact that one makes something provides one with knowledge of that thing. For a discussion of maker's knowledge see Graukroger (1986).

[154] I take the infamous problem of the speckled hen discussed by Chisholm (1942) among a host of others to indicate just this. In brief, the problem involves the fact that we can have an experience of e.g. a 48-speckled hen, and not be able to determine via introspection alone that we are being appeared to in such a manner. Fantl and Howell (2003) take the upshot of the problem of the speckled hen to be that the way a phenomenal property appears to us can "outstrip" our introspective capacities.

[155] Cf. Gertler (2010, 2012).

the way it is, then our awareness of such an appearance will be an awareness of the phenomenal reality.

We can, I think, insure that the way a state epistemically appears to us is the way it actually is by carefully attending to the state in question, and adopting what Gertler (2011) calls an attitude of "scrupulous caution." One adopts this attitude when one is especially concerned to avoid error with respect to what is presented to her in experience. The thought is that by lowering our degree of credulity towards what is presented in experience, we can refrain from reading more into the experience than is actually there. An example will help to clarify this idea: if one is attending to an experience of a Dalmatian, and one is not adopting a scrupulously cautious attitude, one might be inclined to think that one is presented with a 101-spotted experience when one is tokening a 105-spotted experience. By adopting a scrupulously cautious attitude we can avoid being inclined to think we are tokening a 105-spotted experience. In fact if we are being especially cautious, we will not be inclined to think we are tokening an above-75-spotted experience. One especially concerned with avoiding error will only be inclined to think they are being presented with a many-spotted experience, or perhaps just a spotted experience.

How an experience epistemically appears to us will, as Gertler points out, be effected by more than just our level of credulity. Such appearances will in part be impacted by our background beliefs, expectations, epistemic perspectives, and our powers of discrimination. Gertler (2012) suggests, following Horgan and Kriegel (2007), that adopting an attitude of scrupulous caution also entails 'bracketing out" or suspending

presuppositions we might have about any relational information about our experiences, including how they are classified by other subjects, what their typical causes are, etc.[156] Bracketing out these presuppositions enables us to focus on how the experience appears to us at the moment we are tokening it. Such bracketing helps explain how we can insure that our epistemic appearances line up with phenomenal reality.[157]

If by adopting an attitude of scrupulous caution we are able to insure that an epistemic appearance is constituted by the reality of the mental properties in question, then how such properties epistemically appear to us will be how they truly are.[158] We will then be aware of the output of phenomenal simulation as it truly is. Such awareness, however, on my view is not a belief.[159] And in order to possess knowledge* of our attitudes via phenomenal simulation, we need to form a belief about the output of this process, a belief that involves accurately conceptualizing this output.

My proposal for how we move from awareness of the output to an epistemically secure belief about the latter involves the claim that some concepts are partially constituted by the experience the phenomenal concept refers to. Such concepts are formed, according to Chalmers (2010), by singling out the experience in thought and

---

[156] See Horgan and Kriegel (2007; p. 128).

[157] An example might further help to explain what this bracketing involves: in focusing on the experience of a sensation of pain, I might be inclined to believe that this experience is caused by a rotting tooth. This is a presupposition about the cause of the pain, and it can mislead me into attributing a relational property to the sensation that it might not have. The type of suspension of presuppositions that Gertler and Horgan and Kriegel have in mind eliminates this type of misattribution.

[158] A slightly different way of putting this point is that such phenomenal properties will appear to us in a way that we will only be inclined to believe true things about them.

[159] The view that the type of awareness I am referencing is not a belief comports with how other philosophers who embrace the view that we can possess introspective, privileged access of our sensations and attitudes have understood such awareness. See Bonjour (2003), Chalmers (2003), and Gertler (2010) for this view.

"taking up" the experience into the concept.[160] Unlike other concepts, direct phenomenal concepts exist only as long as the experience that partially constitutes them exists.

Direct phenomenal concepts can be further explained by contrasting these concepts with relational concepts. The latter are expressed by ordinary language terms, and have their reference fixed by their relation to things in the world. They also refer in virtue of a description. For example, the relational concept "GREEN" can be expressed by the ordinary language term "green," has its reference fixed by its relation to green objects, and picks out its reference via a description such as "a color with such and such a character." Relational concepts are also standing concepts in the sense that they typically persist over a lengthy period of time even when that which they refer to doesn't exist. Direct phenomenal concepts, on the other hand, refer to mental properties not by means of a description, but rather via a demonstrative act, or demonstration. For instance, one might engender such a concept by attending to the experience and thinking the experience is like *this*, where *this* is not a description but rather a demonstration. So direct phenomenal concepts do not refer in virtue of a description, but rather refer via demonstration. It is in part because of this that such concepts cannot help but refer to that which they are about. Relational concepts, however, can fail to refer because there might not be anything that matches the descriptive component of such concepts.[161]

---

[160] Chalmers (2010) describes the formation of direct phenomenal concepts this way: "The clearest cases of direct phenomenal concepts arise when a subject attends to the quality of an experience and forms a concept wholly based on the attention to the quality, "taking up the quality into the sensation" (267).

[161] It bears mentioning here as well that on Chalmers et al. view, direct phenomenal concepts can only be formed by having a substantive grasp of the property in question. See Chalmers (2003, part 2), Gertler (2011) and Duncan (forthcoming) for further discussion of direct phenomenal concepts and the metaphysical and epistemological work they can be put to.

The existence of what Chalmers refers to as "direct phenomenal concepts", or something very much like them, has been defended by both physicalists and dualists alike.[162] These concepts once formed are deployable in beliefs about the experience in question such as, "I am experiencing *this*,"[163] where *this* is constituted by a direct phenomenal concept. Such beliefs, what Chalmers (2003) calls "direct phenomenal beliefs," insures that our conceptualization of the experience will be accurate. This is because if the very concept that refers to an experience is constituted by that very experience, then such concepts will necessarily refer to the experience in question.

Direct phenomenal concepts and the beliefs that are partially constituted by them enable us to understand how we are able to conceptualize the experience of the output of the PS process in reliable manner. Indeed, in the case in which the output of the PS process is an experience of attraction to a particular content, one who grasps the attraction by adopting an attitude of scrupulous caution, can form a direct phenomenal concept of this attraction, a concept that will partially constitute the phenomenal reality in question. They can then employ that concept in the belief that: "I am experiencing *this* to x", where *this* refers to the attraction and x is a placeholder for whatever the content in question is. As long as one has in fact grasped and conceptualized x appropriately,[164]

---

[162] Physicalists who think that there are phenomenal concepts that are partially constituted by the phenomenal experience they are about include Loar (1990) and Balog (2012). Dualists who embrace this view include Gertler (2001) and Chalmers (2003).

[163] Or even more cautiously, "*this* is being instantiated."

[164] Admittedly, up to this point, I have not said much about how we grasp and conceptualize the mental representation that constitutes the content of the belief. On one, admittedly, controversial view, we grasp our content via an intrinsic, essential phenomenal character they possess. If such a such a view is correct, we might be able to grasp our content via such phenomenology. See Siewert (1998) for a defense of this view. See Kriegel (2013) as well for an overview of the literature on so-called "phenomenal intentionality."

Alternatively, if content does not possess a unique phenomenology, there might be other ways in which we are able to grasp such content that insures we have epistemically secure knowledge of it. One

such a belief will necessarily be true. This is because the demonstrative is constituted by the very experience one is referring to.

