



CONSCIOUSNESS, NATURAL REPRESENTATION, AND FIRST-PERSON WARRANT: REPLY TO DRETSKE

Department of Philosophy
University of California, Riverside
Riverside, CA 92521
U.S.A.
siewert@ucr.edu

Copyright (c) Charles Siewert 2004

PSYCHE 10 (3), November 2004

KEYWORDS: Phenomenal Consciousness, Natural Function, First-Person Warrant, Introspection

REPLY TO: Dretske, F. "First Person Warrant: Comments on Siewert's The Significance of Consciousness." *Psyche* 7(11) April 2001. <http://psyche.cs.monash.edu.au>

ABSTRACT: While finding substantial agreement with my views, Dretske puzzles over how I could take my anti-functionalist blindsight argument to count against his natural representation view of phenomenal consciousness. I explain why I think this, and discuss also a distinct, but related source of conflict between us, concerning whether phenomenal consciousness is specifically a sensory matter. Finally I discuss Dretske's remarks regarding how we know we have conscious experience, offering an outline of my own approach to the topic of introspection.

1. Introduction

I want to thank Dretske for his generous comments, and for his accurate summary of many major points of agreement between us. I am grateful for this partly since I am glad that someone who has thought about these matters so much can find substantial common ground with me (even though, as he recognizes, there are important disagreements). But also I feel fortunate to receive critical remarks so unprickly and fair-minded—especially since I did not give Dretske's views the detailed attention they deserve in my book, a fact to which he rather good-naturedly alludes. Here I can only ask for much the same pardon as I would ask of other symposiasts with whom I have found myself in a similarly embarrassing situation.

In any case, I will try here to atone by putting my views into greater dialogue with Dretske's. My remarks will be organized as follows. First, I will clarify my stance on

whether my blindsight-based argument against standard functionalist theories extends to Dretske's "natural representation" account of consciousness. Second, I want to discuss an important and related area of conflict between our views he does not bring up, but which I think needs to be mentioned in comparing them, having to do with whether phenomenal consciousness should be viewed as a specifically *sensory* affair. (This concerns the issue—discussed in my Chapter 8—of phenomenal thought. (All page and chapter references are to Siewert 1998.) Finally, I want to say something about an issue Dretske does directly raise in his remarks, regarding first-person warrant and introspective knowledge. I will try to relieve at least a little of the puzzlement he expresses about my views on that topic, by summarizing some ideas I have been developing since the publication of my book. Along the way, I raise a few questions of my own about Dretske's views.

2. Natural Function and the Blindsight Argument

As Dretske sees, a great deal of my book's argument hinges on my discussion of a hypothetical case of blindsight, on the basis of which I argue that a number of would-be accounts of consciousness actually promote its intellectual neglect. He sees this argument as helping to dispose of theories that would, by his lights, fail to give due consideration to the idea that one might have an artificial, functional near-analogue of consciousness in the absence of the genuine article. So Dretske says that he can well enough imagine someone (like my hypothetical blindsighter Belinda) who has no conscious visual experience of a given stimulus, but is nonetheless prone to make spontaneous, if crude, verbal and other behavioral discriminations of it, equivalent to those afforded by extremely blurry, low acuity phenomenal vision. To this extent he gladly accepts my contrast between the situations of my characters Belinda and Connie (pp. 105-6). However, he finds himself drawn up short, if in addition he is asked to suppose that the part of Belinda's visual system active in her blindsight is *functioning as nature designed it*. To try to imagine such a thing he finds absurd. Much as there just seems to be nothing else one could want for Belinda to have a genuine *heart*, provided she is blessed with a naturally functioning blood-pump, so there seems to be nothing else one could want for her to have *visual experience*, once she has naturally evolved light discrimination organs that work as well as hers. Thus, for Dretske, reflection on Belinda's blindsight serves to reinforce his view of consciousness. And he wonders how it could fail to do the same for me.

I will explain this, but first I want to be a little bit more precise about just what I would ask us to conceive of here. I do not require us to suppose that Belinda's blindsight visual system is working *just fine* with respect to its natural function. Belinda is described as having normal visual response in her right field, while having only an island of much lower grade blindsight discriminatory ability in her left field. It is indeed at least hard to imagine how a visual system like hers, with its dramatic difference in left and right field visual range and acuity, could be functioning perfectly naturally.

Let us grant then that the visual system that supports Belinda's form of blindsight is inevitably degraded somehow with respect to its natural function. Still, what is not clear to me is that hers must be *more seriously* degraded in this respect than Connie's "blursight" visual system. It would, I think, be natural to suppose that both Connie's

lousy conscious vision in her left field, and Belinda's near-functionally equivalent left field blindsight, are performing at a level below that for which nature made whatever physiological structures render them possible. The question, however, is whether one of them must be a *bigger* failure by nature's standards than another, in the way Dretske requires. I don't see that one of them must. The relevant natural function of Belinda's visual system is to enable her to discriminate the presence, shape, and position of left field light stimuli. Admittedly, it doesn't do this very well—but it does it as well as Connie's. (Note: we could, of course, say that Belinda's visual system fails and Connie's succeeds at their natural function of *creating left field visual consciousness*. But we could not use the fulfilment or non-fulfilment of *that* natural function to define the difference between visual consciousness and its lack without a trivializing circularity.) Now, I can indeed see how the question of whether something is or isn't a *heart* can hang on this question of natural function—to the extent I grasp that notion. But even so, the following seems to me quite conceivable. Things on Belinda's left don't look any way to her, and so her visual system does not perform its job of generating that sort of phenomenal experience. But it is in *other* respects on a par, in producing the left field discriminatory responses for which nature “designed” it, with the visual system of Connie—to whom the corresponding stimuli, in her left field, present a very blurry, achromatic appearance.

