

DAVID H. FINKELSTEIN

---

# Expression and the Inner

HARVARD UNIVERSITY PRESS  
Cambridge, Massachusetts, and London, England  
2003

Copyright © 2003 by the President and Fellows of Harvard College  
All rights reserved  
Printed in the United States of America

*Library of Congress Cataloging-in-Publication Data*

Finkelstein, David H.

Expression and the inner / David H. Finkelstein.  
p. cm.

Includes bibliographical references and index.

ISBN 0-674-01156-2

1. Self-knowledge, theory of. 2. Self-presentation. I. Title.

BD438.5.F525 2003

128'.2—dc21 2003056643

---

## Contents

Acknowledgments

vii

Introduction

1

### I. DETECTIVISM AND CONSTITUTIVISM

#### 1 Detectivism

9

1.1. Old Detectivism 10

1.2. New Detectivism 15

1.3. A Dialogue 19

#### 2 Constitutivism

28

2.1. "A Kind of *Decision*" 29

2.2. Interpretation and Stipulation 40

2.3. The Responsibility Objection 45

#### 3 Between Detectivism and Constitutivism

51

3.1. Experience and the Logical Space of Reasons 53

3.2. The Middle Path 58

3.3. When a Dog Feels Pain 64

3.4. The Phantom Smell Objection 65

3.5. Back to Detectivism? 71

### II. EXPRESSION

#### 4 Meaning, Expression, and Expressivism

75

4.1. Meaning 75

## Detectivism

If you want to know what I think, feel, imagine, or intend, I am a good person—indeed, usually the best person—to ask. It is sometimes said that I enjoy a kind of authority when I talk about what, broadly speaking, might be called my own states of mind—when I say, e.g., “My head hurts,” “I was worried about you,” or “I intend to arrive early.” When people don’t accept my mental state self-ascriptions at face value, it is generally because they take me to be insincere rather than mistaken.

It is easy to be struck by such facts and, once struck, to find oneself inclined to ask questions about them: How is it that we are able to state our own thoughts and feelings so easily, accurately, and authoritatively? How is describing one’s own state of mind different from describing someone else’s? This chapter concerns a natural and widespread approach to answering such questions—an approach that I call “detectivism.” A detectivist is someone who believes that a person’s ability to speak about her own states of mind as easily, accurately, and authoritatively as she does may be explained by appeal to a process by which she *finds out about them*. According to detectivism, I am able to state my own thoughts and feelings because they are conscious, and they’re conscious thanks to a cognitive process by which I have detected their presence. If I’m in an especially good position to say what’s in *my* mind, it is because this cognitive process provides me with better epistemic access to my mental states than other people have to them.

In this chapter, I shall try to persuade you that no form of detec-

tivism provides us with the materials we need in order to understand first-person authority. In what immediately follows, I'll trace some of detectivism's history, focusing in particular on two detectivist positions set out by Bertrand Russell in 1912 and in 1921. These positions of Russell's will serve as representative versions of what I shall call "old" and "new" detectivism.

### 1.1. Old Detectivism

Most detectivists have thought that we can say what's on our minds thanks to a kind of inward observation or perception. Auguste Comte argues that such a view presupposes an incoherent conception of the mind's relation to its own states. In his *Cours de Philosophie Positive*, Comte ridicules the experimental psychologists of his day for thinking that human reasoning could be studied by inward observation:

[A]s for examining . . . intellectual phenomena as they appear, this is manifestly impossible. A thinking individual cannot divide himself into two, one half reasoning, and the other watching it reason. The observed and the observing organ become identical in this case. How could observation take place?

The so-called psychological method therefore is in principle invalid. And consider to what absolutely contradictory antics it leads! On the one hand you are told to insulate yourself, as much as possible, from every external sensation, above all you must refrain from intellectual work; for if you were to do the simplest sum, what would become of *internal* observation? On the other hand, after having by dint of precautions attained the state of intellectual sleep, you must busy yourself contemplating the operations taking place in your mind, when nothing at all is taking place in it! Our posterity will doubtless one day see these pretensions transferred to the comic stage. (Comte 1974, 32)

John Stuart Mill replies to Comte on behalf of introspective psychology:

In the first place, M. Comte might be referred to experience, and to the writings of his countryman M. Cardaillac and our own Sir William Hamilton for proof that the mind can not only be conscious of, but attend to, more than one, and even a considerable number of impressions at once . . . Secondly, it might have occurred to M. Comte that a

fact may be studied through the medium of memory, not at the very moment of our perceiving it, but the moment after: and this is really the mode in which our best knowledge of our intellectual acts is generally acquired. We reflect on what we have been doing when the act is past, but when its impression in the memory is still fresh. Unless in one of these ways, we could not have acquired the knowledge, which nobody denies us to have, of what passes in our own minds . . . Whatever we are directly aware of, we can directly observe. (Mill 1961, 63–64).