The mere fact, though, that these types of beliefs are highly reliable does not entail, on my view, that we have privileged access of the proposition in question. This is because I have already rejected this externalistic account of knowledge in chapter 2. The knowledge in question on my view is not grounded in truth-conduciveness, but is rather based on mental evidence that is in principle accessible. On the view I embrace, such privileged access is possessed by not only forming judgments that are partially constituted by the experience itself, but by attending to the fact that one's judgment that one is experiencing attraction to x *corresponds* with one's awareness of the output of the phenomenal simulation process. The awareness of the output of such a process, when one adopts the attitude of scrupulous caution, will be the very properties that make the belief in question true. On the view being proposed, then, when one is aware of both the truthmaker for one's belief, the belief itself, and an awareness of the correspondence between the two, one has privileged access. After all, what more could one want epistemically then to have awareness of the experience and the content one is having that experience to (call this C), awareness of the second-order judgment that we have C, and an awareness of the fact that our judgment about C corresponds with C? Such awareness affords us with a highly justified belief that we are experiencing attraction to a particular content because we have: (i) the very truthmaker that makes our second-order judgment

---

such way is to embrace an "inclusion" account of the epistemology of content embraced by Burge (1987) and Gertler (2000) among others. On this view, our judgments about our thoughts *inherit* the content of the first-order thoughts one is making a judgment about. The inclusion account *might* explain how our judgments about our propositional attitudes insure a grasp of what the content of our first-order attitudes are. I will return to the issue of the epistemology of content below.

true before consciousness, (ii) awareness of the judgment that that truthmaker corresponds to, and (iii) awareness of the fact that the truth maker corresponds with such a judgment. We have, in other words, everything that constitutes a judgment or belief being true at the fore-of-consciousness. This is as secure, it would seem, as justification, and I would contend, knowledge, gets.[165]

The process just described would also be uniquely first-personal. This is because phenomenal simulation is itself a uniquely first-personal process *of awareness* in the sense that only I can come to be aware of the experience I have to a particular content via such a process. Third-personal parties can become aware of *their* experiences of attraction via such a process, but they certainly wouldn't be able to become aware that *I* have this experience via such a process. Second, only I can form a judgment about that which I am aware of via phenomenal simulation, and attend to the fact that what I am aware of corresponds with my judgment. So the epistemically secure knowledge achieved by this process is doubly first-personal.

At this point one might object that even if the above account is correct, and we are able to possess privileged access to the fact that we are experiencing *this* to x," this does not entail that one has privileged access to the fact that they *desire* x. This is the case, it might be thought, because one can possess privileged access to an experience towards a content, and yet lack awareness of the fact that such an experience is the experience of, e.g. desiring. And if this is the case, we wouldn't have privileged access to the fact that we desire/want a particular thing.

---

[165] Cf. Fumerton (1995): "When everything that is constitutive of a thought's being true is immediately before consciousness, there is nothing more that one could want or need to justify a belief" (75).

In response to this concern, I think we can have privileged access to the fact that we *want* a particular thing. In defense of this claim, consider that most of us have an understanding of what Chang and I call attraction as the feeling of desire.[166] We come to acquire this understanding, I submit, by experiencing attraction early on in our development. We are not only aware of some of these token instances of attraction, but we are able to form phenomenal concepts about these experiences.[167] We are also aware of the fact that some of these experiences are followed by our acting to obtain what we are attracted to. In this way, we come to recognize that such experiences are experiences that lead to action. Perhaps being aware of the attraction and then acting is enough to afford us with the concept of wanting or desiring. This is difficult to determine. And I won't take a stand on exactly how we acquire the concept of desire, or whether it is acquired at all.[168] But I do think it's true that we view wanting/desiring to be, at the very least, intimately connected with action.[169] Given the tight connection between the following—attraction and action, and desiring and action—it is not surprising, that we would be able to recognize that attraction is the experience of wanting

If one needs further defense of the claim that we do come to associate attraction with desiring, consider a child experiencing a strong attraction to a bottle of milk. It is not unreasonable to think such a child's experience (if she has the concept of wanting), might

---

[166] Another way of putting this is that our conceptual understanding of attraction just is an understanding of the latter as the experience of desiring, the experience of being moved to act.
[167] In some cases these concepts will be direct phenomenal concepts.
[168] The empiricism/anti-nativism debate continues to be waged. I am not sure whether there are innate concepts, but if there are I think that WANTING might very well be one of them.
[169] The view that desire should be understood along action-based lines is the standard view in philosophy. Even those who deny the standard view, recognize that there is a strong connection between desiring and action. I should also mention that many non-philosophers I discuss desire/wanting with do not bat an eye when I say that desire is intimately connected up with action.

be followed by the thought, "I want that milk" and if possible, an action to obtain that which is wanted, action the agent would realize was intimately connected up with her desire. If such instances occur enough times for this person, she is going to recognize, I believe, that the phenomenology of attraction just is the phenomenology of wanting. I take it that all cognitively mature agents have not only experienced attraction to a given "thing" followed by the thought or expression that they want that "thing," but that such events have occurred enough times that when persons become attracted to a content, they can and do recognize that experience as the phenomenology of wanting. In other words, we see attraction as a feeling that generates action.

Insofar as we view attraction as the experience of desiring, it is plausible to think that we would be able to form the judgment that the experience we grasp is the experience of desiring. On that basis, we could then come to possess privileged access to the fact that we desire the thing in question by applying the ordinary language concept of desire to our experience. This knowledge would not, it seems, be as epistemically secure as knowledge of the proposition that "I am experiencing *this* to x." But given that we are applying the ordinary language concept of wanting to the experience, this should not be surprising.

The fact, though, that such knowledge is epistemically grounded in our privileged access to the fact that we are experiencing *this* to x, and that we have a secure grasp of attraction as the experience of desire, we would, it seems, possess highly epistemically secure knowledge that we desire x. And it bears mentioning that such knowledge would

also be formed via a process that is uniquely-first person. No one else would be able to have access to this fact in this manner.

I submit, then, that it is plausible to hold we can and in some cases do possess privileged access to the fact that we have a particular desire. Such access would be derived from the privileged access we have of propositions that are partially constituted by direct phenomenal concepts.[170]

Conclusion: PSA & Other Accounts of the Epistemology of Attitudes

In the above section, I defended the view that we come to possess privileged access to some of our desires on the basis of the unique phenomenology that desires possess. The general thought that we become aware of desires via their phenomenology is highly plausible given the view of desires I defended in the last chapter. For if desires really do possess a proprietary felt quality, it would be odd, to say the least, if such phenomenology didn't play a role in helping us to determine what we want and don't want to do.

I would like to conclude this chapter by focusing on a crucial difference between my account of the epistemology of desire, and all other extant accounts of self-knowledge I have discussed in this work. The difference is that the approach offered above is a phenomenologically-based explanation of how we know our desires, while all other extant accounts are what we might call "non-phenomenological views." Indeed, all the theories of self-knowledge surveyed in chapter 2 and 3, with the exception of the type of acquaintance/direct awareness account defended here, are non-phenomenological in

[170] See Chalmers (2003; section 4) and Duncan (forthcoming) for a similar view concerning how we possess knowledge involving ordinary language concepts with respect to our sensations.

nature as well.[171] To take but one example, Moran's (2001) view of how we come to know our desires relies on the contention that we look outward to facts about the world to know what we want. Such a view obviously does not rely on the (alleged) felt quality I have argued desires have in explaining our knowledge of this state. After all, if one thing is uncontroversial, it is that phenomenology is an internal not external feature of the world. Given that it would be odd if desires possessed a unique phenomenology, and such a fact did not play a crucial epistemic role in our coming to know these states, the view of desire I defended in the last chapter, challenges the plausibility of all non-phenomenological accounts of desire.

I want to be clear here: in making the above point, I am not attempting to offer an additional argument for my epistemology of desire. Most of the philosophers who embrace the accounts of self-knowledge discussed in chapters 2 and 3 deny that desires have a unique phenomenology. If they did embrace the metaphysical account of desire defended earlier, I imagine at least some of them would be sympathetic with a view like mine. This re-emphasizes the main point of the puzzle of desire we began with, a point that has been a guiding theme throughout this project—viz. that one's metaphysical views on desire are going to have a profound impact on one's views about how we know our desires. In order to get clear on what the correct view of not just the epistemology of desire is, but of any attitude, I have argued, we must first do the difficult metaphysical

[171] And it bears mention that there is one acquaintance account in the literature that does not invoke phenomenology in an explanation of how we know our attitudes—viz. Bonjour's epistemology of attitudes. Bonjour, as discussed in chapter 2, claims that we have highly justified, non-inferential beliefs about our attitudes on the basis of a built-in-constitutive awareness such attitudes possess, not on the basis of any phenomenology such states might possess. I have refrained from discussing Bonjour's view in this chapter, but I have criticized it elsewhere.

work to figure out what the nature of these states are. In the previous chapter, I worked to do that, and in this chapter I have defended what I take the epistemic consequences of such work to be.

There are, however, a number of objections one can raise to the above account of self-knowledge. In the next chapter I address five objections to this view. Addressing these objections will help further explain and motivate the approach to self-knowledge I have offered here.

Chapter 6: The Epistemology of Desire Part 2

In chapter 5 I offered an account of the epistemology of desire, an account I will refer to here, for ease of reference, as the *Phenomenal Simulation Account* (PSA). PSA, I proposed, has the resources to explain how (2) of our puzzle is true. There are, however, a number of formidable objections that can be raised to this account. In this chapter, I anticipate and respond to five of them. My response to these objections will enable us to have a better grasp of what this view is and isn't committed to. I turn now to the first of the five objections in question.