Here then is the problem. Suppose that (as Dretske is willing to accept) whatever ineliminable manifest functional differences we must think persist between Belinda and Connie, these are unsuitable to constitute the difference between visual consciousness and its absence. Then it seems to me that, as long as we have no adequate grounds for saying their situation, though conceivable, is metaphysically impossible, my blindsight argument will threaten not only the kind of standard functionalist theory of consciousness Dretske agrees it defeats—it will threaten his view as well. For this case in which there is no manifest functional difference between the subjects, out of which their experiential difference can be constituted, can also be thought of as a case in which they are the same with respect to how *natural* their shared visual functioning is. This is how I put the matter in my book:

Both subjects would, I suppose, on a theory like [Dretske's], have internal states whose function it is to indicate the presence of light on their left. And if both are creatures of evolution, I see no reason to suppose they would not both be carrying out their *natural* functions in this regard. Both Belinda's blindsight and Connie's vision would be, we may suppose, abnormal or “unnatural” in some ways—but in neither case would the natural capacity to indicate light in their left field be broken. (p. 145)

Here perhaps Dretske will want to press this point. “But once you've got such a parity of natural function between Belinda and Connie, just what *extra feature*, necessary for boosting Belinda's blindsight up to the status of phenomenally conscious visual experience, are you supposing Belinda still lacks? Surely there must be some such crucial missing ‘extra’ you can name, or else the scenario is unintelligible. And yet you have now robbed yourself of the only plausible candidate feature by which to distinguish them here, when they are so manifestly functionally similar.” But my response to this would be that the “extra feature” I am supposing Belinda to lack is simply this: the feature of its *looking some way to her in her left visual field*. It seems Dretske assumes that I can

conceive of Belinda's situation, only if I can name a feature, identifiable in some *other* way, which I have reason to believe supplies a necessary condition that, when added to Belinda's talents, would give her the visual experience I am supposing Connie to have. But I don't see why I need this.

Let me clarify this point a little further. I don't need to be able to name a further consciousness-supplying difference that I have reason to think would separate Connie and Belinda. In saying that, I don't mean to commit myself to saying there need not *be* any such otherwise specifiable difference between two such subjects. For, as far as my point here goes, I could accept the view that the phenomenal supervenes on the physical. In that case, if Belinda differs experientially from Connie, then there must be *some* incontrovertibly physical difference between them as well. What I would deny, however, is that I would have to say more *specifically what that difference would be*, and claim reason to believe *that* difference would make the relevant experiential difference. So, for example, I don't think I have to concoct a neurophysiological theory that would explain how Belinda's and Connie's brains could be organized and function so as to give rise to the imagined experiential difference between them, merely in order to conceive of Belinda and Connie differing in the way described. That *would* be necessary, if I sought reason to believe this blindsight scenario is nomologically or naturally possible. But that, I take it, is not what is at issue here. Now the question is this. Assuming, for the sake of argument, that there would have to be some other specifiable physical difference between Belinda and Connie, for there to be the imagined experiential difference between them, do we have any reason to think it would have to consist in whether certain of their visual systems' representational powers were or were not rooted in naturally selected indication functions? Thought experiment just cannot, in my view, help us to establish the necessary connection between visual appearance and naturally evolved representational function that Dretske wants. In fact, it seems to me to count against the idea that such a function would be necessarily sufficient to distinguish Belinda and Connie.

Perhaps it would additionally clarify my disagreement with Dretske if I talk a little about how similar issues come up when we raise questions about experience, or the lack thereof, in other animals. Dretske mentions bees. I agree we may suppose that bees' vision allows them genuinely to see colors, in a sense in which some hypothetical blindsight for color detection would not. So it is possible, even natural, to think that colors really look some way to bees. We can say this even if we demur from claiming we can accurately conceive of what it's like to have the color vision of bees, or judge whether colors look the same to them as they do to us, or imagine the world from a bee point of view. However, I can also entertain the notion that bees detect and respond to differences in light reflectance in ways that don't involve colors looking any way to them at all. That is, I think I can entertain the idea that bees are little zombies. And if you say, "Oh, so you must be thinking that bees turn out to be robots, artefacts, not creatures of nature," I will reply, "No, why do you say that? I don't see why there couldn't be some naturally evolved zombie vision, or relatively high grade blindsight that enabled creatures to do what bees do." So here I find, in a slightly different way than with Connie and Belinda, that we do not close the conceptual gap between a mere functional analogue of visual appearance, and the real phenomenon ("real seeming" in Dennett's phrase), by adding that the function is a function of indication prescribed by nature. In the Belinda and Connie case this will be evident, if we find no obstacle to supposing Belinda's

blindsight system is fulfilling the natural function it shares with Connie's vision. In the bee case, this will be evident, if we find no obstacle to supposing that doing what bees do naturally may involve no more phenomenal color vision than blindsight.