Mill's point is not merely that we perceive our own reasonings on odd occasions, e.g., when asked to by psychologists. It is rather that our ordinary, moment-by-moment awareness of what "passes in our own minds" is due to a kind of inward perception (plus memory). A detectivist thinks that our ordinary consciousness of, at least, some significant range of mental states is explained by the fact that we are able to perceive—or anyway, detect—their presence. Mill's reply to Comte is cited approvingly by many of the leading lights in the succeeding generation of philosophers, including Franz Brentano and William James, both of whom defend related positions.<sup>1</sup>

For all of its radical critique of previous philosophy, early analytic philosophy mostly adopts a broadly detectivist conception of a subject's relations to his own inner goings-on. Within the development of analytic philosophy, it is important to distinguish two distinct forms of detectivism, each of which (like so much else in analytic philosophy) can be traced back to a phase in the thought of Bertrand Russell.

In 1912, in *The Problems of Philosophy*, Russell distinguishes between our knowledge of physical objects and our knowledge of the sense-data that make up their appearances. He describes the latter as "things with which I have acquaintance, things immediately known to me just as they are" (Russell 1912, 47). By contrast, when I see or touch a physical object, I am aware of its existence only indirectly—only insofar as I have inferred that such a thing would be the most likely cause of my current visual or tactile sense-data.

According to Russell, we know not only our sense-data but also our own mental goings-on by acquaintance.<sup>2</sup> He writes:

1. See Brentano 1973, 35; James 1890, 189.

2. According to Russell's sense-datum theory, sense-data are not mental—or anyway are "not known to be mental" (Russell 1992, 5; also see Russell 1912, 38–43).

When I desire food, I may be aware of my desire for food; thus 'my desiring food' is an object with which I am acquainted. Similarly we may be aware of our feeling pleasure or pain, and generally of the events which happen in our minds. This kind of acquaintance, which may be called self-consciousness, is the source of all our knowledge of mental things. (1912, 49)

A couple of pages later, Russell introduces the expression "inner sense." He says, "We have acquaintance in sensation with the data of the outer senses, and in introspection with the data of what may be called the inner sense—thoughts, feelings, desires, etc." (1912, 51). A guiding thought of almost all detectivists is that there is an important parallel between my relation to the physical objects before my eyes and my relation to my own states of mind. The expression "inner sense" is meant to capture this parallel. Russell, however, also takes there to be significant disanalogies between the inner and outer senses. (It would be misleading to say, "Russell thinks of self-knowledge as involving a sixth sense—a faculty just like seeing and hearing, only aimed in a different direction.") Here is one of the disanalogies: while inner and outer senses have acquaintance in common, acquaintance is all there is to inner cognition, while it is supplemented by inference in outer perception. If we call physical objects the *termini* of the outer senses—because the purpose of the outer senses is to inform us about physical objects—then we can say that according to Russell, the *termini* of, say, vision are not its immediate objects. Vision's *termini*, e.g., the trees and cars that I see out my window, are explanatory posits. By contrast, the *termini* of the inner sense—thoughts, feelings, and attitudes—are its immediate objects and are experienced directly.

A second, related respect in which inner and outer cognition differ according to Russell lies in the fact that we know our own mental states *better* than we know external objects, with greater—indeed, complete—certainty: "[t]he certainty of our knowledge of our own experiences does not have to be limited in any way to allow for exceptional cases" (1912, 19). The inner sense provides us with perfect knowledge of inner objects. No outer sense yields this kind of access to physical objects.

So while Russell's talk of *inner sense* suggests that we know our own minds by a kind of perceptual mechanism, we should appreciate how different inwardly detecting our own desires is from seeing or hearing

physical objects. We might say that for Russell at this stage, the mechanism by which we know our own minds is *perceptual only* in an extended or metaphorical sense.<sup>3</sup> According to this kind of detectivism, introspecting is far more different from seeing, hearing, tasting, smelling, or feeling than these are from one another. I'll call this kind of detectivism, according to which we know our own minds via an, as it were, *sui generis* perceptual process, "old detectivism." The old detectivist claims that this process provides me, and only me, with a kind of access to my mental states that is complete, infallible, and more direct than my access to the keyboard on which I'm presently typing.

\* \* \*

Old detectivism makes room for a thought that may seem intuitively compelling, viz., that when I say, "I'm in pain," or, "I want a sandwich," something is going on that is fundamentally different from what occurs when I describe another person's state of mind or when I describe the state of my office. However appealing this thought may be, old detectivism accommodates it at too great a cost. The old detectivist's conception of the mind is subject to a pair of related and mutually sustaining pressures: he is forced to picture the mind (1) as an immaterial organ and (2) as epistemically cut off from goings-on outside it. I'll say just a little bit about each of these pressures in order to make intelligible the transition that we find in the work of Russell and others from old detectivism to what I'll be calling "new detectivism." (In so doing, I'll be leading you over a terrain whose features may be familiar to you. We should cover this ground, if only briefly, in order to appreciate how it is that new detectivism has come to appear superior to old detectivism. Soon, I shall argue that this appearance is illusory.)