Objection 1: PSA involves a process that is extremely cognitively tasking for any agent to carry out. In fact, one might wonder whether such a process can be carried out. It seems reasonable to think that we have privileged access to our desires without engaging in such a lengthy, demanding process. Normal people don't need to engage in such a

cognitively demanding process in order to have privileged access to their desires. Insofar as PSA is such a demanding process, we should reject it. Call this: *The Implausibility Objection*.

Response: My response to the Implausibility Objection is two-fold. First, one who offers such an objection is implying that: (i) it is obvious that we have privileged access to our desires *and* that such access is not rare, so (ii) privileged access should be easier to come by than it is. In response to this suggestion, I hold that such access might very well be rare with respect to some agents. Some agents simply might not be reflective or cognitively sophisticated enough to possess privileged access to their desires.

That being said I don't think such knowledge is so incredibly difficult to possess that only the most reflective, cognitively sophisticated, and careful people are able to acquire it. I think such knowledge can be, and in some cases, is achieved by most mature agents. It is not unreasonable to think that when such agents reflect on what they want, they would engage in a process very much like phenomenal simulation. If such agents do reflect on whether they want something in a way like phenomenal simulation, what would need to be the case is that they carried out such a process in a careful, scrupulously cautious manner. I see no reason for thinking mature agents can't do this.

Furthermore, I want to stress that I have focused my attention in this project on a very special type of knowledge. Privileged access is not your everyday, run-of-the-mill type of knowledge. It is a highly epistemically secure, uniquely first-personal standing. And on my view, while it is not implausible to think that agents can, and in some cases

do, possess privileged access to some of their desires in the way outlined above, it bears keeping in mind that we have various ways in which we can possess knowledge or justified belief, as opposed to privileged access, to what we want. Such knowledge might be the result of testimony. Or it might be the result of observing our own behavior. Or it might be the result, as Lawlor (2008) suggests, of observing our internal promptings and reasoning to the best explanation.[172] These are all ways, I submit, in which we might be able to come to possess less epistemically secure, non-uniquely first-personal knowledge of what we want. Granted, these are not ways of coming to know our desires that most philosophers concerned about self-knowledge tend to be interested in. But that does not entail that for all of us, philosophers included, they are not important ways we use to know what we want.[173] The point is just that while privileged access is not, I think, as rare as the above objector makes it out to be, it might very well be the case that these alternative ways of knowing what we want are more common means, for at least some agents, of knowing their desires. Such a concession, however, does not limit the importance of explaining how we have privileged access to some of our desires. Not only is it intuitively plausible to think that we do possess such access, but such access is, as I have argued elsewhere,[174] of great instrumental value to us.

Objection 2: PSA relies on the claim that when we attend to the experience of attraction in a careful, cautious manner, a manner in which we suspend our presuppositions, we will

[172] Such a means, admittedly, in the typical case would only be used in our coming to know our dispositional desires.
[173] This point shares some similarities with the points Cassam (2014) makes with respect to knowledge of dispositional attitudes.
[174] In "The Value of Privileged Access" (unpublished manuscript).

not be mislead by what is given to us in experience. But it is not clear how we can determine whether we are being careful or cautious enough, or whether we have suspended all our presuppositions. Insofar as we are looking for epistemic assurance that an experience won't mislead us, it is far from clear that we can have that. Given this, we should be skeptical about whether we are able to preclude being mislead by experience. And if that's the case, this calls into doubt whether ever do possess privileged access in the way that PSA outlines. Call this the *Speckled-Hen-Revisited Objection*.

Response: The objector appears to be trying to bait us into embracing the view that we need knowledge of the fact that we have adopted a sufficiently careful, scrupulously cautious attitude and truly bracketed out the relevant presuppositions in order to possess privileged access. Such knowledge, the objector implies, is necessary for us to guarantee that we are not going to be mislead by epistemic appearances. In response, such knowledge *might* be required for certain knowledge that one is instantiating the experience in question. But even granting this, I think it suffices that we actually *do* adopt such an attitude and suspension of beliefs, in order to possess *highly epistemically secure knowledge* of the experience in question.

Furthermore, I think there are ways we can have good reason to believe that we have adopted a cautious, careful enough attitude, and have bracketed out the relevant presuppositions about relational features of the experience in question. We can do so, it seems, by working to adopt such an attitude with respect to an experience, and then considering whether it's possible for us to (rationally) doubt that the experience in

question has the phenomenal character it epistemically appears to have. If we believe we genuinely can't, I think this gives us a defeasible reason to believe that (a) we can carry out such a process, and (b) we have done so with respect to the experience in question.

Objection 3: PSA requires that in order to have privileged access to the state in question, we not only have awareness of our experiences to a particular content, but also that we have awareness of our judgments that we have a particular experience to a given content, and that we have awareness of the correspondence or fit between this judgment and the content of our awareness. It seems sufficient for knowledge of attraction that the first-order awareness of our desires provides us with good evidence for the judgment in question. This alternative view would be internalist in nature in that that which justifies the judgment would be a mental state. It would also be internalist in the sense that we would be able, in principle, to access that which justifies our judgment. So, in short, it is not clear why we need to be able to be aware of a correspondence between judgments and the awareness of the output of the simulation. Such an epistemic requirement seems overly demanding. Call this *The Overdemandingness Objection*.[175]

Response: A full response to this objection is going to demand that I develop and defend some highly controversial epistemological positions. I won't attempt to develop and defend such positions here. I do, however, want to explain why I think that one needs to

[175] Some internalists have leveled this objection to Bonjour's account of the self-knowledge of attitudes. For instance, Feldman (2006) objects that Bonjour's requirement that one be aware of the fit between judgment and the built-in-constitutive awareness is unnecessary for possessing knowledge of one's attitudes. See (2006; p. 726) for this objection.

be aware of the correspondence between one's awareness of the experience and content in question and one's judgment about the former. But before doing so, I would like to say something concessive in reply to the above worry. If you are an internalist who is sympathetic with the view that no such awareness of the correspondence between the judgment and the content of the awareness is required in order to possess privileged access to one's desire, then you can still embrace my approach to the epistemology of desire. You could do so by holding that the awareness of the output of the phenomenal simulation process would provide one with the needed evidence for believing that one has the experience and content in question, evidence that if strong enough would provide one, with robust knowledge of the experience. So the contours of the account sketched in the last chapter could still be embraced by "internalists" unsympathetic with what I think is required in order to have privileged access.

Less concessively, though, it has always struck me that the main motivation for being an internalist is that when one has internalistic knowledge, one has a type of epistemic assurance that is simply lacking in the case of externalistic knowledge.[176] When one embraces mentalism of the stripe that many modest foundationalists are attracted to, I am compelled to think that one loses this main motivation for being an internalist. Allow me to explain. Mentalists of the type that would raise the above objection are committed to the view that it's enough that the mental states that serve as evidence for a belief make the truth of that belief probable. But if I am unaware of the

[176] What epistemic assurance amounts to is admittedly a difficult question to answer. Some internalists think that a definition of it is likely going to beg the question against externalists and perhaps some internalists as well (see Fumerton 2014; p. 99). One is tempted to say that when one has epistemic assurance for believing a proposition, one has all one could ask for in terms of epistemic support for their belief.

fact that the evidence for that belief makes that belief (at least) probable, in what sense do I have epistemic assurance that my belief is true? The mere fact that what justifies my belief is a mental state does not in and of itself provide me with such assurance. This is one reason why I am compelled to think that the mentalism-anti-mentalism debate does not really get to the heart of the internalist-externalist debate. Why, it might be asked should it matter if what justifies my beliefs are within the skin or not if I lack epistemic assurance concerning whether my belief is true? What it would take, it seems, to have such assurance, is awareness of the fact that your evidence (at the very least) makes probable (or in the best case scenario, guarantees) the truth of your belief.

Qua internalist, I care about possessing epistemic assurance that my beliefs, and in particular beliefs about my mind, are true. And if this is what internalists should care about, my account affords them with such epistemic assurance. What better assurance could you have then awareness of the fact that the truthmaker for your judgment corresponds with your judgment? For as I suggested in the last chapter, this seems to be as good as justification, and for that matter, knowledge gets.