Here again, in saying this about the bees, I can concede that, if bees, though creatures of evolution, aren't phenomenally sentient, then this must be in virtue of some otherwise-specifiable physical lack. But I don't have to claim to know what that difference would be. And, although we may somehow discover that certain physical mechanisms make bees phenomenally sentient, I don't see what reason we have to believe the evolutionary etiology of these need be considered essential for them to be consciousness-making. Anyway, it doesn't seem to me that thought-experiment can secure that result—or really that it can tell us much of anything about what physical differences make for phenomenal ones.

I think it would help motivate my position here to see how other minds questions about bees can arise for us. I can apparently entertain the notion that artifacts of various kinds, including robots equipped with sensors, are not phenomenally sentient. I can, I think, even entertain the notion of robot-bees that behave in many respects like regular bees, though they are not phenomenally sentient. And once I do, I can pose to myself this question: are *actual, natural* bees phenomenally sentient? Do colors look anyway to them, or are they no more phenomenally conscious than the hypothetical robot-bees? This seems to me an intelligible question. And I don't think it is resolved by pointing out that the sensory mechanisms of the natural bees are functioning in accord with evolutionary design. I can still entertain a rational doubt about whether the bees are buzzing with consciousness, or are just as phenomenally vacant as the hypothetical robot-bees, *even once I fully recognize the bee is functioning as nature prescribed*. For I can entertain the hypothesis that there are naturally evolved non-phenomenal perceptual systems.

When I talk of zombies, robots, and the inner life of bees, I have some fear this may provoke a glib dismissal of the sort that often greets philosophical discussion. Someone might say: "We are after a serious scientific understanding of consciousness here, and you want to distract us with all these absurd, far-fetched possibilities and armchair speculations. Next we will probably be asked to torture ourselves about how we know we're not brains in vats." However, I do not fear Dretske will react in this way, as he shows himself ready and willing to recognize the relevance of contemplating mere possibilities in theorizing about consciousness. And well he should. For such considerations are relevant to assessing a would-be explanation of consciousness as long as it commits itself to telling us what consciousness *is* (and not just what conditions actually give rise to it), and thereby asserts strong claims of impossibility, which surpass what can be established by consulting the actual course of nature alone. If a theory claims to explain consciousness in terms of some feature or features F, by saying that anything with F must have certain conscious states as a matter of some greater-than-nomological necessity, then the conceivability of F occurring *without* such conscious states is not irrelevant to the theory's justification. For the conceivability of F in the absence of consciousness presents a challenge to that theory; it needs to justify the claim that F without consciousness is impossible in some more-than-nomological sense, in the face of this apparently contrary evidence. But now, once we admit the relevance of the merely

conceivable for the study of consciousness, I think it is hard to deny the intelligibility of a contrast between what—like ourselves—genuinely experiences looks, smells, feels, and tastes, and what at most merely acts similarly to what does. Anyone who tries to deny this distinction will come to seem a denier of the very phenomenon of consciousness, trying to return us to the dreary obtuseness of behaviorism. Likewise, skeptical other minds problems that arise in this context are real: there are difficult questions about how we can know which organisms are (and whether various kinds of artifacts would be) phenomenally sentient, and theories of consciousness that make these questions seem empty, confused, or easily resolved, only bring doubt on themselves by doing so.

Dretske may not have a quarrel with these last points. And yet he is apparently confident that the process of reflection that leads us to them will also lead us to recognize the strong necessity with which the line between what *really feels* and what *merely acts in ways inviting that interpretation* coincides with a distinction between what performs its *natural function* and what does *not*. But here is where I have trouble. For I find that the reflection that allows me to appreciate the contrast between real sensory appearance and some close functional analogue does *not* lead me to recognize any necessity that this distinction be made out in terms of whether each, in doing its job, fulfills evolution's natural design. If a natural function to indicate is supposed to be necessary—and necessarily sufficient—to distinguish Connie's visual consciousness from Belinda's blindsight, this necessity is just not evident to me. Dretske jokingly suggests at one point that perhaps I just have a bigger imagination than he does. In the same spirit, I might wonder whether Dretske has a stronger necessity detector than I do.