(1) The old detectivist is under pressure to picture the mind as an immaterial organ for a number of reasons. One is that the mechanism of inner sense that he postulates is supernaturally reliable. I might say of a particularly obvious barn, "You can't miss it"; but a barn, no mat-

3. Russell's extension of the language of perception to relations of acquaintance is typical of sense-datum theorists. H. H. Price, for example, writes, "Data of this special sort are called *sense-data*. And the acquaintance with them is conveniently called *sensing*" (Price 1932, 3). (Price goes on to compare sense-data with "the data of the detective," saying that the detective's data are unlike sense-data in that they are "really the result of inference" [1932, 4].)

ter how large and obvious, might be missed by someone who was distracted or whose eyes were not functioning properly. According to old detectivism, mental items are such that they can *never* go unnoticed by the subject to whom they appear. Unlike eyes or ears, the old detectivist's inner sense *cannot* break down or fail to function. Could such a mechanism be made of mere matter?<sup>4</sup>

(2) One source of dissatisfaction with old detectivism lies in its association with dualism. But just as significant is the fact that old detectivism tends toward scepticism about the external world and about other minds. In *The Problems of Philosophy*, Russell claims that we may justifiably infer that there are external objects because this hypothesis provides the *simplest* way to account for our sense-data: "[E]very principle of simplicity urges us to adopt the natural view, that there really are objects other than ourselves and our sense-data which have an existence not dependent upon our perceiving them" (1912, 24).<sup>5</sup> How should we assess this claim? Consider my present situation—one that I might be pretheoretically inclined to describe as follows: "I am looking at the lamp next to my computer." According to Russell, I have posited the existence of an object (a lamp) external to my present sense-data in order to account for some of them. Is this the simplest way to understand my experience? Why is it simpler to hypothesize that the sense-data in question are caused by one object rather than by two or by many? Am I supposed to be moved by ontological parsimony at this point? Wouldn't the most parsimonious course be to refrain entirely from positing entities that lie outside my experience?<sup>6</sup>

Knowledge of other minds presents the old detectivist with further problems. According to Russell, I learn about the mind of another via *two* problematic inferences. First, I must establish that he—or rather,

4. Lycan (1996) points to the connection between super-reliability and immateriality when he writes (on behalf of new detectivism), "[S]ince an internal monitor is a physical device, and so subject to malfunction, it might report falsely" (17).

5. Russell loses faith in this sort of appeal to simplicity shortly after writing *The Problems of Philosophy*. In unpublished writing from 1912, he admits that we would have grounds for thinking that "of two hypotheses which both fit the data, the simpler is more likely to be true" only if we knew that "the universe had been created for the purpose of delighting mathematicians" (quoted on p. 386 of Hilton 1990).

6. In a paper written in 1914 called "The relation of sense-data to physics," Russell comes around to the view that we ought to refrain from positing objects beyond the sense-data. He writes: "[I]nferred entities should, whenever this can be done, be similar to those whose existence is given, rather than, like the Kantian *Ding an sich*, some-

his body—exists. This is a matter of positing an external cause (the "simplest" explanation) of some of my sense-data. Second, on the basis of analogical reasoning, I infer that the movements of his body are accompanied by thoughts and feelings of the sort I have when my body moves in similar ways. Even a sense-datum theorist who was not terribly bothered by the first inference might find it difficult to countenance the second.<sup>7</sup>

## 1.2. New Detectivism

Eventually, in the face of such difficulties, Russell's detectivism changes dramatically. To see how this goes, it will help to consider the analysis of desire that he puts forward in 1921, in *The Analysis of Mind*. What a desire is, Russell says there, is any mental occurrence that involves "discomfort"—where a mental occurrence's involving discomfort consists in the fact that it causes an organism to engage in "movements tending to produce some more or less definite change involving the cessation of the occurrence" (Russell 1921, 71; see p. 75 for a summary of the view). Hunger, for example, is a feeling that causes an organism to engage in restless movements that are likely to lead to its eating. That hunger is a desire to *eat* consists in the fact that upon eating, the restless activity ceases. "The thing which will bring a restless condition to an end is said to be what is desired" (1921, 32).<sup>8</sup>

thing wholly remote from the data which nominally support the inference" (Russell 1917, 116). Rather than infer the existence of physical objects wholly unlike sense-data, Russell here suggests what he takes to be a more conservative inference. We should infer that the items we call sense-data sometimes exist even when they are not present to a mind. He uses the term "sensibilia" to refer to these items without presupposing that they are data for a mind. (Sense-data are sensibilia that happen to be present to a mind.) Having concluded that our immediate acquaintance is with items that may exist unbeknownst to anyone, Russell argues that the physical world is made up of such items. Sensibilia (and so sense-data) are, he thinks, physical entities; indeed, they constitute the subject matter of physics.

