

OUR KNOWLEDGE  
OF THE INTERNAL  
WORLD

ROBERT C. STALNAKER

CLARENDON PRESS · OXFORD

# Contents

<i>Acknowledgements</i>	vi
1. Starting in the Middle	i
2. Epistemic Possibilities and the Knowledge Argument	24
3. Locating Ourselves in the World	47
3A. Notes on Models of Self-Locating Belief	69
4. Phenomenal and Epistemic Indistinguishability	75
5. Acquaintance and Essence	94
6. Knowing What One is Thinking	112
7. After the Fall	132
<i>References</i>	139
<i>Index</i>	145

# I

## *Starting in the Middle*

Analyze theory-building how we will, we all must start in the middle.

W. V. Quine<sup>1</sup>

The Cartesian picture of the mind, and of the world, was under attack from a variety of directions throughout most of the last century. We were taught to do without private objects, and private languages, the myth of the given, the ghost in the machine, the Cartesian theater, things present to the mind. We became materialists, or at least functionalists. We naturalized our epistemology: instead of trying to build a foundation from the materials we found in our internal worlds, we were advised to start in the middle of things, to observe how people in fact went about justifying their beliefs, and to explain their knowledge in terms of the way they interact with the things in the world that we, as theorists, find there. But the Cartesian beast is a hydra-headed creature that refuses to be slain, and that continues to color our philosophical pictures and projects. Wittgenstein, Ryle, Quine, Sellars, Davidson (not to mention Heidegger) may have cut off a few Cartesian heads, but they keep growing back. Descartes is not the bogeyman

<sup>1</sup> Quine (1960), 4.

he once was; Cartesian skeptical arguments, and arguments for the autonomy of minds and mental states are back in fashion, and philosophers feel free again to observe and contemplate the inner objects that Wittgenstein tried to banish.

The Cartesian target is of course a broad and diverse one: critics of one aspect of the picture may embrace another, and anti-Cartesians sometimes accuse each other of being closet Cartesians. (There is a cryptic and jarring remark in one of Donald Davidson's late papers about naturalized epistemology: "I do not accept Quine's account of the nature of knowledge, which is essentially first person and Cartesian."<sup>2</sup>)

Being myself still mired in the philosophical mindset of the twentieth century, my discussion of our knowledge of the internal world will be in the anti-Cartesian tradition. My subject matter will be that part of our knowledge that the Cartesian internalist takes to be most basic and unproblematic—knowledge of our own phenomenal experience and thought. But I will approach the subjective point of view from the outside. Before getting down to work on the details, I will try, in this first chapter, to set the context by making some "big picture" remarks about the way I see the contrast between a Cartesian philosophical project and an externalist alternative. I will sketch some old themes that are familiar in themselves, but that are not always recognized as playing a role in the details of some of the current debates that I will be discussing.

The contrast I have in mind is a contrast between two kinds of philosophical project, rather than two different metaphysical theses—a contrast between decisions about where to start, between different assumptions about what is unproblematic, and about how to characterize the central philosophical problems. The Cartesian internalist begins with the contents of his mind—with what he finds by introspecting and reflecting. This is what is unproblematic;

<sup>2</sup> Davidson (1991), 192.

these are the things and the facts that we know directly. The internalist's problem then is, how do we move beyond these to form a conception of an external world, and how are we able to know that the world beyond us answers to the conceptions that we form. The externalist, in contrast, proposes that we begin with the world we find ourselves in, and with what either common sense or our best scientific theories tell us about it. Among the things we find are human beings—ourselves—who are things that (it seems) can know about the world, can experience it, have a point of view on it. Our problem is to explain how our objective conception of the world can be a conception of a world that contains things like us who are able to think about and experience it in the way that we do.

The contrasting projects will formulate the central philosophical problems about knowledge and the mind in quite different ways. For the internalist, the central question about intentionality, for example, is this: how can my representational capacities extend beyond my own mental life? I can take for granted, without explanation, my capacity to represent the contents of my mind, and my capacity to reason about what I find there. At this point, there is no problem about the relation between my thought and its subject matter, since they are identical. The problem is to explain how I extend my representational reach beyond this. So the problem is a problem of explaining representational resources for a wider domain in terms of given representational resources for a narrower one. The problem is like the problem of explaining the logical and semantic relation between an observation language and a theoretical language. The externalist sees the problem of intentionality quite differently: we find in the world human beings, with a certain complex physical structure, a certain range of behavioral capacities and causal relations with their environments. What is it about those features, capacities and relations that makes it correct to describe the internal states and verbal behavior of these creatures in terms of intentional relations to propositions,

properties, and individuals? What is it for such complex physical objects to be in states that are *about* the world, and about themselves?