But the real reason for our difference, I suspect, is that Dretske assumes that, whatever it is to be visually conscious of a stimulus (whatever it is for it to *look* somehow to one, for one to have a *conscious visual experience* of it), this is something that *must* occur once one builds up enough conditions out of a set concerning the relations among an organism's environment, its interior, its behavior, and its evolutionary background. But I don't see what entitles us to that sort of assumption. And, when I reflect on the case of Belinda and Connie, this seems to give me reason to think it is *no more than a contingent fact* that naturally evolved powers to respond to light, shape and motion as well as someone like Connie would in her left visual field involves *visual consciousness* of light, shape and motion—involves distal stimuli *looking* some way to an organism. And when I reflect on the case of bee vision, this seems to give me reason to think it is no more than a contingent fact—if it is one—that the exercise of a naturally evolved capacity, of the sort bees have, to respond to light involves the visual appearance of color. Maybe things do, as a matter of fact, look colored to bees. But I don't see why things just *have* to look colored to bees, because of what bees naturally do around them.

3. Natural Function and Phenomenal Thought

In discussing my differences with Dretske, I have been content so far to couch questions about phenomenal consciousness as questions about the character of sensory experience—about what we report by “appearance words”—such as ‘looks,’ ‘smells,’ ‘sounds,’ and so on. But this obscures a fundamental difference in our perspectives which, though Dretske himself passes over it, needs noting here, partly because of its

close connection to the important role played in his view by the idea of natural function. But it will take a little doing to explain this difference—so please bear with me.

Dretske's view of phenomenal consciousness is sometimes styled a form of representationalism. To the extent that this means we are to regard the phenomenal character of sensory experience as inherently intentional, as containing no layer of raw sensation—a “stuff” of experience that must be “shaped” or interpreted somehow, before experience can be said to be accurate or inaccurate experience of where things are—to *that* extent, I would agree with representationalism. But in Dretske's and others' writings it means more. As I understand it, it is committed to the notion that one can say what special type of representation experience is, in a way that will explain why it has phenomenal character. But to that end, it will not be adequate to say what kinds of representations phenomenal experiences are simply by appeal to appearance words like ‘look’ and ‘feel’—that would do nothing to explain why things look or feel to us as they do. So, for the purposes of providing such an explanation, one needs to specify the kind of representation that phenomenal conscious experience is, without simply appealing to the phenomenal appearance concepts we initially use to identify what is to be explained.

One strategy for trying to do this involves confining the range of mental life deemed properly phenomenally conscious to what is distinctively *sensory*. One might say: states of its looking, smelling, or feeling some way to someone—along with whatever correspond to these in the realm of sensory imagery (e.g., visualizing, and “hearing” sounds “in one's head”)—these are the *only* states that have phenomenal character. So the representationalist strategy can focus on what is common to these. States more intellectual or ratiocinative in character—what I will here call *cognitive states in the narrow sense*, do not exhibit phenomenal character of their own; they are not themselves phenomenally conscious, but are at most accompanied by imagery that is. Cognitive episodes that are not properly phenomenal would then include cases of *thinking*, in the sense in which differences in *what* one thinks are not differences in what one *visualizes* or otherwise *images*. More specifically they would include cases of *wondering whether* something is so, *judging* it is so, or *concluding* that it is, *considering* a suggestion, or *raising or entertaining a question*, *doubting* whether something is so, and *recalling* that something is so.

Now it seems to me important to Dretske's brand of representationalism to confine the phenomenal to the sensory, so as to keep out the sorts of cognitive episodes just illustrated. This is because of the important part played in his account by the distinction between mental representations that indicate what they do by way of fulfilling their natural, evolution-bestowed, function, and representations that represent what they do rather on account of a process of learning. If one wants to employ this distinction, then one will have to allow that what I have been calling cognitive states do not, for the most part, represent what they do because evolution has given them the function of naturally indicating those things. If I am thinking about whether to buy a hybrid car, no one will say that this occurs because some part of my brain is performing its natural function of being a *hybrid car indicator*—much less that a structure in my brain is active, which had been selected by evolution to represent Toyotas or Hondas.

Once this is granted, it becomes important to face the question of whether the reach of phenomenal consciousness really does exclude cognitive mental life in the

requisite manner. In Chapter 8.3-6, I have argued in detail that it does not. This bears on the discussion at hand, for if the cognitive is not kept out of the phenomenal, we cannot expect to get an account of phenomenal character from a natural function account of mental representation, as Dretske proposes.

Here I want not only to draw attention to how critical this issue of phenomenal thought is to assessing theories of consciousness. I want also to make a plea that this question be considered fairly and patiently. There is a danger that even before the issue is truly engaged, one will have dismissed the idea as preposterous: “Are you saying there is a special qualitative feel—a kind of cognitive itch—specific to all thoughts about rabbits, a ‘rabbity feel,’ perhaps? Well, I certainly can’t report any such thing!” But to judge the matter fairly, we must slow down here, and allow that this question is not to be settled on the basis of simple and immediate inspection of an example, and by noting how like or unlike thoughts are to sensations like itches. Unexamined, dubious background assumptions and misleading questions and comparisons can make such a seemingly direct approach quite unreliable. The case that phenomenal character exhibits rich variation beyond anything that can be confined to the sensory domain, a variation inseparable from distinctively conceptual thought—the case, as we might put it, for *phenomenal thought* (or as others have said “cognitive phenomenology”), should be approached indirectly, from consideration of a variety of examples and their implications, and with a critical sensitivity to the assumptions behind the questions we ask. One way of doing this is laid out in Chapter 8 of my book. (Some other detailed discussions favorable to my thesis are found in Horgan and Tienson 2002, Pitt 2004, and Strawson 1994.)