7. Russell himself expresses doubt about the argument from analogy in his *Theory of Knowledge: The 1913 Manuscript*, when he suggests that we treat the existence of other minds as a "working hypothesis" (Russell 1992, 14).

8. Wittgenstein makes fun of this analysis of desire: "If I wanted to eat an apple, and someone punched me in the stomach, taking away my appetite, then it was this punch that I originally wanted" (PR, 64; see also PI§440). Later versions of new detectivism prove to be less vulnerable to one-line refutations. (To be fair to Russell, he does try to answer what amounts to this criticism on p. 64 of *The Analysis of Mind*.)

Given this story about what desires are, how do we know our own? Russell writes:

"Conscious" desire . . . consists of desire in the sense hitherto discussed, together with a true belief as to its "purpose," i.e. as to the state of affairs that will bring quiescence with cessation of the discomfort. If our theory of desire is correct, belief as to its purpose may very well be erroneous, since only experience can show what causes a discomfort to cease. (1921, 72)

I believe that the discovery of our own motives can only be made by the same process by which we discover other people's, namely, the process of observing our actions and inferring the desire which could prompt them. A desire is "conscious" when we have told ourselves that we have it. (1921, 31)

I know what I desire by making an inference—based on past experience with my behavior—concerning what it is that will bring my restless activity to an end. As Russell says, this inference may well be erroneous. Indeed, "Our spontaneous, unsophisticated beliefs, whether as to ourselves or as to the outer world, are always extremely rash and very liable to error. The acquisition of caution is equally necessary and equally difficult in both directions" (1921, 122). What Russell commits himself to is not the relatively banal claim that we *occasionally* learn what we desire by making a fallible inference from our own behavior. It is, rather, that our ordinary awareness of our own desires has this character. According to Russell, whenever I have a conscious desire, I have made an inference about myself. Compare this position with his earlier account of self-knowledge. In 1912, in *The Problems of Philosophy*, his view had been that I know my desires thanks to an inner sense that provides me with a kind of access to them that is complete, infallible, and more direct than my access to any external object. Now, in 1921, Russell claims that I know my desires by the same sort of fallible theorizing that allows me to learn the desires of others.

I call this sort of position—one according to which a mental state is conscious by virtue of an awareness of it that is roughly on a par with the sort of awareness one has to external states and events—"new detectivism." In *The Analysis of Mind*, Russell takes our access to our own desires to be on a par with our access to external states and events that we learn about *via inference*. This makes him an *atypical* new detec-

tivist. Most new detectivists take our awareness of our own states of mind to be roughly on a par with our access to external states and events that we learn about perceptually. They claim that we find out about our own mental states via a kind of inward observation or perception—a very different *sort* of inward observation or perception, however, from that invoked by old detectivists. Recall that according to old detectivism, we learn about our own states of mind by a process that can be called "perceptual" only in an extended or metaphorical sense. An old detectivist's inner sense is a supernaturally reliable detection mechanism that provides a person with a kind of access to mental states that is more direct and certain than the sort of access that seeing or hearing could ever provide. New detectivists typically seek to domesticate this story—to render it less Cartesian and more naturalistic—by construing our awareness of our own mental states as involving a species of ordinary, garden-variety perception. Like Russell in 1912, new detectivists sometimes speak of an "inner sense" (or "internal sense" or "inner eye") that allows us to perceive what's going on in our own minds. Unlike Russell's employment of this expression, a new detectivist's use of it should be understood, as it were, literally (or, at least, much less metaphorically)—as referring to a faculty that is on all fours with seeing and hearing, only directed at states and events that happen to be inside the mind. Thus, David Armstrong writes, "By sense-perception we become aware of current physical happenings in our environment and our body. By inner sense we become aware of current happenings in our own mind" (Armstrong 1968, 95); William Lycan defends what he calls the "inner-sense theory" according to which "conscious awareness is the successful operation of an internal scanner or monitor that outputs second-order representations of first-order psychological states" (Lycan 1996, 31); and D. H. Mellor writes, "I can know so much about my own beliefs because I happen to have an internal sense that informs me of them" (Mellor 1977–78, 100). Paul Churchland provides a nice statement of this sort of perceptual new detectivism:

[S]elf-consciousness . . . is just a species of perception: *self-perception*. It is not perception of one's foot with one's eyes, for example, but is rather the perception of one's internal states with what we may call (largely in ignorance) one's faculty of introspection. Self-consciousness is thus no more (and no less) mysterious than perception gener-



ally. It is just directed internally rather than externally. (Churchland 1984, 74)

Finally, expressions of this sort of new detectivism also appear in contemporary psychology. Nicholas Humphrey, a neuropsychologist, writes:

It is as if I, like every other human being, possess a kind of 'inner eye,' which looks in on my brain and tells me why and how I'm acting in the way I am—providing me with what amounts to a plain man's guide to my own mind. (Humphrey 1986, 87)

\* \* \*

The new detectivist looks back on his old detectivist predecessors and thinks: "They were right, of course, to claim that what brings a mental state to consciousness is its being detected. But they were wrong to think this detection such a special business. Ever since Descartes, philosophers have been overly impressed by the subject's *special* relation to his own states of mind. There's nothing so special about this relation. We are able to talk about our own pains, fears, intentions, etc. thanks to the fact that we learn about them, and we learn about them in the same sorts of ways that we learn about the goings-on around us. We happen to have been equipped by evolution with mechanisms that are sensitive to, and inform us about, the goings-on in our own minds."<sup>9</sup> I believe that this gets things precisely backwards. Old detectivists were *right* in thinking that the subject's relation to his own mental states is special. Their *mistake* lay in understanding this as a matter of detection.

That this was where their mistake lay is not easy to show. Most of the anti-detectivist arguments that one finds in the philosophical literature (both contemporary and not so contemporary) can be answered by a smart new detectivist. Consider, for example, the following from a book by John Searle:

9. Armstrong writes: "[F]or the Cartesian tradition the current contents of the mind all intimate themselves to an infallible faculty. What glory for inner sense! But, of course, to many, especially to those with a naturalist bent, it all seems too good to be true. The proper reaction, I believe, . . . is to scale down inner sense to a faculty like any other human faculty of knowledge, a faculty which fails to discern much, and which can fall into error" (Armstrong and Malcolm 1984, 129).

[W]e seem to get a vicious regress if we hold that the phenomenon of bringing unconscious states to consciousness consists in perceiving previously unconscious mental phenomena . . . For the question then arises: What about the act of perceiving—is this a mental phenomenon? If so . . . it would appear that for me to become conscious of that act, I would need some higher-level act of perceiving of my act of perceiving. (Searle 1992, 171)

Someone who thinks that mental phenomena become conscious by being inwardly perceived could reply to Searle's argument as follows: "Even if it's true that in order for an act of inward perceiving to become conscious, a higher-level act of perceiving is necessary, this does not imply any vicious regress. Some mental phenomena, including some acts of inward perceiving, are *not* conscious;<sup>10</sup> they are not themselves inwardly perceived.<sup>11</sup> A vicious regress would threaten only if it were claimed that *all* mental phenomena are inwardly perceived."<sup>12</sup>

### 1.3. A Dialogue

One source of difficulty in trying to argue against new detectivism lies in the variety of forms that it can take: an anti-detectivist argument

10. In his *New Essays on Human Understanding*, Leibniz has Theophilus say: "[I]t is impossible that we should always reflect explicitly on all our thoughts; for if we did, the mind would reflect on each reflection, *ad infinitum*, without ever being able to move on to a new thought" (Leibniz 1981, II, i, §19). The word "all" is doing real work here; no such objection can be raised against the claim that we reflect explicitly on *many* of our thoughts.

11. Armstrong (1968) writes: "If we make the materialist identification of mental states with material states of the brain, we can say that introspection is a self-scanning process in the brain. The scanning operation may itself be scanned, and so on, but we must in the end reach an unscanned scanner. However, the unscanned scanner is not a logically unscannable scanner, for it is always possible to imagine a further scanning operation" (324).

12. New detectivists are committed to the idea that someone's mechanism of inner sense—like the mechanisms that enable a person to see and hear—could conceivably break down. Thus, another strategy for arguing against new detectivism would be to try to show that this sort of breakdown is not conceivable. Sydney Shoemaker pursues this strategy in several of his papers (e.g., Shoemaker 1988 and 1994); he argues that we cannot so much as imagine what he calls "self-blindness." I criticize this argument in my Postscript and, at greater length, in Finkelstein 1999b).



that speaks against one version of new detectivism may do nothing to undermine a slightly different version. It is partly for this reason that the argument against new detectivism I'll be presenting in this section takes the form of a dialogue between a new detectivist (ND) and an anti-detectivist (AD). The dialogue form encourages us to imagine a new detectivist who reformulates his position repeatedly in response to criticism.

AD: Tell me if I'm getting this right. Unlike the old detectivists, who took us to be perfect detectors of our own mental states, you new detectivists admit the possibility of mental states that go unnoticed by their subject.