Here is another way of understanding the above point that I have been trying to make. Fumerton (1995) embraces the view that in order to be justified in believing that P on the basis of E, one must be justified in (i) believing E and (ii) believing that E makes probable P. Most internalists these days reject (ii). In doing so, I submit, such internalists forfeit the main motivation for being internalists. If one is not justified in believing that E makes probable P, then in what sense does one truly have epistemic assurance that their belief is true. After all, externalists embrace (i). Granted, Fumerton's principle applies to

inferential justification. But I think something similar can be said about non-inferential justification. If one lacks justification for believing that what justifies one's belief at least makes probable the truth of what one believes, then I do think that one lacks the type of epistemic assurance that compels many epistemologists to be internalists in the first place. And pointing to the fact that what justifies our beliefs must be mental states does not make (epistemic) matters any better.

This, in short, is why I think that in order to possess privileged access to our attitudes, it must be the case that we are aware of the correspondence that holds between our judgment and the first-order awareness that judgment is about. But as noted above, if an internalist wishes to deny the need for this condition, one can still embrace the broad contours of my account.

Objection 4: PSA requires that we can form concepts that are in part constituted by experiences. But even granting that such concepts exist, it is clear that such concepts can only exist for the duration of the time that the experience exists. And given the lengthiness of the process we must undertake into order to come to be aware of our attitude, it seems reasonable to think that the experience will cease to exist, or cease to exist as it was, when we started this process. If that's the case, then we really can't employ such concepts in order to know our experiences with certainty.[177]

Response: I take this to be one of the more formidable objections to PSA. One response we might offer to it is that while occurrent sensations might be incredibly brief, they can

[177] See Poston (2014) for a more developed objection along these lines.

also exist for some duration. If the experience does continue for a long enough time, then I am not entirely sure why we couldn't carry out the process as I envisioned it. Admittedly, it might be possible for the experience to alter in such a way that we simply could not grasp the experience in the way I described. But, it's reasonable to think that an experience can continue for some time without altering its character. That being said, as I noted above, I take this to be a fairly serious worry for the view I defended, and it is one that I think demands further thought.

Objection 5: Assume for the sake of argument that the above account is correct and we do have privileged access to some of our desires. Such a result seems relatively inconsequential given that our occurrent desires are typically very short-lived. Why should we care that we have privileged access to states that are so fleeting, especially when our *occurrent* desires might change somewhat frequently based on circumstance? What we should care about, it seems, are what our dispositional, long-standing, stable desires are. So even if PSA explains how we have privileged access to such fleeting states, this seems like a fairly hollow victory for those of us interested in whether we do have privileged access to some of our desires, especially those of us interested in an account that can explain how we know the types of states that make us the people we are.

Call this the *Hollow Victory Objection*.

Response: On my view, occurrent desires are the desires that generate action. It is not dispositional desires that do. So occurrent desires along with the appropriate means-ends

beliefs, are the explanatory reasons (and, I would argue, normative reasons) for why we do what we do. Having knowledge to these states, then, puts us in an ideal position to determine why we act the way we act. This does not strike me as knowledge that is inconsequential. True, it might be the case that the desires we know via privileged access might not track the desires that actually end up generating action. This is the case because the desires we become aware of via phenomenal simulation would have to immediately generate action before they ceased to exist. But at least in some cases, the desire that we have privileged access to will be the same *type* of desire as the occurrent desire that actually does generate action.

Furthermore, insofar as it really is important that we know our dispositional desires (and I think it is), I am sympathetic with the view that occurrent desires can, in some cases, be a reliable guide to what we dispositionally desire. For although I argued in chapter 2 and 3 that we can occurrently desire that p while not dispositionally desiring that p, it will, with some regularity, be the case that an occurrent desire is the manifestation of an underlying dispositional desire with the same content. And it seems that we can be justified in believing, at least in some cases, that the underlying dispositional desire exists via inference from knowledge of our occurrent desire.[178] If this is correct, then knowledge of our occurrent desires would play an important epistemic role in helping us determine what our dispositional desires are, a way that might in fact be more reliable than observations of our behavior.

[178] It bears mentioning here that it is plausible to think that many of our dispositional desires are generated by our occurrent desires. If I occurrently desire to go to Hawaii, it seems reasonable to think that such a state generates a more long-standing desire to visit Hawaii.

Conclusion:

There are certainly other objections that can be leveled against PSA. But I hope to have offered reasonable responses to the objections presented in this section. In the next chapter, I turn to a discussion of how the account I presented with respect to desire can be applied to other types of mental states as well.

Chapter 7: Knowing when we are Curious, *or* Extending the Approach [179]

In the previous two chapters I explained how we come to possess privileged access to some of our desires. Such an account relied on the claim that some desires possess the unique phenomenology of attraction. If such desires lacked attraction, and were instead dispositional states of some kind, we would not possess privileged access to them. I think something similar can be said for other attitude types as well. Such a suggestion is not

[179] This chapter was largely inspired by Sanford Goldberg who early on in my project encouraged me to think about whether there is not just a puzzle about desire, but a puzzle about other attitude types as well.

insignificant since a number of philosophers embrace strict dispositionalist accounts with respect to other attitude types in addition to desire.[180]

In this chapter, I work to demonstrate that my approach to the metaphysics and epistemology of desire can be applied, mutatis mutandis, to other attitude types as well. The attitude type I will focus on is the state of being curious, or what I will refer to as a "c-state." I am going to contend that some c-states possess a unique phenomenology, and that we possess privileged access to this attitude on the basis of this phenomenology.[181] As we will see, the claim that c-states possess a unique type of phenomenology, i.e. a phenomenology no other state possesses,[182] runs up against a very standard view of curiosity—viz. that what it is to be curious is to have the desire to know. I argue below that this view is false.

I begin, however, with an example of what I have in mind when I talk about the attitude of curiosity.

The Phenomenology of Curiosity:

I live next to Wrigley Field, and on days when the Chicago Cubs play home games there, I can hear the roar from the crowd when the Cubs do something worth cheering about. On days when I am not so fortunate as to be able to follow what is happening on the field and I hear such roars, I become, qua Cubs fan, naturally curious as to what happened to generate such a cheer. So, for instance, I might hear the roar from the crowd and become

[180] Another way of putting this point is that there is not just a puzzle of desire; there is a puzzle with respect to: (a) any attitude type we might initially think is strictly dispositional in nature, and (b) any attitude it seems plausible to think we have, in some cases, privileged access to. And while I won't defend the claim here, I do think it is plausible that *attitude access dispositionalism*—the view that we have privileged access to some dispositional attitudes—is false.

[181] I want to stress at the outset that, as with desire, I do not think we have privileged access to *all* of our occurrent states of being curious.

[182] Or conjunction of states.

curious as to whether a Cubs' player performed an amazing athletic feat. At such a moment I possess the attitude of curiosity concerning whether it's true a Cubs' player performed an amazing athletic feat.[183]

I think such states possess a felt quality. Christopher Shields (2011) agrees. The latter, however, has more lofty curiosities than I do, and he focuses on one of them in defense of the view that some c-states have a phenomenal character, writing:

> When I am curious as to whether non-carbon-based life is even a possibility in any imaginable universe, then the state I am in has a perfectly familiar phenomenology. Being curious as to whether there is non-carbon-based life in the universe is unlike hoping that this is so, or fearing that this is so, or dreading that this is so. (217)

Shields acknowledges that it is difficult to conceptually and/or linguistically capture what the phenomenology of curiosity is like, but he suggests that the unique phenomenology of c-states feels "like being intrigued" or "having a mental itch" (215). What Shields is referencing here is a type of phenomenology, I think, that is familiar to anyone who has been struck wondering whether God exists, been intrigued about what Pierre Bayle actually believed, or, in a less lofty vein, became interested in determining whether their favorite team won. I am going to use the phrase *felt-intrigue* to refer to the phenomenology Shields and I think states of curiosity have.

[183] I will have more to say below about what the "objects" of curiosity are.

Now Shields does not merely assert that being curious has a unique phenomenology; he offer arguments in defense of the view that it is the attitude of curiosity *itself* that possesses felt-intrigue, arguments he calls "parity arguments." I won't review such arguments here. But suffice to say I think the same role functionalist move I suggested conservatives can make with respect to partial-zombie arguments and contrast arguments, in chapter 4, can be used to respond to Shields' parity arguments.

How, then, should a liberal go about defending the view that c-states possess phenomenology, and in particular, the unique phenomenology of felt-intrigue? I believe the way to defend such a claim is the way I suggested we defend the phenomenology of desire—viz. by arguing that the states that realize the causal role of being curious are states that possess a unique phenomenology. Such an approach is reasonable since it is not just desires we should be realizer functionalists about, but all attitudes. Insofar as this is true, if some c-states really do possess a unique phenomenology then it must be the case that some states that realize the causal role of being curious possess a unique phenomenology. I am now going to argue that this is in fact the case. To offer this defense I first need to explain what I take the causal role of c-states to be. Admittedly, my view on this matter is controversial. But allow me to first begin by offering what I take to be the typical causes of c-states.

