DRETSKE'S WAYS OF INTROSPECTING

William G. Lycan
University of North Carolina

> '[I]ntrospection' is just a convenient word to describe our way of knowing what is going on in our own mind, and anyone convinced that we know—at least sometimes—what is going on in our own mind and therefore, that we have a mind and, therefore, that we are not zombies, must believe that introspection is the answer we are looking for. I, too, believe in introspection.
>
> --Dretske (this volume), p. xx11xx

1. For some years now, Fred Dretske has been seeking a counter-etymological account of introspection—that is, a theory that does not portray introspection as anything like, my *looking inward* and thereby finding out what is going on in my own mind. (Evidently he does not accept James Grier Miller's dictum that ontology recapitulates philology.) He has been uniformly and sharply critical of Lockean "Inner Sense" views of introspection such as those defended by Armstrong (1968, 1981) and me (Lycan (1987, 1996)).

He has worked his way through several different positive views of his own. But throughout, his theme has been the *epistemological primacy of content*: When I am aware of being in a particular mental state, what I am aware of is first and foremost the worldly intentional content of that state, and it is the existent or inexistent features constituting that content, not features of the state itself, that both produce and warrant my belief that I am in the mental state. Dretske (1995) says,

> What one comes to know by introspection are, to be sure, facts about one's mental life—thus (on a representational theory [of the mind]), representational facts. These facts are facts, if you will, about internal representations. *The objects and facts one perceives to learn those facts, however, are seldom internal and never mental…. One becomes aware of representational facts by an awareness of physical objects*. (p. 40, italics mine)

I learn that my thought *t* or experience *e* has a characteristic mental property F by perceiving the (typically) physical features that *t* or *e* represents. Dretske insists, notice, that I do not do this by detecting that *t* or *e* does represent this or that, or by noting any other feature of *t* or *e* itself.

It is also the features represented by my first-order mental state that alone give my introspective awareness of the state its vaunted special authority, my first-person "privileged access."

What has varied in Dretske's various accounts of introspection has only been the way in which these core ideas have been implemented.

2. Dretske (1995, Chapter 2) offered a model based on a perceptual analogy, but unlike the Inner Sense model, on a kind of *displaced* perception. Examples of ordinary displaced perception would be, seeing how much gas there is in your tank *by* (merely) seeing the gas guage, or seeing how close you are to the runway by seeing the instruments in your cockpit, or hearing what Yassir Arafat said by hearing it reported (in English) by an NPR announcer rather than by hearing Arafat's own Arabic utterance. You are able to "perceive" in this very indirect way (let us call it "d-perceiving") when you have a justified belief that the directly and more literally perceived object would not have the perceived feature it does unless the target object "(probably)" had the properties d-perceived as inhering in it (p. 42). Thus, you have d-heard what Arafat said only if you are justified in believing that the radio announcer would not have announced that Arafat said it unless Arafat (probably) had said it. (This is a very liberal criterion of displaced perceiving.)

Turning to introspection, Dretske says, "...an experience (of blue, say) is conceptually represented as an experience of blue via a sensory representation not of the experience, but of... (typically)...the blue object one sees" (p. 44), where the "via" is that of displaced perception. For me to introspect that the experience is one of blue is for me to represent that it is, but only very derivatively, on the sole basis of the experience's representing the blue object as blue plus the justified belief that the object would not be blue unless the experience (probably) were one of blue. More generally, for me to introspect that *e* has F is for me (a) to represent that *e* has F, (b) on the basis of my seeing that *e*'s intentional object has some property P, (c) the former representation being mediated by the justified belief that the object would not have P unless *e* had (probably) had F.

3. Without fanfare, Dretske (1999) abandons the foregoing account. (And with good reason, in my opinion.<**1**>) He switches to a more complicated story. He continues to maintain that the external feature represented by an experience directly "reveals" to the subject "exactly what property...it is that his or her experience has" (p. 113), and that "[i]t is the... [thing] we think *about*...that (when we introspect) 'tells us' *what* we are thinking about..." (p. 116). As before, this would explain the psychological immediacy and epistemic privilege that we associate with introspection, if anything better than Inner Sense views explain those things. (And he explicitly denies that "[o]ne is made aware of what a pumpkin experience is like...by an awareness of the experience" (p. 112).)

But now he offers a new analogy: Consider photographs, specifically, color photographs of orange pumpkins.