Internalists and externalists will each complain that the other is taking for granted what needs to be explained. The internalists see the externalist project as a project motivated by pessimism. Their complaint is this: “Because you see no hope of reasoning your way out of your internal world, you give up and simply *assume* that there is a world that answers to your inner conception. You just help yourself to some additional material, taking it for granted because you see no other way to make progress. You decide that honest toil is so ill paid that theft is the only option.” But the externalists reject this way of understanding their project. “It is not,” they insist, “that we are taking for granted what you take as given, and more besides. It is you, we think, who are taking for granted phenomena that are in need of explanation. In our view, we can make sense of your starting point—the internal world—only by locating it in a wider world. The problem, we think, is not that skepticism is unanswerable, from a purely internal point of view, even though it may be true that it is. (In fact, we argue that the problem of skepticism, seen this way, is worse than you think.) The problem is rather that skepticism about the external world has as one of its sources an uncritical acceptance, and a false conception, of our knowledge of the internal world.”

As will be clear, my sympathies are with the externalist in this debate, but my main concern will be to keep clearly in mind what perspective it is that we are taking. Problems about knowledge and the mind have usually been posed, in recent times, in a way that presupposes the externalist starting point, but Cartesian and traditional empiricist ideas that presuppose an internalist perspective continue to influence the way we think about those problems, and some of the puzzles about our knowledge of our own experience and thought may arise from equivocating between internal and external perspectives. To try to make the

contrast between the two approaches clearer, I will discuss briefly four examples of places in recent and current philosophical debates where I think a shift from internal to external perspectives has played a crucial role. I will start with a look back at Hume's problem of induction, and what he calls his skeptical solution to it. Second, I will look at a discussion by Wilfrid Sellars of contrasting ways we think about the relationship between the qualitative character of visual experience and the properties of things in the world that such experience helps us to detect. Third, I will look at the debates between direct reference theorists and descriptivists, and related debates about anti-individualism, in Tyler Burge's sense of that term. Fourth, I will review what David Lewis called Putnam's paradox, and the response to it that he, following Michael Devitt, defends. Each of the examples deserves much more discussion than I will give them here. My aim at this point is just to highlight some recurrent themes that I see in these familiar examples, themes I will explore in more detail in later chapters.

While internalists and externalists begin at different points, and formulate the central problems in different ways, both are aiming to provide a conception of the world as it is in itself. After sketching the four examples, I will conclude this chapter by considering what Bernard Williams says about how this aim should be understood.

## 1. SKEPTICAL SOLUTIONS TO SKEPTICAL DOUBTS

The classic example of a shift from an internal to an external perspective is Hume's skeptical solution to his skeptical doubts about induction. The problem of induction is first posed from the perspective of the subject: the problem is how to justify the inferences one makes from one's evidence to hypotheses about the external world, and about the future, where the available

evidence is restricted to “the present testimony of our senses and the records of our memory.”<sup>3</sup> The shift (once it is established that the problem, posed in this way, is insoluble) is to view the subjects who are in this predicament as objects in the world who are making inferences about it, and to ask how they do it, why they do it as they do, and why it is that they are as successful as they are. The skeptical solution offers a psychological theory that provides a descriptive account of the conceptual resources that these creatures (ourselves) use to form beliefs, and a causal explanation of how they acquire and use those resources. But the story is not just a descriptive one: we observe not just that these creatures are disposed to behave in certain ways, but that they have a *capacity* to find their way about, reliably, in their environment, and our external theory provides an explanation for that capacity, an explanation for the fact that the methods of inference that they use to form beliefs are reliable methods. Of course the proponent of the skeptical solution is using the very methods that he is assessing in arriving at the conclusion that the world is one that is conducive to the success of those methods, but to acknowledge this is just to acknowledge that the skeptical solution is not a solution to the skeptical problem on the internalist’s terms. The explanation for the reliability of the inferential methods used by these creatures is still a substantive one, and it is not a foregone conclusion that the procedure will result in a positive assessment. What is required is that the story the externalist tells from the middle of things, about what the world is like, be one that is in harmony with the hypothesis that he is a creature who is able to tell this story and to have good reason to believe that it is true. Even this requirement may seem to be out of reach, if one mixes the internal and external perspectives in an inappropriate way. So, for example, suppose one took the Humean external story, and the skeptical solution, to be something like this:

<sup>3</sup> Hume (1748/1977), 16.