It seems that in the standard case (and arguably every case), what causes states of curiosity will be mental states of some kind. My auditory experience of a huge roar from the Wrigley Field crowd, my belief that something went the Cubs' way, and my desire that the Cubs win, might cause me to be curious as to what caused that roar. An

aesthete's visual experience of an aesthetically impressive sculpture along with the belief that the artist in question is a 12-year old, might naturally cause such a person to be curious about the name and further background information about the sculptor in a question. A wanderlust's desire to visit Honolulu along with the belief that she is going to have to fly to get to that city, will perhaps cause this person to be curious about the price of flights to Honolulu. In the usual case, then, what is going to cause c-states are other mental states. If one wants to consider earlier links in the causal chain leading to c-states as typical causes of such attitudes as well, then one would be inclined to include physical objects as causes of being curious as well. And in one clear sense, it is reasonable to say that what caused the art lover to be curious about the sculptor in the above example, was in part the sculpture itself.

One might wonder, though, whether there is anything more precise we can say about the typical causes of c-states. There might be. But I suspect that the type of precision in question might only be found by investigating particular agents and their typical interests, preferences, proclivities, background knowledge, etc. For the fact of the matter is, the types of experiences, beliefs, desires, physical objects etc. that cause states of curiosity are going to depend on the person tokening such mental states. My neighbors might not be curious about what caused the roar from a Wrigley Field crowd if they are not baseball fans. They might instead be annoyed by such roars and attempt to ignore them. Most young children, I take it, lack the cognitive maturity and background knowledge to be curious about the writing of James Joyce. So while it might be nice to have something more precise to say about the typical causes of c-states, I would not be

surprised if this precision wasn't to be found. Still, I think it is at least relatively uncontroversial to suggest that in the typical case, what causes c-states are mental states of some kind.

It is more controversial what the typical outputs of c-states are, i.e. what c-states cause. On the view I am sympathetic with, such outputs will typically be the desire to find out the answer to a question one is curious about, and at least in some cases, this will in turn engender action to put to rest one's curiosity. If I am curious as to what caused the roar from Wrigley, this will in turn cause me to desire to find out what caused it, a desire that might lead me, if I am not too busy, to check for an update to the game on the web, or send a text to a friend who is following the action. If a person, for instance, is curious about the price of a ticket to Honolulu, this will naturally cause her to desire to answer this question, and perhaps, surf an online travel site to find out. C-states, then, on my view, typically cause us to desire to find out an answer to that which we are curious about, and in some cases, generate action to put to rest that question. The action, though, is going to be generated by the desire, not the c-state.

Now I mentioned that my account of the causal role of c-states is controversial. This is because a number of philosophers think that c-states in the typical case, don't cause, e.g. the desire to know; rather, being curious just is the desire to know. And it must be acknowledged that it is difficult, to come up with a clear example in which a person is curious and yet completely lacks the desire to find an answer to that which they are curious about. There seems to be, at the very least an intimate connection between c-states and desires, a relationship that has led many to think that being curious just is a

type of desire. If this is in fact the case, then the functional nature of c-states offered above is flawed, and the epistemology of curiosity reduces to a particular version of the epistemology of desire, the type of view defended in chapter V.

But while the view that c-states are a species of desire is arguably the standard view of curiosity, such a view, as Dennis Whitcomb (2010) notes, is rarely defended.[184] When philosophers do defend an account of the metaphysics of curiosity, they tend to presuppose that c-states are desires and focus on what curiosity is a desire for.[185] Whitcomb, himself, only offers two lines in defense of the claim that being curious is a type of desire. His defense rests on the relatively undefended contentions that: (i) c-states motivate and (ii) c-states are satisfiable.[186] He takes (i) and (ii) to be good reason to think that such states are desires, since he holds that desires motivate and are states that have satisfaction conditions.

Now some might think that Whitcomb's reasons fail to provide us with a strong reason for believing that c-states are desires because desires are not the only states that motivate and/or desires are not the only states with satisfaction conditions. I am somewhat inclined to embrace the latter claim. But as I argued in chapter 4, desires are the only states that generate action. So if (i) is true, I take (i) to be sufficient reason to embrace the view that desires are c-states.

---

[184] Foley (1987) and Goldman (1991) appear to think being curious is a desire for true belief. Williamson (2000) appears to view curiosity as a desire for knowledge.

[185] Kvanvig (2003; p. 145-146), for example, in what Whitcomb notes is the only contemporary defense of the metaphysical nature of curiosity, argues that curiosity is a desire for perceived truth.

[186] Whitcomb (2010; 671) writes: "Curiosity is motivational; sometimes we read or listen *because* we are curious. It is also satisfiable; it gets satisfied in the same way that thirst gets satisfied, namely by getting what it is a desire for. Hence we should keep the view that curiosity is a desire" (671).

But I don't think (i) is true. And my reason for this is based part in part on my views on the metaphysics of desire. I argued in chapter 4 that attraction is the unique phenomenology of desire and that there is good reason to think that only desires generate action. But if that's the case, and if c-states really do possess a phenomenology, then they would have to possess the phenomenology of attraction in order to move us to act. But felt-intrigue, I think, doesn't feel like attraction. So since felt-intrigue seems to be the phenomenology that is at least associated with c-states (not attraction), this provides us with a defeasible reason for thinking that c-states really aren't desires after all.

Admittedly, much more can and needs to be said in defense of the view that c-states are *not* desires. There will certainly be those who remain unconvinced by the above line-of-reasoning. For those sympathetic with the standard view of curiosity, the account of the epistemology of desire I offered in chapter 5 is, it seems, applicable to c-states. The idea would be that we could be aware of an attraction[187] to the object of our desire, where such an object would be knowledge of the answer to a particular question.[188] In part, on the basis of such awareness we could come to possess privileged access to the fact that we had a particular desire to know (or have a true belief) about P, which would on the view we are considering, just be a c-state.

But insofar as no one in the extant literature on curiosity to my knowledge has offered a well-developed defense of the view that c-states are desires, and given that my metaphysical views on the nature of the latter type of state combined with the above phenomenological appeal, give me reason to believe that c-states aren't desires, I am

[187] Perhaps a particular type of attraction.
[188] Or a true belief about the answer to a particular question.

going to proceed under the assumption that c-states are not a particular type of desire, and that the causal profile of being curious is roughly what I outlined above. Recall that on the approach offered above, the causal profile of c-states is that they are states that are caused by experiences, beliefs, desires and other types of mental states, and in turn cause the desire to seek an answer to that which one is curious about, the belief that such an investigation would be interesting, etc. But if this is the causal role of c-states, why should we think that some c-states possess the phenomenology of felt-intrigue?

Consider again the states I typically token when I hear roars go up from Wrigley Field. Ostensibly, what causes me to be curious about such sounds are, at a minimum: (i) the auditory experience of a crowd cheering, (ii) my belief that something went the Cubs' way, and (iii) my desire to know what happened.[189] And typically, I am in turn (iv) caused to be struck with a strong urge to investigate matters further. Now it does not seem unreasonable to think that what (i)-(iii) cause and what is caused by (iv) is, at least in some cases, states that possess felt-intrigue. When I hear such sounds come from the stadium, such an experience is usually followed by the type of experience Shields references, an experience that is shortly followed by an experience of attraction I have to switch screens on my laptop to discover what happened. It does not seem unreasonable, then, to think that what is playing the causal role of curiosity in at least some of these cases are states with felt-intrigue. And insofar as realizer functionalism is true, such states just would be c-states.

[189] There are likely other states that causally contribute to my being curious, but the mental states listed above, I think, are the least controversial.

As in the case of desire, there are those who will want to deny that what I am calling felt-intrigue is non-sensory in nature. I pointed out in chapter 4 that partial-zombie arguments and contrast cases force those who want to deny the existence of non-sensory phenomenology to explain how the phenomenology of the agents in such scenarios reduces to sensory phenomenology (or deny the possibility of such cases). As I implied in chapter 4, these arguments really do force the NSAP conservative to determine what sensory experience or combination of sensory experiences felt intrigue reduces to. An NSAP conservative might attempt to rise to the challenge by arguing that felt-intrigue is a combination of anticipation, excitement, or some variety of pleasure.