> It is sometimes enough to look at what these pictures are pictures *of* (viz., the pumpkins) in order to tell what the picture is like.... [Similarly, t]o know what the experience is like, what properties it has, it is enough for the experiencer to 'look at' what the experience is an experience *of* (something that, as the experienc*er*, he cannot

help but be doing). *That* will tell him what the (relevant) properties of the experience are. (p. 112)

To the obvious objection, that in order to know what a pumpkin photograph is like (without looking at it) one must know something about the camera, the camera settings, the film and so on, Dretske replies,

> [That] is an important disanalogy. It is, in fact the reason why one cannot tell what it is like to be a bat by looking at what the bat 'looks' at (perceives)--a moth, say. We don't know enough about the bat's 'camera' and 'film' to tell much about what the bat's internal 'pictures' of the moth are like. But this point, though important, does not upset the usefulness of the analogy. For in the case of experience, the person having the experience, the experien*cer*, looks at objects *with* the 'camera' and 'film' whose characteristics are (or may be) unknown to third parties. The person sees objects, as it were, *through* the lens and *with* the pre-loaded film--hence, *as* these objects are represented in his experience of them. (p. 123, n. 12)

Lycan (1999) argued against this, contending that the disanalogy does upset the usefulness, if not the correctness, of the analogy: It makes it hard or impossible to see how introspective beliefs are justified. Dretske is trying to explain how we are "made…aware of what our experiences are like" (p. 112). And that awareness is a kind of knowing. At least, surely, introspection justifies our introspective beliefs about our experiences. Now, in the case of a photograph I have not seen, I can be justified in holding certain beliefs about how the photograph looks, because I know something about the camera and film and I have previous experience, gained through inspection of samples, of what sort of photographs are produced by that sort of camera and film in such-and-such circumstances. That knowledge is of course inferential. But Dretske's new view of introspection affords no parallel. Though I do look at external objects "*with*... [my visual] 'camera' and 'film'," how might that justify my belief that my resulting experience is F? Not because I know something about my visual processes and I have prior experience, gained through inspection of samples, of what sort of experiences are produced by those processes in such circumstances. (I may remember what an orange pumpkin looked like to me on one or more past occasions, but how did I know at *those* times what my experience was like?) My belief about my experience does not seem inferential at all—as Dretske himself emphasizes.

Notice that epistemological Reliabilism will not help.**<2>** The pumpkin and its orangeness (themselves) are not reliable indicators that one's experience is F, or indeed of anyone's experiencing anything in any way at all. Up till this point, Dretske has nowhere answered the question of how introspection yields justified belief.

4. Instead, Dretske (in press) shifts focus. He addresses a particular version of the problem of externalism and self-knowledge, the type of anti-compatibilist argument made by McKinsey (1991), Brown (1995) and Boghossian (1998), according to which if externalism is true, then we

can achieve purely introspective knowledge of brutely empirical facts about the external world: **<3>**

> If, as some externalists hold, you cannot think that something is water without having stood in causal relations to water, then it seems to follow that to know, in a special authoritative way, that you think you are drinking water is to know, in that same authoritative way, that there is (or was) water. But you cannot know, not at least in *that* way (the way you know that you *think* you are drinking water), that there *is* water…. So externalism is false. (p. xx3xx)

To his credit, Dretske is an externalist about nearly everything, but as before he does not want to deny the special authority of introspection; so he must produce an explicitly externalism-friendly account. The externalism-friendliness itself is hardly a problem, given his uniform emphasis on the primacy of intentional content, but he must respond specifically to the foregoing anti-externalist argument. His responses are two.

First, he contends that the argument assumes something false:

> It assumes that knowledge of what is in your mind is, or requires, knowledge that you have a mind. It assumes that knowledge of what you think—for instance, that there is water—is (or requires) knowledge that you think. This is false. The special authority we enjoy about our own mind is an authority about *what we think*—that, for instance, there is water—not about the fact *that we think it*. (p. xx3xx)

Second, accordingly concentrating on knowledge of attitude content rather than of the attitude itself, he advocates a vivid version of the Burge-Heil line on self-knowledge of content (Burge (1988), Heil (1988)), that our beliefs about our attitudes' contents simply and automatically inherit those contents: If I believe that there is water in the fuel tank, and I introspectively believe that I believe that there is water in the fuel tank, whatever externalist psychosemantics makes the first of those mental tokens of "water" designate H2O also (of course) makes the second of them designate the same substance; so my introspective belief does not require extra justificatory steps to make sure that my first-order belief really is about H2O rather than about XYZ.

Dretske offers the analogy of a simple measuring instrument whose job is to announce the value of some magnitude Q (temperature, voltage, velocity or the like). The instrument's output is thus labelled "VALUE OF Q." But that output can, additionally but at no extra cost, be labelled "VALUE THAT Q IS REPRESENTED AS HAVING." In this way, the device represents its own represented content as well as that same content in the first place.