ND: That's right. If we're going to say that we detect our own thoughts and feelings, we have to admit the possibility of our failing to notice some of them. And we do fail to notice lots of them.<sup>13</sup> When a person is *unconsciously* angry or jealous, she's failed to detect the presence of one of her own mental states. I speak with first-person authority about my states of mind not because I have *perfect* epistemic access to them, but simply because, as a rule, I have *better* epistemic access to them than other people have.

AD: Well then, the first step toward seeing what's wrong with your way of accounting for first-person authority is to mark the distinction between two issues: whether or not some mental state is *known* to its subject and whether or not it is *conscious*. I might know that I am angry, even though my anger is unconscious. Imagine that my therapist, whom I trust and have good reason to trust, tells me that I've been unconsciously jealous of my sister since I was a child. I might, thereby, learn a fact about myself—that I harbor unconscious jealousy—without becoming consciously jealous.

A way to put this point would be to say that there's a distinction between "conscious *of*" and "consciously." I may be con-

13. Armstrong (1968) writes: "When we perceive, there are many (indeed innumerable) features of our environment that we do not perceive. In the same way, when we are aware of our current mental states, there are mental states and features of mental states of which we are unaware. These are mental states or features of mental states of which we are unconscious" (326).

scious *of*, i.e., aware of, my jealousy without being consciously jealous. For me to be consciously jealous—i.e., for my jealousy to be conscious and not merely something I am conscious *of*—it does not suffice for me merely to *know* that I am jealous.<sup>14</sup>

ND: So there's a distinction between someone's knowing that he's jealous and his being consciously jealous. What bearing do you take this to have on my way of explaining first-person authority?

AD: Let me say a little bit about first-person authority, about the sort of phenomenon it is; then, I'll answer your question. Consider a case in which I speak about someone else's state of mind. Imagine I tell you that a mutual friend of ours—call her Ruth—is angry with her uncle. You might just accept my assertion, or you might ask me to provide evidence in support of it. Whether or not you ask for evidence, if my claim about Ruth is responsible, then I should be able to provide some. It could be that I saw Ruth shouting at her uncle or slamming a door in his face, or perhaps she told me that she was furious with him.

Imagine now instead that I tell you that I'm angry with my uncle. One difference between this sort of case and one in which I'm speaking about another person's anger is that here, it would be odd for you to ask me for evidence supporting my assertion without your having some genuine reason for supposing that I wasn't angry with my uncle. But let's imagine that you ask anyway; you simply insist that I cite some evidence supporting my claim that I'm angry with my uncle.

I might be able to provide evidence. Perhaps I remember recently shouting at my uncle or slamming a door in his face. But such incidents seem beside the point, and whether or not I am able to think of any it would be reasonable and natural for me to reply: "What do you mean? I'm just really angry with him."

There's an asymmetry between speaking about someone else's anger and speaking about one's own. I am able to ascribe mental states to myself responsibly without being able to cite evidence in support of the ascriptions. This is a central feature of first-person authority.

ND: And so?

14. For further discussion of this point, see §5.4.

AD: Notice that we don't speak with that sort of authority about our own mental states unless they are *conscious*. Imagine that I say, "I'm unconsciously angry with my uncle," whereupon you ask what the evidence is for this assertion. Here, I cannot reasonably respond: "What do you mean? I'm just really angry with him." If someone's mental state is not conscious (i.e., if he's not consciously angry, afraid, intending to visit Paris, believing that his love is true, or whatever), then although he may be aware of it, he cannot speak about it with first-person authority. The claims that I make about my unconscious states of mind are only as good as the evidence that stands behind them.

And now, I can say what's wrong with your position. In order for me to speak with first-person authority about some mental state, it's not enough that I *know* about it; it must be conscious. What you can *explain*, however, simply by positing a mechanism that enables us to detect our own states of mind is, at most, our *knowledge* of them. You cannot thereby explain how it is that we come to be consciously angry (or afraid, or intending to visit Paris . . .). And since merely knowing one's own state of mind is compatible with not having first-person authority about it, you cannot explain first-person authority either.

ND: I grant that your therapist might inform you that you are angry without your becoming consciously angry. If your knowledge that you are angry is based on *testimony*, then you aren't consciously angry and you won't speak with first-person authority about your anger.

What's needed in order to be consciously angry—what's needed for first-person authority—is *perceptual* self-knowledge. It isn't enough that you *somehow* find out that you're angry. Being consciously angry is like having a thing *in view*, not like learning something at second hand. When you are consciously angry, you can *track* changes in your anger as they occur. First-person authority is like the authority of an eyewitness. It can be explained if we say (as most new detectivists have) that the mechanism which enables us to detect our own mental states is a *perceptual* one.