An NSAP liberal can reasonably object, though, that we can token experiences of felt-intrigue without experiencing any of the above types of phenomenology. We can, it seems, be struck by an experience of felt-intrigue without also being struck with a feeling of excitement. We can, it appears, token an experience of felt-intrigue without experiencing delight. At the very least, then, the burden is on the conservative, I think, to come up with a reasonable combination of sensory experiences that capture what I have been calling felt-intrigue.

Up to this point I have argued that states of curiosity possess a unique phenomenology, phenomenology that cannot be reduced to more familiar sensory phenomenology. Such a claim is, as I suggested above, not innocuous. If I am right and c-states possess a unique phenomenology unlike the phenomenology of desiring, then the common view concerning what being curious is, is flawed. The phenomenology of curiosity does not feel like desiring. And on my view, it does not feel like the

phenomenology of any other attitude type (or combination of attitude types). The proprietary phenomenology I am suggesting c-states possess, can help us explain how we have privileged access to such states. Before I offer such an account, I need to say a bit more about what I take the objects of curiosity to be.

I think it is reasonable to hold that what we stand in a relation to when we are curious is not a proposition, but rather an interrogative.[190] This suggestion admittedly runs up against the popular but typically undefended claim that all attitudes are propositional attitudes. But in defense of this claim, it seems at the very least strange to even try to *model* the attitude of being curious in the same way that we model the attitude of believing or hoping. "Curious that", as a locution, is forced at best. We are curious *whether* a proposition is true. We are curious *about* what the answer to a question is. But it does not seem plausible to suggest that we are curious *that* some proposition is true. Perhaps the linguistic evidence is misleading. But I am inclined to think that what we are curious about, what we experience felt-intrigue to are not objects with truth values, but rather interrogatives.[191] On the view I am sympathetic with, when a person is curious about whether it will rain today, she adopts the attitude of curiosity towards the question: will it rain today? To adopt such an attitude is just to be curious about *what* the answer is to that question.

---

[190] Whitcomb (2010; p. 671-672) holds a similar view arguing that the contents of desires are interrogatives. But as noted above, Whitcomb thinks that all c-states are desires.

[191] Even in cases where we are curious about a proposition, assertion, belief, etc. we appear to be curious concerning a particular question about that proposition, assertion, belief, etc. Consider the following example: I might be reading McDowell's *Mind and World* (1996) and come across a proposition I am curious about. I am not, though, "curious that" with respect to that proposition. I am curious about some feature of that proposition. I might for instance be curious whether the proposition means what I take it to mean. Or I might be curious about whether McDowell really believes that proposition. In other words, I seem to be curious about what the answer is to particular questions about the proposition.

Let's proceed as if this is correct.[192] How then, if at all, do we come to possess privileged access to our c-states? I think that the way we have privileged access to our c-states can be explained in a manner that parallels the explanation offered in chapter 5 concerning how we have privileged access to our desires. What we hold before consciousness with respect to c-states, though, is a particular interrogative. Experiencing felt-intrigue toward this question enables us to be aware that we are curious about a particular question, and more specifically, that we are curious about what the answer to *that* particular question is. An example will help: Shields is curious about whether non-carbon-based life is possible. How could Shields come to possess privileged access that he is curious about whether non-carbon-based life is possible? My suggestion is that he can come to possess such access by holding before his mind the question: "Is non-carbon-based life possible?" and then attending to his experience of felt-intrigue towards that question. Phenomenal simulation puts Shields in a prime position to know in a highly epistemically secure way that he is curious about the answer to that particular question.

We of course need to explain how Shields is able to refrain from forming false beliefs about felt-intrigue, as well as how he is able to become aware of such an experience in a way that enables him to conceptualize it accurately. I have detailed how I think we accomplish the above with respect to desire, and it will not surprise the reader to learn at this point that I think a similar story can be offered with respect to c-states. By carefully attending to the phenomenal features of the experience in question, adopting the attitude of scrupulous caution, bracketing out presuppositions about the relational features of this experience, the way such an experience epistemically appears to Shields

[192] If it is not, one can still embrace the broad contours of the epistemology of curiosity I offer.

will be constituted by the experience's phenomenal reality. His awareness of such an appearance would enable him to be aware of the phenomenal reality in question. He could then conceptualize this experience via the way the experience appears to him, a conceptualization that will enable his judgment that he is experiencing "this" to the particular question to be accurate, where "this" refers to the felt intrigue. The final step of the process would involve Shields becoming aware that his judgment corresponds or fits that which the judgment is about.

If one is worried that Shields doesn't truly possess privileged access to the fact that he is curious about a certain question, we can advance, I believe, a similar account to the one advanced in chapter 5 concerning ordinary language concept application. I will not outline the details here, but I think a similar story can be told with respect to the concept of curiosity as was told with respect to desire. We can become aware, that is, of the fact that felt-intrigue is the experience of being curious. If this is correct, Shields could then come to have epistemically secure knowledge of the fact that he is curious about the answer to a particular question, knowledge that is grounded in his privileged access to his experience of felt-intrigue.

Now it is certainly the case that more work needs to be done with respect to both the metaphysics of curiosity as well as the epistemology of such states, but I hope to have defended enough substantive claims here that the broad contours of how I think we have privileged access to c-states is clear.

Conclusion:

In this chapter I have defended the view that (a) c-states possess a unique phenomenology, and (b) we have privileged access to c-states in a similar way that we have privileged access to our desires. I want to conclude with the suggestion that the way I have explained how we have privileged access to these two attitude types can serve as a template for other attitude types as well. Judgments[193], for instance, on my view have a unique phenomenology, phenomenology that plays a crucial epistemic role in how we come to know these mental events in a uniquely first-personal, epistemically secure manner. As with desire, taking into account the unique epistemology of judgments, I believe, can have an impact in areas beyond epistemology as well. Future work, I believe, will bear this out.

[193] One might think of judgments as mental actions and not as attitudes proper. This, in part, is one reason why I refrained from focusing on judgments in this chapter. I did not want to engage in that debate. But even if judgments are mental actions, I take it we still want a plausible story of how we know such actions.

Bibliography

Alston, William. “Varieties of Privileged Access.” *American Philosophical Quarterly* 8 (1971): 223-241.

Alvarez, Maria. *Kinds of Reasons: An Essay in Philosophy of Action*. Oxford: Oxford UP, 2011.

Armstrong, David. *A Materialist Theory of the Mind*. London: Routledge P, 1968.

Ashwell, Lauren. “Deep, Dark, or… Transparent? Knowing our Desires.” *Philosophical Studies* 165.1 (2013): 245-246.

---. “The Metaphysics of Desires and Dispositions.” *Philosophy Compass* 9.7 (2014): 469-477.

Audi, Robert. "Dispositional Beliefs and Dispositions to Believe." *Nous* 28 (1994): 410-434.

Balog, Katalin. "Acquaintance and the Mind-Body Problem." *New Perspectives on Type-Type Identity Theory*. Ed. Simone Gozzano & Christopher Hill. Cambridge: Cambridge UP, 2012. 16-42.

Bayne, Timothy & Michelle Montague. "Cognitive Phenomenology: An Introduction." *Cognitive Phenomenology*. Eds. Timothy Bayne & Michelle Montague. Oxford: Oxford UP, 2011. 1-34

Bickle, John. "Multiple Realizability." *The Stanford Encyclopedia of Philosophy*. Ed. Edward Zalta. Fall 2013.

Block, Ned. "Introduction: What is Functionalism?" *Readings in Philosophy of Psychology Vol. 1*. Ed. Ned Block. Cambridge: Harvard UP, 1980. 171-183.

Boghossian, Paul. "Content and Self-Knowledge." *Philosophical Topics* 17 (1989): 5-26.

Bonjour, Laurence. "Back to Foundationalism." *Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtues*. Laurence Bonjour and Ernest Sosa. Malden: Wiley-Blackwell P, 2003. 60-76.

Boyle, Matthew. "Two Kinds of Self-Knowledge. Philosophy and Phenomenological Research 78.1 (2009): 133-164.

Braddon-Mitchel, David. & Frank Jackson. *Philosophy of Mind and Cognition*. 2nd Edition. Oxford: Blackwell P, 2007.

Burge, Tyler. "Our Entitlement to Self-Knowledge." *Proceedings of the Aristotelian Society* 96, 1996: 91-116.

Byrne, Alex. "Introspection." *Philosophical Topics* 33.1 (2005): 79-104.