The explanation of authority and privilege is simple:

> About this topic—the content of its own thoughts about Q—it has authoritative (because infallible) and privileged (no *other* instrument enjoys this kind of authority about *this* Q-meter's representational states) information. As long as we assume that the way a Q-meter represents something as having a Q of 5 is by pointing at the numeral '5' on a suitably calibrated scale…, Q-meters are infallible about what they

'think' (that is, represent) about their own representational efforts. (pp. xx5-6xx)

However, this is to give the device authority only over the content, "*how* it is representing the world," and not over the mental fact "that that, in fact, is what it is doing—*representing* Q" (p. xx7xx). So too with human introspection:

> Whatever way we have of telling *what* it is we think and experience is not a way of telling *that* we think and experience it…. [M]y privilege and authority do not extend to the fact that I think and experience. (pp. xx12-13xx).**<4>**

Self-ascription of thoughts and experiences must be justified in some other way—and as always, not by reference to our being aware of those thoughts and experiences themselves. Dretske recommends the positive views of Evans (1982) and Shoemaker (1988).

5. I have a number of doubts about and objections to all this.

Regarding the first response, I do not see how the McKinsey-Brown-Boghossian type of incompatibilist argument does presuppose that knowledge of what I think requires knowledge that I think. McKinsey and Boghossian assume, merely for the sake of argument, that I do know introspectively that I am thinking about water or the like, and they argue from there that if externalism is true then I gain illicitly automatic nonempirical knowledge that water exists.

Probably what Dretske has in mind is, rather, that McKinsey's and Boghossian's stipulative assumption is not what should be at issue. Since he thinks the assumption is false, he does not much care whether it is compatible with externalism.**<5>** He is concerned to show only that introspective knowledge of content itself is compatible with externalism.

But that brings up a larger question: What does Dretske suppose is knowledge "of content itself"? According to him, I know my intentional contents infallibly, but I do not so know the attitudes whose contents they are. How to express this disparity? I infallibly know the content that P, but I do not know in the same way that I am believing that P or that I am thinking that P, or anything of the form "I am V-ing that P." How am I supposed to know <that P>, the content, all by itself, without knowing in the same way that I am in any way entertaining the proposition that P? (Of course, this knowing the content <that P> is putatively a kind of self-knowledge, sharply distinct from *knowing that P*.)

Dretske may suspect that I am committing a kind of subtraction fallacy. Perhaps his idea is that although (granted) I cannot know <that P> without knowing or at least believing that I am V-ing that P, for some attitude V, the basis and status of my knowledge that I am V-ing that P are not uniform across that fact that I know. I know infallibly that it is *that P* I am V-ing, but I know only fallibly, if at all, and by a different method, that it is *V-ing* that P that I am doing. I am not entirely sure how to make sense of that, but I expect one could.**<6>**

6. I have deeper reservations about the Q-meter analogy. First, Dretske's new model of introspection entails that introspection is *hyper*infallible—not just nomologically infallible, or metaphysically. According to the analogy, to represent the content that P and introspectively to know that content are one and the same state of affairs, separated only in the merest thought by

the facile affixing of a second label. But the introspection of content is not infallible at all, not even nomologically. One can be mistaken about what one hopes for or about what one desires or about what one believes. And this does not mean just that one can mistake the attitude, as when I think I hope that P when I do not in fact hope that P but only wish that P; I can be aware of a particular hope I cherish, and still think it is a hope that P when actually it is a hope for something else.

One can also mistake the content of what one believes. Belief in singular propositions is the obvious example. Suppose you briefly catch sight of (it seems) your old acquaintance Ted, and you believe that he has put on weight. In fact, the man you see is Ted's twin brother Ned, who you did not know existed. What you believe is the singular proposition that Ned has put on weight, but you mistake it for the proposition that Ted has. (Of course, you also believe the latter proposition, since you believe Ned *is* Ted. But still you believe that Ned, the man you saw, has put on weight, and you mistake this for the proposition that Ted has.) Less tendentious content mistakes are possible as well. I once thought I believed that my daughter's cat Rudy was stupid, until I realized that what I was actually believing was that Rudy was mulishly uninterested in abetting any project of any human being.<**7**>

7. My second reservation about the Q-meter model: Since Dretske's position is a version of Burge-Heil compatibilism, how does he propose to respond to the standard objections to that view, especially to those based on planet-switching? Simple switching will not bother him, for familiar reasons. E.g., suppose you are transported to Twin Earth during your sleep, and you awaken believing yourself to be still on Earth just as always. When you look at some Twin "water," XYZ, you will think that it is water, and you will be wrong. But you will think you are thinking that it is water, and in that you are right. Even after you adapt after some years on Twin Earth (assuming that you do adapt), and your "water" thoughts are then about XYZ rather than H2O, of course your thoughts that you are thinking about "water" will be about XYZ as well, and you are right again. Your introspective beliefs cannot go wrong. And as before, that explains the special security of your introspective knowledge.<**8**>

However, Boghossian (1989, pp. 22-23) offers a more complicated switching case that is more troublesome for the externalist. You are transported to Twin Earth during your sleep, and things go much as before. You adapt. But then someone reveals to you that you have been unwittingly transported and that you are now on Twin Earth. But that person refuses to tell you when the switch occurred. Now you are asked, were your "water" thoughts one year ago about Earth water (H2O), or about XYZ? It seems clear that you do not know; you have no idea. Boghossian then argues that if you do not know now, then you did not know then, because you have not *forgotten*, or lost previous knowledge in any other recognizable way. (Boghossian stipulates that you have not lost the Earth concept of water, either, but in adapting have merely added a second concept, the Twin-Earth concept.) So it seems that Burge-Heil makes the wrong prediction in this case.