I would not—and certainly should not—try here to recapitulate that earlier course of reflection. But I would urge that anyone who is attracted to a form of representationalism that purports to explain phenomenal consciousness as specifically a sensory affair needs to think about the examples and questions I discuss there before reaching their conclusions. It is important not to do this hastily. For example, when I point out that there is a difference between the way it seems (a difference in “what it is like”) for you to hear an utterance *with understanding*, and the way it seems to hear the same type utterance *as a mere pattern of sound*, it won’t do to think only of the contrast between how speech sounds in a foreign tongue before one has learned to interpret it, and how it sounds when later one understands what is said. For then one might think it’s enough to say (as I have heard Michael Tye say in this connection) that we have here simply a difference between the way speech sounds when it appears as an unbroken phonetic stream, and its sounding differentiated into recognizable units—words and phrases. That may seem to keep us safely in the realm of the strictly sensory. But the kind of contrast to be considered also can be appreciated even where one already understands the language in question. When one reads a text in such a language to oneself or recites it aloud, while ceasing to follow or understand its meaning, there will not always be a problem with perceptual recognition of patterns and units of sound. And yet this experience differs from the experience when one re-reads the passage so as to follow and comprehend its meaning. (For a more detailed discussion of this specific point, see pp. 275-6.) This is just one illustration of how we need to take care, in considering examples pertinent to addressing the issue of phenomenal thought.

In addition to pleading for such care in considering the kind of examples I and others have adduced in connection with this issue, I would suggest also that we ventilate a little skepticism about what basis we have for assuming (if we do) that examples of sensations (e.g., pains, itches), or the experience of a thing's "sensible qualities" (colors, sounds, odors, flavors), examples which serve as such vivid paradigms of differences in phenomenal or qualitative character, completely *exhaust* differences of this kind they *illustrate*. Differences in the way things look in cases of Gestalt shifts in drawings or figures also furnish perfectly good examples of differences in phenomenal looks we illustrate with appearances of colors. And yet it is unclear that these are differences in what "sensible qualities" of objects are being sensed. At least it seems we should admit Gestalt qualities among the kind of qualities pertinent to the "qualitative character" of experience. But now, once we admit that these kind of *interpretive* differences (between taking the figure as duck or rabbit, vase or profile, M or lazy sigma, bisected diamond or pyramid) are contrasts in phenomenal character, we may be less likely to assume that something has to be very like *feeling a sensation* to have phenomenal character (so as to conclude that since the occurrence of thought is not *very like that*, it has none). The sudden occurrence of a new thought may indeed strike us as rather more like the experience of a Gestalt switch—but that doesn't suggest it is not phenomenal, for Gestalt perception should be as much for us a paradigm of phenomenal consciousness as the onset of an itch. Furthermore, notice that even if we do guide our conceptions of phenomenal character by "feeling" talk, this can help us to dissolve, not reinforce, the barrier that would keep the cognitive out of the phenomenal. For we may not only speak, in connection with our *sensory* lives, of feeling pain or itches. We also may speak, in connection with our thinking, cognitive lives—of *feeling confident* or *sure* of what we think, upon reflection; similarly we may speak of *feeling doubtful* or *uneasy* about what we think. Is it clear that only the former (sensational) "ways of feeling" are to be classified along with differences in the look of a Necker cube, as genuinely phenomenal differences, while the latter (cognitive) "ways of feeling" are not?

Here I am only raising suggestions and posing a few pointed questions—a decision about what to conclude from them calls for more sustained thought. But what I want to do in this section of my reply to Dretske is just to motivate this concern: When considering representationalist accounts of phenomenal consciousness, such as Dretske offers us, we need to take seriously the objection that, by conflating the phenomenal with sensory appearance, they mistakenly suppose that by distinguishing the sensory from the (narrowly) cognitive, we find what makes the phenomenal phenomenal.

4. First-Person Warrant

When, in his comments on my book, Dretske raises questions about my view on the source of first-person warrant, he worries a little that he is "dragging me somewhere I don't want to go." But now, having just insisted on dragging him into the topic of phenomenal thought, I'm not in much of a position to complain or beg off. Anyway his epistemological issue is hardly one I can responsibly avoid. So I will now try to relieve his puzzlement somewhat, by laying out the elements of my own view, such as it currently is. But first let me register a little puzzlement of my own about his views.

I am puzzled by his claim that we don't have a distinctively first-person warrant for what *I* think we have it for. Here is what I think we have. I think we have a warrant for first-person beliefs and judgments (e.g.) that something looks to us a certain way, that we feel some way, that we are thinking of something or other—a warrant for these that differs in kind from that available for corresponding second and third person beliefs/judgments. So, for example, you see a look of concern cross my face and ask, “What’s the matter?” I say, “I was just wondering whether I locked the car before we came here.” Now I already believed and knew that I was wondering about this even before I answered your question. And I had warrant for believing that. But then evidently it was a warrant distinct in kind from the sort available for third-person beliefs about my state of wondering—since none of that was then on offer.