X (the defender of the skeptical solution): “There is really no such thing as causation, so the world is like a random sequence of states, but it is a sequence that happens (by sheer chance) to have exhibited, up to now, a certain pattern of regularity, *and it will continue to do so* (still by fortuitous coincidence) so we can be confident that our inductive methods will continue to work.”

S (the internalist skeptic): “But what reason do you have to be confident that the pattern will continue?”

X: “I can’t give you a reason, but I can give you an explanation for my confidence. I am a creature of habit, and the regularity of the pattern up to now has irresistibly caused me to expect it to continue. I can’t help having this belief, and it is a good thing too, since the pattern *will continue.*”

One might, with good reason, find X’s line here to be not just unsatisfying, but incoherent, since he purports to be giving a causal explanation for a certain belief, while rejecting the applicability of causal concepts. But the real Humean does not reject causation, and emphatically affirms the central role of causal hypotheses in inductive reasoning. What is rejected is only a certain theory of causation that (according to the Humean diagnosis) tries to explain a relation between events in terms of a relation (necessary connection) that applies only to ideas. The Humean also will reject the conclusion that we can have reason, grounded only in what is available from the internal perspective, to believe any causal claims. So much the worse for the internal perspective.<sup>4</sup>

<sup>4</sup> I say “so much the worse for the internal perspective”, but I can’t claim that Hume says this. He remains, I think, profoundly ambivalent, taking his skepticism as seriously as his naturalism. There is some suggestion that he thinks it is a weakness that we (and he) are unable to stick consistently with the unmitigated skepticism that he argues for, but also a suggestion that it is a good thing that we are weak in this way.

I will leave it to the Hume scholars, who have long argued about the tensions between the naturalist and skeptical strains in Hume’s thought, to determine whether there is a stable position, faithful to the texts, that reconciles these two strains. But whether there is or not, I think it is clear that Hume’s skeptical solution makes the kind of externalist shift that I am trying to illustrate. (Thanks to Robert Fogelin for helpful discussions about Hume’s skepticism, and his so-called skeptical solution.)

## 2. VISIBLE PROPERTIES AND VISUAL EXPERIENCE

On the traditional empiricist picture, ideas of visible properties—of color properties, for example—derive from visual experience, which is then in one way or another projected onto the world. This picture can be developed in various ways, and is compatible with very different theoretical accounts of the nature of the properties that we are detecting, or at least take ourselves to be detecting, when we have visual experiences. On one view, color is a confused concept that involves attributing to things in the world properties that are really properties of our experience; on another, color is a power or a disposition to cause us to have experiences with a certain character, a power that resides in the physical objects to which we ascribe color properties; on a third view, colors are whatever the categorical properties are, the possession of which by an object in a perceiver's vicinity tend to cause her to have experiences with a certain phenomenal character. What these ways of developing the empiricist picture have in common is the assumption that our concepts of color properties are derivative from concepts of certain types of phenomenal experience.<sup>5</sup> On a contrasting externalist view, as developed for example by Wilfrid Sellars,<sup>6</sup> the ascribers of color properties begin with a naive view of an objective world, with things in it to which our most basic color concepts are applied. We don't, to begin with, have a theory about how we are able to determine the colors of things, or about the nature of the color properties that we can see that things

<sup>5</sup> The contrasting views of the nature of the color properties themselves are not tied to this empiricist thesis about the conceptual priority of a concept of color experience. One might, for example, combine a physicalist, or even a dispositionalist view of color properties with the thesis that our concepts of colors as properties of things in the world are prior to our concepts of the experiences that those properties tend to cause in us.