I do not doubt that Dretske has replies available to objections of this kind, but to my knowledge he has not offered them.<**9**>

Third reservation: How, exactly, does Burge-Heil apply to McKinsey-Brown-Boghossian (the

argument that externalism predicts illicit a priori knowledge of empirical facts)? Burge-Heil was originally meant to undercut the assumption that if externalism is true, then in order to know my own thought contents I would have to know exotic empirical facts about causal-historical chains or teleology or the like. And so it did undercut that assumption.**<10>** But where is it supposed to intercept the argument from externalism to the thesis that we can achieve purely introspective knowledge of empirical facts about the external world? According to Dretske, I can know the content <that there is water in the fuel tank> infallibly. Without empirical input I can infer from that content that there is water. So it seems I still have my illicit nonempirical knowledge, Burge-Heil notwithstanding.

My earlier concessive suggestion about knowledge "of content" may help. Suppose I do know infallibly that it is *that there is water in the fuel tank* that I am V-ing, even though I know only fallibly that it is V-ing that I am doing. Perhaps in order to infer that water exists, I must exploit not just the content itself but that I am V-ing it, or at least that I am taking some attitude toward it. And suppose we agree with Dretske that the latter, fallible knowledge is not introspective or a priori in any sense at all. (Does he hold that I infallibly know the content without introspectively knowing even that I am taking some attitude or other toward it?) That would block the implication that I have a priori knowledge of water, since an empirical premise was required for the inference. But to take this line, Dretske would still need to make sense of the idea of selective bases and status of knowledge across the fact known, and he would also need to provide an argument for the claim that in order to infer that water exists, I must appeal to a premise about my V-ing the content regarding water.

8. Fourth reservation: There is an activity naturally and rightly called "introspect*ing*." It is something we do, ostensibly as a means of learning things about ourselves and our mental states. If I have never particularly noticed what Pittsburgh tap water tastes like to me, I can attend carefully the next time I drink some, and find out. If my doctor asks me to describe my wrist pain, I can resist distraction, focus my attention on the pain, and give the doctor more detail than I had previously been aware of. If a cognitive psychologist asks me how I solve a certain puzzle involving conditional sentences, I can introspect and detect the steps in my reasoning (this is not to say that such experiments show their subjects' introspective reports to be very reliable).

For that matter, if I now ask you to pay attention to the upper right quadrant of your visual field, without shifting your gaze, and to describe its contents, you can do that. Now (in real life), please attend to your left forearm, wrist and hand; are you experiencing any sensations in them? Now tell me whether you are now hearing any sound, however faint, that approximates a musical tone. Notice that your having done all that was entirely under your voluntary control. And undoubtedly that sort of introspecting is subserved by some brain mechanism, since it is a matter of deliberate attending and focusing, which are occurrent (if not at all modular) psychological events.

Now, *that* is introspecting, with a vengeance. It produces belief. It affords learning. It is a way of finding out. It justifies the belief, too, or so I would argue.**<11>**

Dretske has an obvious account to give of such introspecting, consistently with his thesis of the

epistemological primacy of content and with his rejection of Inner Sense:**<12>** He can say that all the busy attending and focusing gets done, all right, but what are attended to and focused on are, as always, just the intentional contents of the states being introspected. We must remember that those contents are normally real physical objects and their properties in the external world. To attend to them is to attend to them *perceptually*, using our sense organs. Even when the intentional objects are intentional inexistents such as illusions or hallucinations, our attention and focusing are perceptual, not internal. That "introspecting" is an activity of this sort does nothing to show contra Dretske that introspection makes us directly aware of thoughts and experiences and sensations themselves. (This account is a very strong version of what has come to be called the "transparency thesis" regarding introspection and sensory qualities.**<13>**)

But, I would maintain, introspecting reveals more than just intentional contents. There are any number of examples of this. Start with beliefs. I can introspect (fallibly) what it is I believe. But I can also tell (fallibly) by introspection how strongly I believe it—tentatively, or pretty confidently, or am I unalterably convinced? Or take pains. I firmly agree with Armstrong (1968), Pitcher (1970) and Dretske (1995) that pains are proprioceptions of bodily states; they have intentional objects, real or inexistent, which are damaged or otherwise unsalutary conditions of body parts. But those objects are not all I can introspect about a pain. I can also introspect its hurtfulness and the urgency of my desire that it cease.