AD: But a person might perceive that he's angry without being consciously angry. Imagine someone who can tell that he is angry by watching his own face in a mirror. (Maybe he's a genius at read-

ing people's faces.) He might be able to track subtle changes in his anger this way, but if his self-awareness were based on what he saw in the mirror, he would not be consciously angry, and he would not speak about his anger with first-person authority.

It's precisely here that old detectivism was more appealing than new detectivism. The old detectivist said, "We know our own minds by inner sense," but what he meant by "inner sense" was something very different from seeing or hearing. At least such a view registered a kind of appreciation of the fact that being consciously angry is very different from being aware of objects and events that one sees or hears. (The problem with old detectivism was the *way* that it registered its appreciation of this difference—by representing our relation to inner states and events as involving a supernaturally reliable and direct kind of detection.) But you new detectivists try to explain consciousness by appealing to a cognitive mechanism that is on all fours with seeing and hearing. So against your kind of detectivism, the case of someone who looks in the mirror and sees that he's unconsciously angry may be adduced as a counterexample.

ND: I never claimed that just *any* sort of perception would suffice for consciousness. My position is that I am consciously angry when I perceive that I'm angry *in the right way*—let us say: via the operation of an *inner* sense. The word "consciously" indicates that a *particular* mode of perceptual access is in play.

AD: But you don't *explain* anything by so using the word "particular." You need to say what the relevant difference is between this inner sense and the outer senses. You need to make it intelligible that this inner sense, unlike any of the outer senses, gives rise to consciousness and first-person authority. I don't think you can.

The importance of consciousness just isn't like the importance we attach to learning something by this, rather than that, mode of perception—by, for example, hearing rather than seeing. I might pay a therapist for years with the aim of bringing emotions to consciousness that I already know myself to harbor unconsciously, e.g., with the aim of becoming consciously angry at my parents or consciously sad about my violent childhood. Imagine someone who is almost never consciously angry, sad, lustful, proud, or jealous, but who is such a keen observer of his own behavior that he can say more about the unconscious occurrence of

these emotions in himself than most people can say about the occurrence of these emotions (conscious or unconscious) in themselves. The life of such a person would be importantly incomplete: it would be thin, cold, and impoverished. The difference between a life impoverished in this way and a life of ordinary self-awareness cannot be understood as the difference between learning a set of facts by one mode of perception rather than another. *ND*: I agree that a person who was never consciously sad or angry (etc.) would have a thin life. But my conception of self-knowledge as a mode of perception *is* able to do justice to the kind of incompleteness that would characterize such a life. A blind man suffers a related kind of incompleteness; he lacks a certain sort of phenomenology. Each mode of perception provides us with phenomenology as well as information. We *savor* moments of staring out over the ocean, hearing sea gulls call, smelling the sea breeze, etc. The inner sense is no exception; it provides us with a particular sort of phenomenology. This accounts for much of its importance in our lives.

*AD*: Imagine a person with blindsight across his entire visual field.<sup>15</sup> He would not savor moments of staring out over the ocean. Only conscious seeing involves the sort of phenomenology that you're talking about. I take it that given your other views, you ought to understand conscious seeing as: seeing plus perceiving that one is seeing via the inner sense. So really, it's not the case that on your view, each mode of perception provides its own distinctive phenomenology; only the inner sense provides us with phenomenology. Introducing the topic of phenomenology only makes your view more difficult to maintain. How can we understand the inner sense as just one among several senses given its *unique* connection to phenomenology?

15. Someone is said to suffer from blindsight if his primary visual cortex has been damaged in such a way that he retains the ability to visually detect features of stimuli in the affected region of his visual field even though he claims to be unable to see the stimuli. (So: the neurologist asks, "In which direction is the light moving?" whereupon the patient replies, "I can't see anything on that side of me." The neurologist then instructs the patient to just *guess* at the direction of movement, and the patient's guesses are correct.) Such people are often said to see the stimuli unconsciously. *AD* is imagining a person who has blindsight across his whole visual field. Actually, no such person has been documented, but this may be merely because no one has suffered the right sort of damage bilaterally.

Anyway, I don't think we should be talking about phenomenology. You yourself want to say that the inner sense discloses such mental states as believing, intending, and hoping. Yet, there isn't any characteristic phenomenology associated with consciously believing, intending, or hoping. Still, just as there is an important difference between my being consciously jealous and my knowing that I'm unconsciously jealous, there is an important difference between my consciously believing, for example, that my parents never wanted me, and my knowing that I believe this unconsciously.<sup>16</sup> You can go on saying that somehow, it's the mode of perception that makes the difference. But how? Once we've seen that what's at issue is not phenomenology, how can we think that the mere fact that I know my state by this, rather than that, mode of perception could matter so much to my life?