<sup>6</sup> Sellars (1956/1997).

have; we just learn how to tell that things are red or green, blue or yellow, and the ability that we acquire constitutes our possession of the concepts that we are applying. When we become more critical and self-conscious about the nature of our capacities to detect these properties, and of the limitations of those capacities, we theorize that our ability is explained by the fact that we are sometimes in certain internal states that tend to correlate with the presence of the property detected, and we also learn that the correlation is not perfect. As a result, we come to distinguish *being* red from merely *looking* red. The new, more sophisticated concept of looking (to one) to be red (or of there looking to be something red before one) applies when one is in the hypothesized internal state, even when the normal correlation fails to hold. On this Sellarsian, externalist picture, it is the objective properties, or our concepts of them, that have *conceptual* priority; the idea that we can be in internal states corresponding to the colors of things, and our concepts of the qualitative character of those internal states, derive from a quasi-theoretical hypothesis about our relation to those properties of visible things. But while our concepts of the qualities of our experience are derivative, the qualities themselves have a kind of *explanatory* priority: they play an essential role in the explanation of our capacity to detect, by looking, the colors of things, and an essential role in the causal explanation for our acquisition of the concepts that we are applying when we detect color properties. The internalist's mistake, according to the Sellarsian diagnosis, is to conflate the two kinds of priority, and this conflation distorts the epistemic role that something like sense contents play in our perceptual knowledge.

Quine makes the same distinction, and paints a similar picture, most explicitly in the introductory chapter of *Word and Object*. "There is every reason to inquire into the sensory or stimulatory background of ordinary talk of physical things. The mistake comes only in seeking an implicit sub-basement of conceptualization, or

of language . . . . Our ordinary language of physical things is about as basic as language gets.”<sup>7</sup>

The issues about priority that Sellars discussed remain controversial. They are complicated, not only by different ways of spelling out the relevant notions of priority, but also by different views about the nature of the relevant experiential properties. Christopher Peacocke for example, defends the apparently anti-Sellarsian thesis that experiential concepts are *definitionally* prior to our concepts of the colors of things in the world.<sup>8</sup> But he also disclaims a commitment to the consequence that possession of color concepts requires possession of a concept of experience. “All this experientialist requires for the possession of the concept of redness is a certain pattern of sensitivity in the subject’s judgements to the occurrence of red’ experiences” (where “red’” ascribes the relevant experiential property).<sup>9</sup> This sounds like a causal, rather than a definitional dependence, and it might be a commitment that Sellars would have accepted. But Peacocke’s priority thesis, as I understand it, does have the consequence that one whose normal way of detecting the property *red* was by having an experience with a different qualitative character (as in the notorious inverted spectrum case) would thereby have a different concept of the property. In this sense, the concept essentially involves a certain type of experience, according to Peacocke’s priority thesis.

But what exactly is this experiential property, *red’*? According to *intentionalists* or *representationists*, the phenomenal character of experience is to be explained in terms of the intentional content of experience—the way an experience represents things to be.<sup>10</sup> Peacocke’s priority thesis is tied to a rejection of intentionalism, and the assumption that experiences have an intrinsic qualitative

<sup>7</sup> Quine (1960), 3.

<sup>8</sup> Peacocke (1984).

<sup>9</sup> *Ibid.*, 59.

<sup>10</sup> Intentionalism can be spelled out in different ways. For a defense of one of them, see Byrne (2001). ‘Representationism’ is Ned Block’s term. He characterizes and criticizes it in Block (2003).

character that is prior to any representational role that experience may play.<sup>11</sup>

The externalist story, as told by Sellars and Quine, does not imply that the qualitative properties of experience are representational properties, but it does imply that our conceptions of those properties are derivative from their representational role. First comes the naive capacity to detect, and then a proto-theoretical account of representation (not a general account of what it is to represent, but just a recognition of a difference between the way things are and the way they seem to be, and a recognition of a difference between something represented and something in oneself that is doing the representing). The theorist of the mind hypothesizes that there are these internal properties—qualia—that explain our capacity for visual detection. So according to this story, our recognition of qualia derives from our recognition that we are representing in a particular way.

### 3. DESCRIPTIVISM AND THE CAUSAL THEORY OF REFERENCE

The received view of reference that Saul Kripke criticized in *Naming and Necessity* has its origins in an internalist picture of representation, and even though at least some of the post-Kripkean neo-descriptivists would disclaim any allegiance to a Cartesian project, I think that intuitions from that project play a role in motivating defenses of this account of reference, and that it is useful to see the parallel between the Kripkean critique and the kind of externalist project promoted by Sellars and Quine.