Or go back to perceptual qualia. I introspect a red patch in my visual field, and a tone whose pitch is tenor D# coming from my left. Let us grant that that is in part to detect external, physical redness and a vibration in the air of such-and-such a frequency. But those properties themselves determine nothing about any sensory modality, much less any finer-grained mode of presentation. I also introspect that the redness is visual rather than something I am being told about, and that the tone is aural rather than a frequency being exhibited on an oscilloscope. To see the point more clearly, consider an uncontroversially primary quality such as squareness that can be either felt with the fingers or seen. I can introspect that I am seeing something square rather than feeling it, or vice versa. In a case of illusion, I can introspect both that I am feeling something square and, at the same time, its not looking square at all.**<14>**

I do not see that Dretske's present view yields any account of such introspectings.

9. Dretske (this volume) asks how we know we are not zombies. It is not clear what he means by a "zombie." He says (p. xx18xx, n. 1) that zombies are "human-like creatures who are not conscious…of anything." But "conscious" notoriously has many different meanings, and some writers have used "zombie" as meaning simply a human-like creature that has no mental properties whatever—no beliefs or desires, for example. Dretske seems to mean something more specifically sensory, that "zombies" do not perceive or sense anything. In any case, his main purpose in the article is to extend his primacy-of-content thesis, in the form defended in Dretske (in press), more explicitly to sensory experiences: Just as "what you think…doesn't tell you that you think it," "[w]hat you see—beer in the fridge—doesn't tell you that you see it" (p. xx2xx). The puzzle is that "there is nothing we perceive that tells us we are conscious" (p. xx5xx).

Since I too hold a representationalist theory of all sensory experience (Lycan (1987, 1996)), I

agree that the issues for experience and for thought are exactly parallel. Dretske adroitly fends off one apparent difference between them. It may seem, and some have asserted, that at least some sensory experience is *necessarily* conscious; for example, since pains and itches are by definition feelings, it is conceptually impossible to have a pain or an itch without being aware of it. Then the argument goes: Since as a native speaker I know that, and I have experienced pains and itches, I thereby know that I am conscious in the sense that I sense things.

Dretske responds (p. xx7xx) that this merely changes the way in which his problem is posed. For now the question becomes, how do I know that I have had *pains* and *itches*, as opposed to bodily events that are like pains and itches in every respect except that we are not aware of them? So even if it is true that pains and itches are necessarily conscious states, that does not help with the epistemological issue.

I think Dretske's response is quite right, but as I shall explain later on, I implement it differently. He coins the notion of "protopain" (p. xx8xx), as a state

> that has all the properties you are aware of when you experience pain except for the relational one of your being aware of it. Protopain is what you have left when you subtract your awareness of pain from pain.

Those two characterizations differ importantly. The second is logically stronger than the first, if likewise understood in terms of properties: It says that protopain has *all* the properties of your pain except for your awareness of it, not just all those you are aware of. Since when you are in pain, you probably are not aware of very many separate properties of the pain, the second characterization seems to me more useful. In my speech, then, a "protopain" would have all the distinctive physiological, functional, dispositional and representational properties of a pain, except those that would be excluded by its subject's awareness of it. For reasons that will become clear, though, that is not (at all) Dretske's use of the word; he really means a particularly restrictive version of the first characterization. But the difference does not matter to the correctness of his response to the argument from necessary awareness.

There is an ambiguity, and a dialect difference, in our use of the word "pain." In one sense, as the foregoing argument has it, pain is *by definition* a feeling; the notion of an "unfelt pain" is simply contradictory. Yet ordinary people quite frequently speak of pains that go unfelt, without any sense of contradiction. Even quite a bad pain may not be felt if attention is distracted by sufficiently pressing concerns. (To defend this against theorists whose dialects insist otherwise, I have amassed a small but rewarding library of quotations from popular works.**<15>**) When one person's commonplace sounds to another contradictory on its face, there is equivocation. And in this case there demonstrably is: Sometimes the word "pain" is used to mean just protopain, which can go unnoticed. But sometimes "pain" means a conscious feeling or mode of awareness, and on that usage the phrase "unfelt pain" is indeed simply self-contradictory; it comprehends both the protopain and the awareness together.**<16>** Neither usage is correct to the exclusion of the other.

10. But now, what is a protopain? Like me, Dretske follows Armstrong and Pitcher in

thinking of pains as representational. What they represent (veridically or not) are, roughly, damage or disorder or a harmful condition in a particular body part; Dretske says, "injury, stress, or irritation to some part of the body" (p. xx9xx). Damn right.

But at this point he and I part company. Dretske makes an unexpected announcement: "[W]hat I am calling protopain is simply the bodily condition we 'perceive', the physical condition we are made aware of, when we are in pain" (p. xx9xx). Moreover, zombies have protopains even though they do not sense at all. So for Dretske the relevant psychophysical states are just two: The physical condition of disorder, injury etc., and the awareness of it that, in his concessive usage, is the *pain*.