Dretske says he disagrees with me regarding what we have “special warrant” about, and expresses this by saying roughly that, on his view, we have special warrant regarding *what* we think, not *that* we think. But if this is opposed to my view, it must come to saying that (for example) I don't have a distinctive warrant for first-person judgments that I am (or was) *wondering whether I had locked the car*—but only a distinctive warrant for judgments regarding *something that I'm wondering about* (whether I locked the car, perhaps?). But then, I don't see what reason Dretske has for thinking that cases of solitary self-knowledge like that just invoked (and those discussed in more detail in Chapter 1.7) don't show what I say they do. I also have some difficulty understanding what special access he thinks we do have, and why he thinks we have it.

Consider his example of a small child who thinks that Daddy is home. Dretske denies she has special access to the fact that *she believes that Daddy is home*. (She isn't even able yet to form beliefs about her own (or anyone else's) beliefs.) However, in some sense consistent with this denial, she is said to have special access to *what she believes*. If this is consistent, I suppose it must mean something like: she has access to *Daddy's being home*, or to *the fact that Daddy's home*.

But now, first, I don't see how such an example conflicts with my thesis of first-person warrant—if it's intended to. It is quite consistent to claim, as I do, that those who *do* have the wherewithal to form beliefs about their own wonderings, believings, and so on, have a distinctive warrant for these beliefs, and *also* to claim (as I am happy to admit) that two-year-olds don't have this warrant for such beliefs about themselves (since they simply don't even have the beliefs in question). (In this connection, see my discussion of Gopnik's research, pp. 43-46.) An additional question: why are we supposed to think the child has special access to *Daddy's being home*? Partly my puzzlement here is with what is supposed to be distinctively first-personal about her warrant for believing Daddy's home, if that is the idea. (After all, her warrant, we may suppose, is in a sense of the same sort as her brother Bobby has for *his* belief that Daddy is home—they both see his car in the driveway, or some such.) Maybe someone will say her warrant is specially first-personal in that it is based *on her own experience* of Daddy's car (not Bobby's experience). I doubt that is what Dretske had in mind. Anyway, saying that would be perfectly compatible with my thesis of first-person warrant.

I am puzzled why Dretske did not state his view somewhat differently. I might have thought he wanted to say—not that we have no distinctively first-person warrant for beliefs about our believing, wondering, seeing, feeling, etc.—but rather that the

distinctive warrant we have for such beliefs is somehow derived from a *more basic* warrant we have regarding things *other than our own states of mind*. So then: it's not that I don't know I see an apple in a way different from that in which *you* know I see an apple. Rather, my special manner of knowing this is based on my more primary knowledge of the apple itself in a way yours is not. I know that it looks to me like there's an apple, not by somehow directly perceiving its looking that way to me (whatever that could mean)—but rather by seeing the apple there. This is supposed to be, I think, a case of knowing about one thing (my state of mind) by perceiving something else (the state of the world) somehow like that in which I can know about the state of my car's fuel tank (it's _ full) by perceiving something else (the fuel gauge). And, this fuel-gauge-like way I can know my own state of mind by knowing the state of my surroundings will not work for knowing the state of *anyone else's* mind. That's why this is an account of what's special about first-person warrant or authority. At least, that's the picture I had of the account from Dretske's other writings. But maybe I've misunderstood them. Or maybe I've misunderstood his remarks here in *Psyche*. Anyway, I am still not clear on why the sort of Dretskean picture I've just sketched should seem incompatible with my thesis of first-person warrant.

One thing Dretske and I agree on in this area; as he notes, I, like him, cannot accept anything like the idea that I know that I have a visual experience of an apple by (directly, introspectively) “seeing” my experience of an apple. Still, I am not clear on just how it is he thinks that I know that it looks like an apple to me, by seeing (not the experience but) *the apple*. I take it the idea is not that, by seeing, I know there is an apple, and then from the fact that there *is* an apple, I *infer* that it *looks* to me like there's one. But then just how does this work?

Maybe this is not the place to inquire further into Dretske's views about self-knowledge. In any case, I still owe him some kind of an answer about what my *own* (developing) view is on the topic. So let me now turn to that. To explain my position I need to say a little about the difference between what we might call “*first order thought*” (i.e., thought about something other than one's own experience or states of mind) and “*reflective thought*” (i.e., thought the expression of which involves referring to, or attributing to oneself, an experience or other state of mind). Consider first the expression of first order thought. Where *that* is concerned, we can identify objects of reference via their appearance to us. And we can recognize that these objects change in their *appearance*—change in how they look, sound, feel, etc., to us—even while they are *otherwise unchanging*. This makes it possible for us to distinguish sensory appearance from our *thought or judgment* about things that appear to us, and to understand the possibility of a *distinctively sensory error*—an illusion or hallucination (as opposed to a false judgment or belief). That then allows us to consider the possibility of a sensory error so radical as to give us a putative object of appearance, hence of thought, where in fact none exists. (This would cover certain kinds of holographic illusions, as well as extreme forms of hallucination.) So we arrive at the possibility of thinking of something, when there exists no object of which we are thinking, such that anything we are inclined to predicate of it is true of it. (If I am subject to a holographically induced illusion, and it looks to me like there's a coin there, where there actually isn't one, I may think there is a coin there—even say intelligibly ‘I think *this* is a coin’—though nothing is there to which my utterance of ‘this’ refers.)