Reference to individual concrete things, such as human beings, is particularly problematic, from an internalist point of view, since

<sup>11</sup> Though Peacocke explains the primed properties such as *red'* as properties of a visual field, and I would have thought that a visual field is a feature of an essentially representational mental structure.

such objects are paradigm cases of things that are not denizens of the internal world, and so not things to which we might have direct access from the inside. The descriptivist strategy is to explain the capacity to refer to concrete individuals in terms of a capacity to refer to the properties and relations that are exemplified by such individuals, things that might more plausibly be thought of as internal to the mind, or at least as things that the mind could grasp from the inside. Of course Frege was clear that the contents of thought are not themselves mental objects—they are something more abstract that can be objects of the thoughts of different thinkers—but he still seems to have assumed that the contents of speech and thought must be, in some sense, internal to the mind. Frege was famously incredulous at the idea that physical objects like Mt. Blanc (with all its snowfields) might be constituents of a proposition. Russell disagreed, holding to the view that propositions might indeed have physical objects as components. But in the end Russell took the bite out of this externalist doctrine by combining it with the view that propositions could be grasped only by someone who was *acquainted* with all of their constituents, where acquaintance required the kind of perfect and complete knowledge that we could have only of mental objects or of universals. There are propositions with Mt. Blanc as a component, and we can describe such propositions, but they cannot be the contents of what we are saying or thinking when we talk or think about Mt. Blanc. So while Frege and Russell had different conceptions of a proposition, if we restrict ourselves to propositions that are candidates for the contents of speech and thought, then both of these founding fathers of the received view of reference will agree that singular reference to physical objects must be mediated by general concepts that apply to those objects.

Kripke's externalist critique begins with arguments against the descriptive adequacy of the descriptivist project: in some cases that seem, intuitively, to be examples of successful reference, the speakers lack the conceptual resources that the analysis requires

them to have; in other cases, it was argued that the analysis implied the intuitively wrong conclusion about what the referent is. A second part of the critique argues that even if a descriptive analysis were correct, it could not provide a satisfactory account of reference without an explanation of how we are able to refer to, or to express, the properties and relations that are expressed in the descriptions that constitute the analysis. What is questioned here is the internalist presupposition that our intentional relations to properties and relations are unproblematic. A descriptivist analysis just passes the buck from one kind of expression to another. This point was supported by the arguments that Tyler Burge gave against what he called individualism. If general terms, along with names and other singular referring expressions, depend for their semantic values on environmental conditions, then our intentional relations to them cannot have the kind of foundational status that the internalist project requires. Speakers and thinkers cannot have the kind of “perfect and complete” acquaintance with properties and relations that is necessary (according to the internalist) to grasp the propositions expressed in the descriptivist analyses, and so further reduction is required for the success of the internalist project. Here it is important that the anti-individualist arguments apply to a wide range of general concepts—not just to a few natural kind terms and theory-laden scientific terms, but even to purely qualitative predicates. If only a relatively narrow range of terms and concepts are “twin-earthable” (to use David Chalmers’s term), then there might be a prospect of a reduction of the concepts that are in this narrow range to those that are not. But the externalist argues that the phenomenon brought out by the anti-individualist thought experiments is ubiquitous. There is no foundation. We need an explanation of another kind.

At this point, the externalist makes a distinction that parallels the distinction made by Quine and Sellars between conceptual and explanatory priority. Singular reference with a proper name is *conceptually* direct, but that should not be taken to imply that

there is no explanatory story to be told about what it is in virtue of which a name refers. Just as it is a mistake to confuse explanatory with conceptual priority in the case of visible properties and visual experience, so it is a mistake (according the causal theorist of reference) to confuse an explanation for the fact that a name refers as it does with a conceptual analysis of what is expressed by that name. A definite description of an individual named might play an essential role in the explanation for the fact that the name refers to that individual even if the propositions expressed with the name are determined as a function of the individual itself, and not of some concept expressed by the description. Kripke took Frege's notion of sense to involve an equivocation between these two roles of a descriptive concept in the explanation of the relation between a name and its referent.<sup>12</sup>

Despite the influential critique by Kripke and others, the descriptivist program remains alive. "Description theories of reference are supposed to have been well and truly refuted," David Lewis wrote in 1984. "I think not: we have learned enough from our attackers to withstand their attacks."<sup>13</sup> Lewis was sensitive to the distinction between a conceptual or semantic role for a description and an explanatory, or metasemantic role, and he acknowledged that a causal descriptivist analysis—one that builds the description of the causal process by which the reference of a term is determined into the semantics for the term—just passes the buck to the terms used in the description. He nevertheless argued that such an analysis was defensible, and preferable to an account that located the causal story in the external account of the facts in virtue of which thought and talk has the content that it has.