Considered as an account of "unfelt pain" (I am not sure that Dretske intended it as such), this will not do. What is characteristic of unfelt pain is that, as I said in the previous section, nearly all the distinctive properties of a pain are there--physiological, functional, dispositional *and* representational, except for awareness of it. The unfelt pain produces many of the usual effects: wincing, cramping of movement, favoring, and possibly subvocal cries (though not *reports*, or considered analgesic behavior, which require awareness). The physical condition of disorder etc. in itself falls far short of this, though it is a distal cause of the causing of those usual effects. It is not an unfelt pain, because it is not a pain at all.

But more to the point, the physical condition is not a property of the pain state itself, but is only the state's intentional object. (Of course, the state does have the correlative property of *representing* the physical condition.) In my view, what makes a state of pain a state of pain is its functional and its representational properties as well as its represent*ed* properties. Consider the case of psychosomatic pain, or of phantom pain, in which there is no actual bodily disorder even though it seems to the subject that there is. On Dretske's usage, there is no actual protopain, no protopain save the nonactual one being represented. Yet the subject's pain mechanisms may be firing away like crazy, constituting a protopain as I would have used the word. (Not wishing to neologize, since it is Dretske's word and not mine, I should choose a different one. How about "underpain"? But then, since I am perfectly happy with the idea of unfelt pain, I may just stick with "pain.")

Verbal matters aside, does Dretske's view allow for unfelt psychosomatic or phantom pain (underpain)? If not, that would seem to me a serious defect. I suppose, though, that it does. The subject's pain system could be representing nonexistent bodily disorder, yet because of distraction or whatever, the subject could be unaware of that.

11. In the concluding section of his paper, Dretske considers and criticizes the Inner Sense answer to his title question. In this concluding section of mine, I shall reply.

His first criticism:

> …I did not become aware of the fact that I have conscious experiences by awareness of the conscious experiences themselves in the way I become aware of the fact that there is beer in the fridge by seeing the beer. I have experiences of beer bottles but not experiences of beer bottle experiences. (p. xx13xx)

Reply: Granted. Even for the Inner Sense theorist, as Dretske goes on to acknowledge, introspecting is not fully analogous to sense perception. It is just more like sense-perception than it is like mere thinking or describing. And we do not need to use the odd expression, "experiences of experiences"; "introspective awareness of experiences" will do fine.

Second criticism: Dretske continues,

> I think those who suppose they are introspectively aware of their own experiences are simply confusing a fact they are aware of—the fact, namely, that they have experiences—with objects they are not aware of—the experiences they have.

Reply: That is a canard. *I* have not ever confused the fact that I have experiences with the experiences I have, much less "simply" confused them. (And I am fairly sure that David Armstrong has never done so either.) So I do not engage in the corrupt "double dipping" that Dretske goes on to attribute to people (if any) who have fallen into that confusion.

Third criticism: Acknowledging our acknowledgement that "inner sense" is not fully analogous to perception, Dretske offers a charitable interpretation of my version of the Inner Sense theory (p. xx14xx). According to his interpretation, the introspective "scanner" gives us knowledge that we have experiences without ever making us aware of the experiences themselves; "[t]he experiences themselves are, so to speak, *invisible* to the introspective scanner." That is how to avoid the objectionable analogy with becoming aware of the fact that there is beer in the fridge by seeing the beer. But then I am impaled on the other horn of a dilemma (p. xx15xx): If the scanner does not ever make me aware of my experiences themselves, by what other means does it furnish me with the knowledge that I have them? "Inner sense" understood in this charitable way is no answer to the question of how I know I have experiences; it becomes merely a label for the admitted fact that I do know I have them. So, whatever other virtues the Inner Sense theory may have, it does not yield an answer to Dretske's title question.

Reply: Quite right. On the charitable interpretation, Inner Sense would merely label the problem. What is not quite right is the charitable interpretation. Armstrong and I very definitely hold that the introspective scanner makes us aware of the experiences we are having, and not just of the fact that we are having some.

Out of virtuous concern to set the record straight (and thus not out of *sheer* self-advertisement), I shall close by showing how the interpretation goes wrong.

According to the view I hold, introspection is the operation of an internal attention mechanism that monitors experiences and produces second-order representations of their properties, including especially their feature-representational properties. David Rosenthal (1990) objected to it that if this attending or monitoring is supposed to be a form of perceiving, I come to grief over the fact that "perceiving always involves some sensory quality." (If introspecting itself involves some sensory quality, either it is the same quality as is represented by the state being introspected or it is a different quality, and either way lies trouble.)

I replied (1996, pp. 28-29) by simply denying that introspecting per se does involve a sensory quality. I excused the disanalogy by pointing out that the sense organs have the function of

detecting specific environmental features—it is those features that, by being perceptually represented, constitute perceptual qualia—while the function of introspection is entirely different, and so we would not expect introspection to involve a sensory quality.

Dretske interprets that distinction as follows.

> Our inner sense does not reveal qualities of the objects (the experiences) being scanned. [Lycan]…tells us that these (first-order) experiences (of beer bottles) do not (like beer bottles) have 'ecologically significant features' and so our introspective 'scanning' of them does not represent them as having properties.