*ND*: Let me try another tack. Seeing a thing is different from being told about it—not just because I can track what I see and not just because of phenomenology, but because there is a kind of informational or aspectual richness that comes with seeing. When I'm describing an object that I have in view, I can go on and on about it. When I see, e.g., my dining room table, I see it *as* brown, *as* dirty, *as* round . . . (Moreover I see the brown *as* dark, *as* saturated . . .) There is no end of aspects here.

Each mode of perception comes with its own sort of aspectual richness. I see my guitar as shiny, reddish, unusually small; I hear it as slightly out of tune, warm, old-sounding; I feel it as cool, dry, and cracked. Perhaps what's important about the difference between learning that I'm jealous via the inner sense and learning this via some other sense lies in the aspectual richness associated with the inner sense.<sup>17</sup>

*AD*: Well, I *can* go on and on, both when I'm describing an object in my visual field and when I'm talking about my state of mind. I don't know that I could go on *endlessly* about my anger, but I can

16. A related point: there is nothing Moore-paradoxical about saying, "I unconsciously believe that my parents never wanted me, but actually, they wanted me quite a lot."

17. Lycan (1996) writes: "I must reemphasize my presumption that internal monitoring normally or often does give rise to introspective belief, and I should also note that, again on the model of external perception, introspection presents its object under an aspect, *as* being a certain way." (27).

usually say how extreme it is, at whom it is directed, and what it's about. But it is a mistake, nevertheless, to assimilate mental state avowals to descriptions of observed objects. To see why, notice first that my seeing an object as  $\phi$  does not preclude my also hearing or feeling it as  $\phi$ . I might both see and feel some object as, for example, cracked. If I understand what you mean by "aspects," bats and dolphins *hear* many of the aspects of things that we see. They hear things as large, curved, rough, off to the left, etc.<sup>18</sup>

Now given a new detectivist's conception of the mind, we should be able to imagine someone's *seeing* the very aspects of his own, say, anger that are, typically, inwardly sensed—someone who visually inspects his own brain (or face or behavior) and sees, e.g., that he is mildly angry at so-and-so for such-and-such. Consider a case in which someone is able to assert that he is mildly angry with his uncle only on the basis of some such visual feat. If what mattered about consciousness lay in a subject's inwardly perceiving aspects of his own inner state, then such a person would have what matters. He would be as good as consciously angry. But such a person would not be as good as consciously angry; he'd be utterly detached from his anger. Here's another example that brings out the same point: if I were able to talk about some belief of mine (including my degree of conviction and whatever other aspects I'm normally able to report) only thanks to a visual inspection of my brain or behavior, the belief would not be *mine* in the way that conscious beliefs are mine. If it became apparent to me that the belief was false, I would not be *surprised* in the way that a conscious believer is surprised when he finds that his belief is false. The significance of consciousness in our lives does not lie in the aspectual richness associated with some inner sense.

I'll sum up. In order for me to speak with first-person authority about one of my states of mind, it is not enough that I *know* of or about the state; it must be *conscious*. I might know a good deal about one of my unconscious states of mind; in such a circum-

stance, I would not speak about it with first-person authority. Objects of perception (or more broadly, detection)—e.g., the table before me—are said to be "known" rather than "conscious." (My seeing a table doesn't make it correct to say, "The table is conscious.") Therefore, merely positing an *inner* mechanism of perception (or detection) does not explain first-person authority. This is not to deny—nor, of course, is it to affirm—the existence of a mechanism by which we inwardly perceive our own states of mind. The point is that the discovery of such a mechanism wouldn't account for consciousness or first-person authority.

You have claimed that your postulated inner sense differs from the outer senses in that it makes its objects conscious and so makes them items about which we have first-person authority. You've failed, however, to explain *how* a mechanism of perception (or of detection) could have this power. In the absence of such an explanation, you simply have no account of either consciousness or first-person authority.

The shift from old to new detectivism could be compared to what happens in bad psychotherapy: the patient gives the appearance of improving while his neurotic defenses, which need to be broken down, are strengthened. The old detectivist's *problem* was that he understood a mental state's being conscious too much on the model of a table's being perceived—even if "inner sense" was taken to be radically different in nature from "outer sense." With the emergence of new detectivism, philosophers tried to avoid the twin pitfalls of dualism and scepticism by understanding the relations we bear to our own conscious mental states to be even *more* like those we bear to tables and chairs. I have argued that the result is a picture of mind that leaves first-person authority looking like a mystery.

If we reject detectivism, both old and new, how are we to make sense of first-person authority? In Chapter 2, I'll discuss what can appear to be the only answer to this question.

18. If you're inclined to object that a bat or dolphin cannot hear anything as, e.g., curved because bats and dolphins lack concepts, then let the point concern an imagined creature who has both the concept *curved* and a capacity for echolocation.