<sup>12</sup> Kripke (1972), 59: "Frege should be criticized for using the term 'sense' in two senses. For he takes the sense of a designator to be its meaning; and he also takes it to be the way its referent is determined. Identifying the two, he supposes that both are given by definite descriptions."

<sup>13</sup> Lewis (1984), 60.

Even though Lewis wanted to defend what is, in a sense, an internalist project, he accepted the externalist's formulation of the problem of intentionality, and he argued that any solution to it will require a move that, I will suggest in the next section, parallels the move in Hume's skeptical solution to the problem of induction.

#### 4. PUTNAM'S PARADOX AND ITS SKEPTICAL SOLUTION

Lewis's externalist shift, like Hume's, is a response to a skeptical problem that is posed from the subject's point of view. The problem is what Lewis calls Putnam's paradox, an argument that Hilary Putnam posed first in 1977.<sup>14</sup> The rough idea is this: Start with the fact that any consistent theory has many interpretations according to which it is true. All that needs to be assumed for this result is that there are enough things in the world; nothing need be assumed about what those things are like. But actual theorists claim more than that their theories are true on some interpretation or other: they *intend* a certain interpretation, and the claim is that the theory propounded is true on that interpretation. What Putnam's skeptical argument challenges is the assumption that this provides any constraint at all on interpretation. For I might formulate my referential intentions (in my public language, or in my language of thought), and add them to my total theory, and the resulting augmented theory, incorporating statements expressing all of my referential intentions, will still be true on many interpretations, no matter what the world is like. The point applies quite generally: suppose that there is some condition C that we might propose as a constraint on admissible interpretations of our language (or on whatever the objects or events are that represent our thoughts). C itself could be incorporated into one's theory, and

<sup>14</sup> Putnam (1977).

the argument applied to the resulting theory. “Constraint C is to be imposed by accepting C-theory, according to Putnam. But C-theory is just more theory, more grist for the mill, and more theory will go the way of all theory.”<sup>15</sup> The point is that all that any such constraints can do is to restrict the range of consistent theories that are candidates to represent a subject’s corpus of beliefs. But since any such theory will be true, on many interpretations, the restrictions do not help to constrain the content of the claim that the theory makes about the world.

But Lewis replies: “C is *not* to be imposed just by accepting C-theory. That is a misunderstanding of what C is. The constraint is *not* that an intended interpretation must somehow make our account of C come out true. The constraint is that an intended interpretation must confirm to C itself.”<sup>16</sup> The constraint is imposed, not on oneself from within, but on the objects we find in the world, who are in fact ourselves.

Like Hume’s skeptical solution, this response to Putnam’s paradox does not answer the internalist skeptic on his own terms. The conclusion of Putnam’s argument is that all reference is radically indeterminate, and Lewis’s strategy can succeed in stating a determinate condition only if this conclusion is false, so the response might be thought to beg the question. Lewis does not take this worry very seriously: who gave the skeptic the license to set the terms of the debate? But he takes more seriously what he describes as “a deeper and better reason to say that any proposed constraint is just more theory.”<sup>17</sup> He thinks that it is tempting to believe, of whatever theory of reference is correct, that “somehow, implicitly or explicitly, individually or collectively, we have made this theory of reference true by stipulation.” And he thinks that if this tempting belief were accepted, Putnam’s conclusion would be unavoidable. “The main lesson of Putnam’s Paradox,” Lewis

<sup>15</sup> Lewis (1984), 62.

<sup>16</sup> *Ibid.* See also Devitt (1983), which Lewis cites in this context.

<sup>17</sup> *Ibid.*, 63.