That is where Dretske goes astray. What inner sense does not do is, present external environmental features—(physical) colors, sounds, textures and the like--as such. So it does not "involve sensory qualities" in the way Rosenthal meant, the way in which seeings and other perceivings do; and introspectings do not have qualia even though they represent first-order states as having them. But introspection does represent our experiences as having properties. In particular, it classifies them; it assigns them to kinds. We are indeed "made aware of them, as we are of beer bottles, as objects having properties that serve to identify them" (p. xx14xx), though of course experiences are events, not physical objects like bottles.

Introspection's typology is proprietary and distinctive. It is a system of phenomenal kinds. Lycan (1996) argues that our scanners' identificatory outputs are untranslatable into natural language, and that that untranslatability accounts for several otherwise mysterious aspects of conscious experience, including its ineffability and Levine's (1983) "explanatory gap."

I am not as happy with the Inner Sense view as some of my remarks may have sounded. In particular, it runs into trouble over a certain kind of false positive (Neander (1998), Lycan (1998)), to which problem I believe Dretske's theory is immune. His view may have other advantages over mine as well. The moral is that introspection will not be well understood any time soon.

Notes

1 For one thing. I doubt that many people ever have such beliefs as specified in clause (c), much less justified ones--that, in the present case, an object's being blue counterfactually depends on one's own experience's being a certain way. One can see the point of the counterfactuals in the ordinary examples: We are able to d-hear Arafat through the radio because the announcer's saying what she does is a reliable auditory sign of Arafat's having said what he did. But a blue object's being blue is not a sign, reliable or unreliable, of one's experience's being of blue.

2 It is no big secret that Dretske himself is a Reliabilist; see particularly Dretske (1981).

3 There are other versions: The simple argument that if content is determined by external

environmental factors such as a complicated causal-historical chain, then in order for you to know your own content, you would have to know the details of the causal-historical chain; the even simpler argument that you cannot introspectively know anything that is constituted by anything outside your own head.

4 Cf. Bernecker (1996). A further distinction is needed as well. Bar-On and Long (this volume) emphasize that in addition to authority regarding content vs. authority regarding attitude, there is also the different question of authority regarding the subject: How do I know that it is I who am doing the thinking and experiencing?

5 Bernecker (1996) argues ingeniously that it is not.

6 It occurs to me that he might do this by appeal to his notion of hyperdyperintensionality as revealed through contrastive stress (Dretske (1972)).

7 If you are tempted to think that that is not what I thought I believed, then you must admit that I am fallible, indeed wrong, in what I think I thought I believed.

8 It has been complained that the Burge-Heil approach simply assumes a Reliabilist theory of epistemic justification. But to raise that complaint against Dretske would be churlish.

9 Gibbons (1996) offers an ingenious Reliabilist reply to Boghossian, but it is elaborate and tricky.

10 The assumption is unfounded in any case, because it is not generally true that in order to know that P, one must know the metaphysical truth-maker of the fact that P.

11 Based on the theory of justification defended in Lycan (1988).

12 A similar suggestion is made by Carruthers (2000).

13 Defended by Harman (1990). For an excellent review and elaborated defense, see Tye (forthcoming 2002).

14 If in these cases I am not introspecting the intentional content of my experience, what aspect of the experience am I introspecting? For example, what constitutes the difference between the tactile presentation of squareness and the visual presentation of the same property-instance? My answer (Lycan (1996)) is that they are distinguished functionally, both by the architectures of the respective sense modalities and by the functions of those modalities themselves.

Goldman (1993) raises the skeptical question of how one might introspect a property that is at bottom a functional property. Functional properties are relational properties, and involve as relata objects in the external world as well as the subject's physical behavior; but surely introspection can reveal only what is in the head. For replies to this, see Armstrong (1993), Rey (1993), and Lycan (in press).

15 "From time to time, she winced slightly as she moved in front of the jury to the easel and back to a stool. For the most part, she was so involved in her intricately constructed argument that she didn't feel the pain" (Ann Rule, *A Rose for Her Grave and Other True Cases* [New York: Pocket Books, 1993], p. 335). For another example, see Lycan (1996, p. 164, n. 6). Rosenthal (1991) offers a nice defense of unfelt pain. See also Palmer (1975) and Nelkin (1989).

16 Since in this paper Dretske agrees to reserve the word "pain" for conscious pains and coins "protopain" for the kind that can be unfelt, he seems to take the conservative side of the dialect difference. We may think this might be only for rhetorical purposes, but it is not; read on.

References

Armstrong, D.M. (1968). *A Materialist Theory of the Mind*. London: Routledge and Kegan Paul.

Armstrong, D.M. (1981). "What is Consciousness?" In *The Nature of Mind and Other Essays*. Ithaca, NY: Cornell University Press.