Now contrast this with the case of reflective thought—in particular, thought about ways in which it appears to one. Where the objects of reflective thought are concerned, we *cannot* recognize the occurrence of changes in *their* appearance, while they themselves are otherwise unchanging. We cannot, I maintain, distinguish a category of higher order “appearance” or “sensing” here at all, as distinct from reflective thought or judgment. And so there are neither correct nor incorrect sensings of one’s own experience—there are not, properly speaking, sensings *of* one’s own experience at all, if the ‘of’ here is taken to be the ‘of’ of intentionality. There is no such thing as inner sense.

However, on my view, in order to understand what has been called introspection, we not only need to take full cognizance of this difference, which makes the very label ‘introspection’ so misleading; we must also take heed of special forms of reflective thought, whose occurrence nonetheless contributes to the *appeal* of inner sense doctrines. For this, we need to recognize that we are able to think thoughts about our own experience, and modes of sensory appearance, which are of a sort that invite various forms of demonstrative expression—for example: ‘this feeling’; ‘the way this looks to me’; ‘The color (or shape) this looks to me’; ‘The way this tastes’; ‘This taste’; ‘This way of tasting’. Furthermore, we need to see that thoughts susceptible to such irreducibly demonstrative expression play a special role in “anchoring” reflection about experience. Through thinking them one can consider various ways of classifying one’s experience, recognize their aptness, or reject or correct them. (So, I may first notice a peculiar smell, then wonder, of this way it smells to me: *what is that way?* I consider calling it ‘buttery,’ and then reconsider—it smells more like *peanut butter* to me, although not so *much* so that I think what now smells that way to me actually *is* peanut butter.) I think it is fair to regard the special sort of reflection involved in such consideration of one’s olfactory experience as the exercise of a special form of *cognitive* attention—where cognitive attention more generally encompasses all ways of attending to something just by *thinking of or about* it, as opposed to attending by looking, listening, touching, etc. The particular kind of cognitive attention just illustrated can be directed to experience, and modes of appearance, generally. (Thus I would affirm—as against some expressions of the doctrine that consciousness is “transparent”—the feasibility of attending to one’s own experience.) As this form of attention in some way enables one to make one’s experience into the object of diverse thoughts, it may seem analogous to *sensory* attention—though again, I urge us to resist this analogy by noting the absence, in the case of objects of reflection, of anything corresponding to sensory appearance.

The crucial point to recognize then is this. Because *there is cognitive attention* to experience, we can understand what we’re talking about when, expressing the thought constituting such attention, we say ‘the way this feels to me,’ ‘the way this looks to me,’ and so on. But because *this involves no inner sense*, we cannot (as we can with sensory appearance and first order thought) make sense of the possibility that, even so, *nothing* we are inclined to say about what we’re thinking of when we attend in this way is, in fact, true of it. There is here no conceivable analogue of the case where illusion or hallucination gives us something to think about only falsely. Thus, if the expressions of our acts of cognitive attention are intelligible to us, they find genuine referents that we cannot conceive ourselves utterly helpless to classify correctly.

Now this, I propose, is key to understanding first-person warrant. For given all this, then, if we are entitled to think we do know what we're talking about, in talking as we can in expressing thoughts constituting the sort of cognitive attention illustrated, we are entitled to presume we have some correct way of classifying the objects of such thoughts—our own experience. And then, I would argue, since we are indeed entitled to think we know what we're talking about, in the absence of special reason to doubt (for we are entitled to such a presumption of semantic self-knowledge), we are entitled to a presumption of correctness in our attentive classification of our own experience. Hence I would maintain we possess a distinctive and strong—though defeasible—form of warrant for judgments about our own experience—distinct in kind both from what we have for first-order judgments about our surroundings, and from what we have for second and third person judgments about experience.

Dretske asked after my account of our first-person knowledge of our conscious experience, and here it is. It offers a way to address the question he raises about how one is in a special position to think with warrant about one's own conscious states, but not someone else's. For the form of cognitive attention I have identified that puts me in a position to think with warrant about my experience, enables me to understand what I am thinking of, when I have thoughts expressible by such formulations as 'the way this looks to me,' 'this feeling,' and so on—but not via any appearances of the referent (since there are none)—nor indeed via any appearances at all whose falseness might affect my warrant for the thoughts I have about the referents of these expressions. But you, by contrast, are not able to exercise such a form of cognitive attention to understand what I am talking about when I say 'the way this looks to me,' 'this feeling,' and so on. Nor are you able to do so, to understand your reference to the same things by phrases like 'the way it looks to you,' or 'that feeling,' and so on. You can only understand these references, via appearances of *me*, my acts, and my effects—appearances whose falseness *could* affect the warrant you have for your thoughts about the (putative) referents. Now the form of cognitive attention I am speaking of here can be directed *only at conscious states* (for expressions of this form of attention in phrases like 'the way this looks to me,' 'this feeling,' etc., have phenomenal appearances, hence conscious states, as referents, if they are intelligible at all). And it can be directed *only at one's own conscious states* (for one cannot conceive it to be directed at *another's* experience without ceasing to distinguish that experience from one's own). Thus, the warrant it brings to judgment is based in consciousness, and is a uniquely first-person affair.