writes, “is that this purely voluntaristic view of reference leads to disaster.” But I think this is a misleading diagnosis. We don’t need Putnam’s paradox to see that any general solution to the problem of intentionality that tried to be purely voluntaristic would be incoherent. An intention is an intentional state, and a stipulation is an intentional act whose content is determined by the content of the intention with which it is performed. One obviously cannot explain what makes an intention that is directed at Osama bin Laden be an intention that is directed at *him* in terms of the agent’s intention that it should be directed at him. A purely voluntaristic theory of reference makes sense only as a theory that aims to explain *linguistic* intentionality in terms of the intentionality of thought, and a project of this kind (Grice’s project, for example) is untouched by Putnam’s paradox. I think the main lesson of Putnam’s argument should instead be put this way: a formulation of the problem of intentionality as a problem for the subject of the intentional states (“how should *I* establish a connection between my thoughts and what they are to be about”) is hopeless. A clear view of the problem requires that we distinguish, conceptually, between (1) ourselves as theorists attempting to explain our intentional relations to things in our environment and (2) ourselves as the objects whose relations to things in their environments we are studying. But as in the case of Hume’s skeptical solution, our two views of ourselves must be in harmony: a satisfactory account must explain how it is possible for us, as objects in the world, to be the kind of thing that can have a theory of the kind that we, as theorists, have, and it must explain how such theories can succeed in saying things about the world.<sup>18</sup>

Each of these four examples involves a dialectical shift from the subject’s perspective to the perspective of a theorist. A problem is formulated, or reformulated, as a problem about the relations

<sup>18</sup> Putnam’s paradox is often compared with the skeptical puzzle about rule following that is posed by Kripke’s Wittgenstein. And Kripke does refer to Wittgenstein’s solution as a “skeptical solution”. See Kripke (1982).

between beings found in the world who are only accidentally the same as the ones who are posing the problem. The questions are not, how should *we* justify our inductive practices, or bring it about that our thoughts extend beyond our minds, or that our words attach to things in the world, but how are *their* capacities to learn about the world, or to talk or think about it, to be explained, where *they* are of particular interest because they happen to be us. But since they *are* us, the shift, in each case, raises potential problems about circularity. Responding to these problems requires distinguishing between different kinds of priority, and imposes a demand that the theorists' explanations of the cognitive and epistemic capacities of their objects of study be in harmony with the fact that they themselves are able to give the kind of explanation that they are giving.

## 5. THE ABSOLUTE CONCEPTION OF REALITY

It is tempting to think of this external standpoint as a view of reality from outside, or from above. We retreat into our objective selves, leaving behind our empirical selves, and take on the view from nowhere.<sup>19</sup> This image is reinforced by the language of perspective: the external standpoint seems to be a perspectiveless perspective. But in a way this image gets things exactly backward. It is essential to the view from the middle of things that there is no place from which to observe and reflect on the world other than our place within it. It is essential that the theorist viewing himself as an object in the world is the same as the object being viewed. It is not that we are looking for a platform outside of the world on which to build our conception of it; instead, we are trying to do without foundations at all.

<sup>19</sup> Cf. Nagel (1986).

Both internalists and externalists are aiming at a conception of a reality that exists independently of our conception and knowledge of it. They differ about whether such a conception can be built from within, and perhaps also about what such a conception requires. Bernard Williams, who was concerned with Descartes's project of generating such a conception, suggested that there is something puzzling and problematic about an absolute conception of reality. Here is the problem, as he saw it:

Suppose A and B each claim to have some knowledge of the world. Each has some beliefs and moreover has experiences of the world, and ways of conceptualizing it, which have given rise to those beliefs and are expressed in them: let us call all of this together his *representation* of the world (or part of the world). Now . . . A's and B's representations may well differ. If what they both have is knowledge, then it seems to follow that there must be some coherent way of understanding why those representations differ, and how they are related to one another.

We need, that is, to understand how the different representations "can each be perspectives on the same reality." This requires one to form a conception of the world which *contains* A and B and their representations; . . . but this will still itself be a representation, involving its own beliefs, conceptualizations, perceptual experiences and assumptions about the laws of nature. If this is knowledge, then we must be able to form the conception, once more, of how this would be related to some other representation which might, equally, claim to be knowledge; indeed, we must be able to form that conception with regard to *every* other representation which might make that claim.

But the idea that there might be such a conception, Williams goes on to argue, poses a dilemma:

On the one hand, the absolute conception might be regarded as entirely empty, specified only as 'whatever it is that these representations represent'. In this case, it no longer does the work that was expected of it . . . . On the other hand, we may have some determinate picture of what the world is like independent of any knowledge or representation in thought; but then that is open to the reflection, once more, that that is only one

particular representation of it, our own, and that we have no independent point of leverage for raising this into the absolute representation of reality.<sup>20</sup>

The first step in defusing this dilemma is to distinguish the content of a representation both from the particular means used to express that conception, and from the act of expressing it. The absoluteness we are looking for is in the content: we want a representation of the world as it is in itself (or as Williams puts it, “of what is there *anyway*”) and not just of the world as it appears from a certain perspective. But of course any representation of the world as it is in itself will use certain means to say that the world is that way, and the saying of it will take place at a certain time and place in the world.