Armstrong, D.M. (1993). "Causes are Perceived and Introspected." *Behavioral and Brain Sciences* 16: 29.

Bar-On, D., and D.C. Long (this volume). "Knowing Selves: Expression, Truth, and Knowledge."

Bernecker, S. (1996). "Externalism and the Attitudinal Component of Self-Knowledge." *Noûs* 30: 262-75.

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

Boghossian, P. (1998). "What the Externalist Can Know A Priori." In E. Villanueva (ed.), *Concepts* (Atascadero, CA: Ridgeview Publishing).

Brown, J. (1995). "The Incompatibility of Anti-Individualism and Privileged Access." *Analysis* 55: 149-56.

Burge, T. (1988). "Individualism and Self-Knowledge." *Journal of Philosophy* 85: 649-63.

Carruthers, P. (2000). *Phenomenal Consciousness*. Cambridge: Cambridge University Press.

Dretske, F. (1972). "Contrastive Statements." *Philosophical Review* 81: 411-437.

Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge, MA: Bradford Books / MIT Press.

Dretske, F. (1993). "Conscious Experience," *Mind* 102: 263-83.

Dretske, F. (1995). *Naturalizing the Mind*. Cambridge, MA: Bradford Books / MIT Press.

Dretske, F. (1999) "The Mind's Awareness of Itself." *Philosophical Studies* 95: 103-24. Reprinted in *Knowledge, Perception, and Belief: Selected Essays* (Cambridge: Cambridge University Press).

Dretske, F. (in press). "Externalism and Self-Knowledge." In Ü. Yalçin and S. Nuccetelli (eds.), *Externalism, Self-Knowledge and Knowledge of the World*. Totowa, NJ: Rowman and Littlefield.

Dretske, F. (this volume). "How Do You Know You Are Not a Zombie?"

Evans, G. (1982). *The Varieties of Reference*. Oxford: Oxford University Press.
Gibbons, J. (1996). "Externalism and Knowledge of Content." *Philosophical Review* 105: 287-310.
Goldman, A.I. (1993). "The Psychology of Folk Psychology." *Behavioral and Brain Sciences* 16: 15-28.
Harman, G. (1990). "The Intrinsic Quality of Experience." In J.E. Tomberlin (ed.), *Philosophical Perspectives, 4: Action Theory and Philosophy of Mind*. Atascadero, CA: Ridgeview Publishing.
Heil, J. (1988). "Privileged Access." *Mind* 47: 238-51.
Levine, J. (1983). "Materialism and Qualia: The Explanatory Gap." *Pacific Philosophical Quarterly* 64: 354-61.
Lycan, W.G. (1987). *Consciousness*. Cambridge, MA: Bradford Books / MIT Press.
Lycan, W.G. (1988). *Judgement and Justification*. Cambridge: Cambridge University Press.
Lycan, W.G. (1996). *Consciousness and Experience*. Cambridge, MA: Bradford Books / MIT Press.
Lycan, W.G. (1998). "In Defense of the Representational Theory of Qualia (Replies to Neander, Rey and Tye)." In Tomberlin (1998).
Lycan, W.G. (1999). "Dretske on the Mind's Awareness of Itself." *Philosophical Studies* 95: 125-33.
Lycan, W.G. (in press). "Goldman on Consciousness." *Philosophical Topics*.
McKinsey, M. (1991). "Anti-Individualism and Privileged Access." *Analysis* 51: 9-16.
Neander, K. (1998). "The Division of Phenomenal Labor: A Problem for Representational Theories of Consciousness." In Tomberlin (1998).
Nelkin, N. (1989). "Unconscious Sensations." *Philosophical Psychology* 2: 129-41.
Palmer, D. (1975). "Unfelt Pains." *American Philosophical Quarterly* 12: 289-98.
Pitcher, G. (1970). "Pain Perception." *Philosophical Review* 79: 368-93.
Rey, G. (1993). "Why Presume Analyses are On-Line?" *Behavioral and Brain Sciences* 16: 74-75.
Rosenthal, D. (1990). "A Theory of Consciousness," Report No. 40, Research Group on Mind and Brain, Zentrum für Interdisziplinäre Forschung (Bielefeld, Germany).
Rosenthal, D. (1991). "The Independence of Consciousness and Sensory Quality," in E. Villanueva (ed.), *Philosophical Issues, I: Consciousness*. Atascadero, CA: Ridgeview Publishing.
Shoemaker, S. (1988). "On Knowing One's Own Mind." In J.E. Tomberlin (ed.), *Philosophical Perspectives, Vol. 2: Epistemology*. Atascadero, CA: Ridgeview Publishing.
Tomberlin, J.E. (ed.) (1998). *Language, Mind, and Ontology* (*Philosophical Perspectives*, Vol. 12). Atascadero, CA: Ridgeview Publishing.
Tye, M. (forthcoming 2002). "Representationalism and the Transparency of Experience." *Noûs* 36.