---. "Knowing that I am Thinking." *Self-Knowledge*. Ed. Anthony Hatzimoysis. Oxford: Oxford UP, 2011. 105-124.

---. "Knowing what I Want." *Consciousness and the Self: New Essays*. Eds. JeeLoo Liu and John Perry. Cambridge: Cambridge UP, 2012. 165-183.

Carruthers, Peter. *Consciousness: Essays from a Higher-Order Perspective*. Oxford: Oxford UP, 2005.

---. "Introspection: Divided and Partly Eliminated. *Philosophy and Phenomenological Research* 80.1 (2010): 76-111.

---. *The Opacity of the Mind: An Integrative Theory of Self-Knowledge*. Oxford: Oxford UP, 2011.

Cassam, Quassim. *Self-Knowledge for Humans*. Oxford: Oxford UP, 2014.

Chalmers, David. *The Character of Consciousness*. Oxford: Oxford UP.

---. *The Conscious Mind*. Oxford: Oxford UP, 1996.

---. "The Content and Epistemology of Phenomenal Belief." *Consciousness: New Philosophical Perspectives*. Eds. Aleksandar Jokic and Quentin Smith. Oxford: Oxford UP, 2003. 220-272.

Chang, Ruth. "Can Desires Provide Reasons for Action?" *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*. Ed. R. Jay Wallace. Oxford: Oxford UP, 2008. 56-90.

Chisholm, Roderick. "The Problem of the Speckled Hen." *Mind* 51 (1942): 368-373.

Chudnoff, Elijah. *Cognitive Phenomenology*. New York: Routledge, P, 2015.

Churchland, Paul. "Eliminative Materialism and Propositional Attitudes." *Journal of Philosophy* 78, 1981: 67-90.

Cohen, Stewart. "Justification and Truth." *Philosophical Studies* 46 (3), 1984: 279-95.

Conee, Earl & Richard Feldman. "The Generality Problem." *Evidentialism*. Oxford UP, 2004.

Crane, Tim. *Elements of Mind*. Oxford: Oxford UP, 2001.

Dreiser, Theodore. *Sister Carrie*. 1900. New York: Penguin P, 1994.

Dretske, Fred. "The Epistemology of Pain." *Pain: New Essays on its Nature and the Methodology of its Study*. Ed. Murat Aydede. Cambridge: MIT P, 2005. 59-73.

---. "How Do you Know you are Not a Zombie?" *Privileged Access*. Ed. Brie Gertler. Aldershot: Ashgate P. 1-13.

---. "Introspection". *Proceedings of the Aristotelian Society* 94, 1994: 263-278.

Evans, Gareth. *The Varieties of Reference*. Oxford: Oxford UP, 1982.

Fales, Evan. *A Defense of the Given*. Lanham, MD: Rowman and Littlefield P, 1996.

---. *Causation and Universals*. New York: Rouledge P, 1990.

Feldman, Richard. "Bonjour and Sosa on internalism, Externalism, and Basic Beliefs." *Philosophical Studies* 131: 713-728.

---."Foundational Justification." *Ernest Sosa and his Critics*. Ed. John Greco. Malden: Blackwell P. 42-58.

Fernandez, Jordi. "Desire and Self-Knowledge." *Australasian Journal of Philosophy* 85.4 (2007): 517-536.

Foley, Richard. *The Theory of Epistemic Rationality*. Cambridge: Harvard UP, 1987.

Fumerton, Richard. *Metaepistemology and Skepticism*. Lanham, MD: Rowman and Littlefield P, 1995.

---. *Knowledge, Thought, and the Case for Dualism*. Oxford: Oxford UP, 2013.

Funkhouser, Eric. "Multiple Realizability." *Philosophy Compass* 2.2 (2007): 303-315.

Gertler, Brie. "In Defense of Mind Body Dualism." *Reason and Responsibility*. Eds. Joel Feinberg & Russ Shafer-Landau. 13th Edition. Belmont: Thomson-Wadsworth P, 2008. 285-296.

---. "The Mechanics of Self-Knowledge." *Philosophical Topics* 28 (2) (2000): 125-146.

---. "Renewed Acquaintance." *Introspection and Consciousness*. Eds. Declan Smithies & Daniel Stoljar. Oxford: Oxford UP, 2012. 93-127.

---. *Self-Knowledge*. New York: Routledge P, 2011a.

---. "Self-Knowledge and Rational Agency." *Philosophy and Phenomenological Research* (2016): 1-19.

---. "Self-Knowledge and the Transparency of Belief." *Self-Knowledge*. Ed. Anthony Hatzimoysis. Oxford: Oxford UP, 2011b. 125-145.

Goldman, Alvin. "Interpretation Psychologized." *Mind and Language* 4 (1989): 161-185.

---. *Knowledge in a Social World*. Oxford: Oxford UP, 1999.

---. "The Psychology of Folk Psychology." *Behavioral and Brain Sciences* 16 (1993): 15-28.

---. *Simulating Minds*. Oxford: Oxford UP, 2006.

Gordon, Robert. "Folk Psychology as Simulation." *Mind and Language* 1 (1986): 158-171.

Graukroger, Stephen. "Vico and the Maker's Knowledge Principle." H*istory of Philosophy Quarterly* 3.1 (1986): 29-44.

Horgan, Terence. "Nonreductive Materialism and the Explanatory Autonomy of Psychology." *Naturalism: A Critical Appraisal*. Eds. Steven Wagner & Richard Warner. Notre Dame: Notre Dame UP, 1993. 295-320.

Horgan, Terrence & Uriah Kriegel. "Phenomenal Epistemology: What is Consciousness that we may Know it so Well?" *Philosophical Issues* 17 (2007): 123-144.

Horgan, Terence & John Tienson. "The Intentionality of Phenomenology and the Phenomenology of Intentionality." *Philosophy of Mind: Classical and Contemporary Readings*. Ed. David Chalmers. Oxford: Oxford UP, 2002. 520-533.

Jackson, Frank. "Epiphenomenal Qualia." *Philosophical Quarterly* 32 (1982): 127-136.

James, William. *The Principles of Psychology*. Cambridge: MIT P, 1891/1981.

Kim, Jaegwon. "Multiple Realization and the Metaphysics of Reduction." *Philosophy and Phenomenological Research* 52 (1992): 1-26.

---. Philosophy of Mind. 3rd Edition. Boulder: Westview P, 2011.

Kind, Amy. "Shoemaker, Self-blindness, and Moore's Paradox." *Philosophical Quarterly* 53, 2003: 39-48.

Kornblith, Hilary & Derk Pereboom. "The Metaphysics of Irreducibility." *Philosophical Studies* 63.2 (1991): 125-145.

Kriegel, Uriah. "The Phenomenal Intentionality Research Program." *Phenomenal Intentionality*. Ed. Uriah Kriegel. Oxford: Oxford UP, 2013. 1-26.

---. *The Sources of Intentionality*. New York: OUP, 2011.

---. *Varieties of Consciousness*. Oxford: Oxford UP, 2015.

Kripke, Saul. *Naming and Necessity*. Cambridge: Harvard UP, 1981.

Kvanvig, Jon. *The Value of Knowledge and the Pursuit of Understanding*. Cambridge: Cambridge UP, 2003.

Lehrer, Keith. "Self-Presentation, Representation, and the Self." *Philosophy and Phenomenological Research* 64.2 (2002): 412-430.

Levine, Joseph. "On the Phenomenology of Thought." *Cognitive Phenomenology*. Eds. Tim Bayne & Michelle Montague. Oxford: Oxford UP, 2011. 103-120.

Lawlor, Krista. "Knowing What One Wants." *Philosophy and Phenomenological Research* 79.1 (2009): 47-75.

Lewis, David. "An Argument for the Identity Theory. *Journal of Philosophy* 63 (1966): 17-25.

---. "Desire as Belief." *Mind* 97 (1988): 323-332.

---. "Desire as Belief II." *Mind* 10 (1996): 303-313.

---. "Mad Pain and Martian Pain." *Readings in Philosophy of Psychology Volume I*. Ed. Ned Block. Cambridge: Harvard UP, 1980. 216-232.

Lycan, William. *Consciousness and Experience*. Cambridge: MIT P, 1996.

---. "In Defense of the Representational Theory of Qualia." *Philosophical Perspectives 12: Language, Mind, and Ontology*. Ed. James Toberlin. Atascadero, California: Ridgeview P. 479-486.

Medina, José. "What's so Special about Self-Knowledge?" *Philosophical Studies* 129 (2006): 575-603.