Suppose I am A, forming a conception of the world as it is in itself. It is part of the content of my conception that there are conceivers forming conceptions of the world (as it is in itself), and that those conceptions are formed from a particular point of view within the world. If my conception is correct and reasonably inclusive, then among those conceivers will be someone who is me (A), and someone else who is B. My account will recognize that A and B are conceiving of the world from different perspectives, and will include an account of how those perspectives differ. But since the particular conceptions being formed by A and B that we are interested in are conceptions of the world as it is in itself, it will not be part of the *content* of A’s conception that it is A who is forming that conception (though it will be part of A’s conception that A is, at a certain time and place, forming *a* conception with that content). It could be that A and B form exactly the same conception of the world as it is in itself. In this case, there will be distinct acts of conceiving, each a conception formed from a certain point of view, but they will have the same content.

<sup>20</sup> Williams (1978), 49, 50.

Now, as Williams suggests, it may be that the conceptions of A and B are *not* the same, in either form or content, even if both are correct (both count as knowledge), and the two conceptions may differ even if both are correct representations of the world as it is in itself. Each may tell only a part of the story, or they may, as Williams suggests, tell the story in different but equivalent ways.

Some things Williams says suggest that the absolute conception he is looking for must be *comprehensive*, incorporating all possible representations of the world. It is not entirely clear what this would mean, and I don't think that a conception of the world as it is in itself requires that completeness is achievable, or even intelligible, but suppose we can make sense of the idea. Consider a possible world that contains a representation of itself that says enough so that anything else that might be said would be redundant. There is, in this world, a book (with very small print) in A's library. Since the book is complete, it will tell us that there is a book on the shelf of A's library that tells the complete story, and it must also tell us exactly what the book says. One might be tempted to imagine an infinite regress here, like a picture of a room that has a picture of the room on the wall, and so of smaller and smaller pictures nested within one another. But self-representation need not require this regress. It is easy for a book to tell us, among other things, what the book itself says. At the appropriate point, the book might say: "On the third shelf of A's library, there is a book that contains the following text: (now turn to the top of page 1 of *this* book, and read through to the end; then return to this point, to finish the story of what else there is in the world)". Is this a cheat? Does A's book really give us the complete story? Well, imagine a description of this possible universe that is not in the universe at all. A's world is, after all, a mere counterfactual possibility. Suppose *we* have a complete description of A's counterfactual world. Our book is just like the book in A's library, except that at the appropriate place it puts A's whole book in quotation marks in place of the parenthetical

remark. No circularity here, or hint of a cheat. *Our* story of this counterfactual world might be complete, whether or not A's story is. But isn't the *content* of our book the same as the content of the book in A's library? (If so, we could save a lot of trees by using A's more efficient method of telling the story.)

Of course the story A tells, even if comprehensive, will be told in a particular language. If B's story is also comprehensive, it will be equivalent to A's story—identical in content—but it might still tell the story in a different way. We should resist the temptation to think that it detracts from the absoluteness of the content of a representation if the representation doesn't present a pure proposition, detached from any means of expression. The search for a representation, freed from any means of representation, will face a dilemma that parallels the one posed by Williams for the absolute conception of reality. Paraphrasing Williams: what a statement says (the proposition it expresses) must be independent of any linguistic item that expresses it. But here we face a dilemma: either the pure proposition is entirely empty, specified only as "whatever it is that these linguistic items (in Russian, English, etc.) express". In this case, it no longer does the work that was expected of it. On the other hand, we may have some determinate way of saying what the statement says, but then it is open to the reflection that our characterization of the proposition, once more, is only one linguistic representation of it, and again we have no independent point of leverage for raising it into a pure proposition.

I trust that no one will take this dilemma seriously, in this bald form, but there are real problems in the vicinity. It is a recurrent problem, in all of the attempts to view the philosophical terrain from the middle, that we "have no independent point of leverage". We want to theorize about the relation between representations and their content, but of course we can do so only by using other representations. We need a conceptual distinction between the content of a representation and the vehicle in which that content rides, but there may be more than one way to make the distinction,

and controversies about how to make it can interact with substantive issues about the subject matter that is represented. It is sometimes frustrating to have to start in the middle, but that is where we are.