Of course, this account needs a lot of filling out, argument and clarification. And I need to confront the task of bringing together what I have said about phenomenal thought with what I am now saying about first-person warrant, so as to show something like my account of first-person warrant regarding sensory appearances applies also to what I might call *cognitive* appearances. Now perhaps Dretske can see why I said I thought providing a positive account of first-person warrant is not quickly done. While it would be silly of me to pretend what I've just said constitutes a fully adequate answer to Dretske's question, I offer it to show that I do have something to say in this regard, which would distinguish my account of introspection both from something along the lines of his "displaced perception" approach, as well as from an "inner sense" theory.

Pieces of this account can be found in a couple of papers I have published (Siewert 2001, 2004). But I would no longer endorse everything I say there, and my account of first-person knowledge is still a work in progress. I am now trying to be a little more optimistic than I was when I said in my book (as Dretske quotes) that it may take several more books to get it spelled out. Maybe *one* more could do it.

5. Conclusion

Admittedly I have gone on at some length in replying to Dretske's comments. But I have some legitimate reason for this. I did not want simply to give minimally responsive answers to his questions or rebut criticisms implicit in them. I wanted to use this occasion to try to learn something about the geography of the issues, by situating my own position vis-à-vis Dretske's, and discerning major sources of agreement and discord. His essay has helped me to understand better the common ground. We both, I think, are deeply dissatisfied with theories that would hold that the difference between having and lacking phenomenal experience is exhausted by manifest functional roles, and we are both convinced that sensory experience is, in a certain sense, richly significant (or, as he would be happier to put it, richly representational).

However, in Dretske's case, the fault in standard functionalist thinking is to be remedied by appeal to a notion of natural representational function, adequate also to account for the inherently intentional nature of phenomenal character. Here we differ, since I think my blindsight argument, which Dretske agrees counts against his functionalist opponents, also has adverse consequences for his appeal to natural function. Dretske seems ready to allow that Connie (with her experience) and Belinda (with her lack of it) could, nonetheless, match one another in terms of manifest functional roles to the degree I say. But then I don't see why one of them must be thought to satisfy the relevant role more in accordance with nature's design than the other—though it appears that Dretske's account requires this. His discussion of bees brings out a closely related argument. While bees may, for all I know, be phenomenally sentient, it also seems possible to me that something with the light responsive powers of a bee could be a visual zombie or a blindsighter, and that such powers could have arisen by a process of natural selection. But it appears Dretske wants to say that this just couldn't be the case. However, I don't see why. And further, I wonder just why, if he is right, an apparently intelligible skepticism about bee consciousness is not laid to rest by the application to bees of evolutionary theory. It may be laid to rest by *something*, but I don't think that does it.

Another source of conflict between our views is to be found on the topic of phenomenal thought. Here too though, the question of natural representational function is involved, since it seems Dretske's use of this notion in a theory of consciousness requires us to keep the phenomenal confined to the sensory side of a sensory/cognitive divide, as I would not do.

Finally, there is the matter of introspection. Though Dretske and I are united in opposing the idea of inner sense, and though we both perhaps somehow try to extract a positive account of the mind's self-knowledge from the very absence of what traditionally has been thought its provider, my way of doing this requires recognition of some kernel of insight hidden in the inner sense doctrine—to which Dretske seems much

less inclined to concede even a dim apprehension of truth. But on my view, getting rid of even any vestige of inner sense does not leave us with enough to work with.

This undoubtedly is not all I have to learn from Dretske. But at least I can say it is more than I had learned when I wrote my book.¹

References

Horgan, T. & Tienson, J. 2002. "The Intentionality of Phenomenology and the Phenomenology of Intentionality," in D. Chalmers, ed., *Philosophy of Mind: Classical and Contemporary Readings*, Oxford: Oxford University Press.

Pitt, D. 2004. "The Phenomenology of Cognition, Or *What Is It Like to Think That P?*" *Philosophy and Phenomenological Research*, 69/1:

Siewert, C. 1998. *The Significance of Consciousness*. Princeton: Princeton University Press.

_____ 2001. "Self-Knowledge and Phenomenal Unity," *Noûs*, 35/4: 542-68.

_____ 2004. "Is Experience Transparent?" *Philosophical Studies* 117: 15-41.

Strawson, G. 1994. *Mental Reality*, Cambridge: MIT Press.

¹ Special thanks to David Pitt for criticisms and suggestions.