Montague, Michelle. "Against Propositionalism." *Noûs* 41.3 (2007): 503-518.

Moran, Richard. *Authority and Estrangement: An Essay on Self-Knowledge*. New Jersey: Princeton UP, 2001.

---. “Responses to Shoemaker and O’Brien.” *European Journal of Philosophy* 11, 2003: 402-419.

---. "Self-Knowledge, 'Transparency,' and the Forms of Activity." *Introspection and Consciousness*. Eds. Declan Smithies & Daniel Stoljar. Oxford: Oxford UP, 2012. 211-236.

Mumford, Stephen. *Dispositions*. Oxford: Oxford UP, 1998.

Neander, Karen. “The Division of Phenomenal Labor: A Problem for Representational Theories of Consciousness.” *Philosophical Perspectives 12: Language, Mind, and Ontology*. Ed. James Toberlin. Atascadero, California: Ridgeview P.

Nelkin, Norton. "Propositional Attitudes and Consciousness." *Philosophy and Phenomenological Research* 49.3 (1989): 413-430.

Neta, Ram. "The Nature and Reach of Privileged Access." *Self-Knowledge*. Ed. Anthony Hatzimoysis. Oxford: Oxford UP, 2011. 9-31.

Nichols, Shaun and Stephen Stich. *Mindreading*. Oxford: Clarendon P, 2003.

O'Brien, Lucy. "Moran on Agency and Self-Knowledge." *European Journal of Philosophy* 11 (2003): 375-390.

Oddie, Graham. *Value, Reality, and Desire*. New York: Oxford UP, 2005.

Parfit, Derek. "Rationality and Reasons." *Exploring Practical Philosophy: From Action to Values*. Eds. Dan Egonsson, Jonas Josefsson, Bjørn Petersson, & Toni Ronnow-Rasmussen. Burlington: Ashgate P, 2001. 17-39.

Peacocke, Christopher. *Being Known*. New York: Oxford UP, 1999.

---. "Our Entitlement to Self-Knowledge: Entitlement, Self-Knowledge, and Conceptual Redeployment." *Externalism and Self-Knowledge*. Eds. Peter Ludlow and Norah Martin. Stanford: CSLI P, 1998. 265-303.

Pitt, David. "Introspection, Phenomenality, and the Availability of Content." *Cognitive Phenomenology*. Eds. Tim Bayne & Michelle Montague. Oxford: Oxford UP, 2011. 141-173.

---. "The Phenomenology of Cognition or What is it Like to Think that P?" *Philosophy and Phenomenological Research* 69, 2004: 1-36.

Polger, Thomas. "Putnam's Intuition." *Philosophical Studies* 109 (2002): 143-170.

Poston, Ted "Direct Phenomenal Beliefs, Cognitive Significance, and the Specious Present." *Philosophical Studies* 168.2 (2014): 483-489.

Price, Huw. "Defending Desire as Belief." *Mind* 98 (1989): 119-127.

Prinz, Jesse. *Gut Reactions: A Perceptual Theory of Emotion*. New York: Oxford UP, 2004.

---. "The Sensory Basis of Cognitive Phenomenology." *Cognitive Phenomenology*. Eds. Tim Bayne & Michelle Montague. Oxford: Oxford UP, 2011. 174-196.

Putnam, Hilary. "Psychological Predicates." *Art, Mind, and Religion*. Eds. W.H. Capitan and D.D. Merrill. Pittsburgh: Pittsburgh UP, 1967. 37-48.

Raz, Joseph. *Engaging Reason*. Oxford: Oxford UP, 1999.

Rosenthal, David. *The Nature of Mind*. New York: Oxford UP, 1994.

Ryle, Gilbert. *The Concept of Mind*. London: Hutchinson P, 1949.

---. Speckled hen.

Russell, Bertrand. *The Analysis of Mind*. New York: Macmillan P, 1921.

---. "On the Nature of Acquaintance." *Logic and Knowledge*. Ed. R.C. Marsh. New York: G.P. Putnam's Sons' P, 1956. 127-174.

---. *Problems of Philosophy*. New York: Henry Hart & Company P, 1912.

Scanlon. T.M. *What We Owe to Each Other*. Cambridge: Harvard UP, 1998.

Schroeder, Timothy. "Desire." *The Stanford Encyclopedia of Philosophy*. Ed. Edward Zalta. Fall 2015.

---. *Three Faces of Desire*. Oxford: Oxford UP, 2004.

Schueler, G.F. *Desire*. Cambridge: Cambridge UP, 1995.

Schultheiss, O.C. and J.C. Burstein. "Goal Imagery: Bridging the Gap between Implicit Motives and Explicit Goals. *Journal of Personality* 61 (1999): 1-38.

Schwitzgebel, Eric. "A Phenomenal, Dispositional Account of Belief." *Nous* 36.2 (2002): 249-275.

---. "The Unreliability of Naïve Introspection." *Philosophical Review* 117.2 (2008): 245-273.

Searle, John. *Intentionality*. Cambridge: Cambridge UP, 1983.

---. *The Rediscovery of the Mind*. Cambridge: MIT P, 1992.

Senor, Thomas. "Internalist Foundationalism and the Justification of Memory Belief." *Synthese* 94.3 (1993): 453-476.

Shields, Christopher. "On Behalf of Cognitive Qualia." *Cognitive Phenomenology*. Eds. Tim Bayne & Michelle Montague. Oxford: Oxford UP, 2011. 215-235.

Shoemaker, Sydney. "Moran on Self-Knowledge." *European Journal of Philosophy* 3.3 (2003): 391-401.

---. “On Knowing One’s Own Mind.” *Privileged Access*. Ed. Brie Gertler. Aldershot: Ashgate P, 2003. 111-129.

---. “Self-Knowledge and Inner Sense.” *Philosophy and Phenomenological Research* 54 (1994): 249-315.

Siewert, Charles. *The Significance of Consciousness*. New Jersey: Princeton UP, 1998.

Siegel, Susanna. "Which Properties are Represented in Experience?" *Perceptual Experience*. Eds. Tamar Szabo-Gendler & John Hawthorne. Oxford: Oxford UP, 2005. 481-503.

Silins, Nicholas. "Judgment as a Guide to Belief." *Introspection and Consciousness*. Eds. Declan Smithies & Daniel Stoljar. Oxford: Oxford UP, 2012. 295-327.

Smith, Michael. *The Moral Problem*. Malden: Blackwell P, 1994.

Smithies, Declan. "The Nature of Cognitive Phenomenology." *Philosophy Compass* 8 (2013): 744-754.

Stalnaker, Robert. *Inquiry*. Cambridge: Cambridge UP, 1984.

Stampe, Dennis. "The Authority of Desire." *Philosophical Review* 96 (1997): 335-81.

Stich, Stephen. *From Folk Psychology to Cognitive Science*. Cambridge: MIT P, 1983.

Strawson, Galen. “Intentionality and Experience: Terminological Preliminaries.” *Phenomenology and Philosophy of Mind*. Eds. D.W. Smith and Amie Thomason. Oxford: Oxford UP, 2005. 41-46.

---. *Mental Reality*. Cambridge: MIT P, 1994.

Stocker, Michael. Desiring the Bad: An Essay in Moral Psychology." *Journal of Philosophy* 76 (1979): 738-753.

Tye, Michael. *Consciousness, Color, and Content*. Cambridge: MIT P, 2005.

---. "Qualia." *The Stanford Encyclopedia of Philosophy*. Ed. Edward Zalta. Fall 2015.

Tye, Michael & Biggs Wright. "Is There a Phenomenology of Thought?" *Cognitive Phenomenology*. Eds. Tim Bayne & Michelle Montague. Oxford: Oxford UP, 2011. 326-344.

Whitcomb, Dennis. "Curiosity was Framed." P*hilosophy and Phenomenological Research* 81.3 (2010): 664-687.

Wilson, Timothy and E.W. Dunn. "Self Knowledge: its Limits, Vaule, and Potential for Improvement." *Annual Review of Psychology* 55 (2004): 493-518.

Wilson, Timothy & Dolores Kraft. "Why Do I Love Thee?: Effects of Repeated Introspections about a Dating Relationship on Attitudes toward the Relationship." *Personality & Social Psychology Bulletin* 19.4 (1993): 409-418.

Williamson, Timothy. *Knowledge and its Limits*. Oxford: Oxford UP, 2000.

Zimmerman, Aaron. "Basic Self-Knowledge: Answering Peacocke's Criticisms of Constitutivism." *Philosophical Studies* 128 (2006): 337-379.