

---

Ronald McIntyre, "Naturalizing Phenomenology? Dretske on Qualia," in Jean Petitot, *et al.*, eds, *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science* (Stanford: Stanford University Press, 1999), pp. 429-39.

---

## Naturalizing Phenomenology? Dretske on Qualia\*

by Ronald McIntyre

Since phenomenological concerns are often perceived as irrelevant to, or even incompatible with, "naturalism" in the philosophy of mind, a naturalistic account of mind that treats such concerns seriously would be most welcome. Fred Dretske's recent book, *Naturalizing the Mind*,<sup>1</sup> promises to offer such an account. The central notion in the book is that of mental representation, and Dretske claims that his "Representational Naturalism . . . provides a satisfying account of the qualitative, the first-person, aspect of our sensory and affective life" (p. xiii).

My discussion has three parts. First, I briefly characterize Dretske's particular naturalization project, emphasizing his naturalistic reconstruction of the notion of representation. Second, I note some apparent similarities between his notion of representation and Husserl's notion of intentionality, but I find even more important differences. Whereas Husserl takes intentionality to be an intrinsic, phenomenological feature of thought and experience, Dretske advocates an "externalist" account of mental representation. Third, I consider Dretske's treatment of qualia, because he takes it to show that his representational account of mind succeeds in naturalizing even the "subjective" features of experience. I argue that Dretske characterizes the notion of qualia in an ambiguous way. I conclude that he succeeds in naturalizing qualia only if qualia are understood as nonphenomenological features of experience and that he therefore has less to say than he thinks about the subjective life of beings like us.

### I. DRETSKE ON REPRESENTATION

Phenomenologists who are interested in the philosophy of mind will welcome some features of Dretske's naturalization project. Whereas functionalism, computationalism, behaviorism, and eliminative materialism, for example, ride roughshod over the features of mentality that phenomenologists emphasize, Dretske seems to place them at the very center of his naturalization project. This impression is furthered by what he claims to achieve: a naturalistic understanding of representation, of introspection (including the nature of first-person authority), of consciousness, and of the subjective qualities of experience.

The central notion in Dretske's account is that of "representation." His strategy for "naturalizing" this notion seems to proceed in two parts: first, to describe a natural property – a property that systems which are obviously physical can have – that shares some of the features of intentional experience and thought; and, subsequently, to suggest ramifications of this property that will ultimately describe full-fledged mental representation as it occurs in conscious experience and thought. My focus will be on the first part of this strategy.

Dretske begins with an account of representation appropriate to such nonmentalistic representations as a thermometer's measurement of temperature or a speedometer's measurement of speed: "The fundamental idea is that a system, *S*, represents a property, *F*, if and only if *S* has the function of indicating (providing information about) the *F* of a certain domain of objects" (p. 2). A system performs its indicating function, Dretske says, by occupying different states corresponding to the different determinate values of the property in question. Thus, if *S* is an old-fashioned thermometer, the different levels of mercury in its tube correspond to different determinate degrees of temperature. The thermometer, by occupying a specific mercury-level state, thereby indicates that the temperature is a certain value.

Importantly, on this account, a system can *indicate* a property without *representing* it. A thermometer represents temperature, rather than merely indicating it, because indicating temperature is a thermometer's *function*: thermometers are designed to indicate temperature. Thermometers also indicate other things that they are not (ordinarily) designed for. Prolonged readings below 0° C, for example, may indicate that sweater sales will rise, but thermometers (normally) do not have the function of indicating sweater sales and thus do not represent that property. This functional feature of Dretske's notion of "representation" makes it an *intensional* notion and marks it as closer to the mental than mere indication is.

Thermometers and speedometers, like most representational systems, are conventional systems. They are designed by human beings for the specific purpose of indicating some property that they are capable of indicating, and only thereby do they acquire the *function* of indicating that property. But Dretske believes there is also a meaningful notion of "natural representation": a system of indication may achieve, via natural selection or learning, the *natural* function of indicating some particular property. A system does so, he says, when the system's indicating that property is useful to the system, the information thereby delivered up is available for further use by the system and so integratable into its higher-order responses to its environment, and the system's indicating that property is selected for by evolutionary development. (Cf. pp. 9-10, 162-68.)

Dretske takes conscious experience to be a species of natural representation, but he does not identify the two. Conscious experience, he claims, requires a conceptual element that is not generally present in natural representation. I shall not develop this extension of the notion of

representation, however, for the features and problems I will be discussing occur at a more basic level.

Even at the basic level of representation, Dretske distinguishes the “sense” of a representation from its “reference.” A representational system, typically, represents some *object* as having the property the system has the function of indicating. The represented object Dretske calls the “reference” of the representation; the property it is represented as having – the “way” the object is represented – he calls the “sense” of the representation.

Let us not pass over this sense/reference distinction too hastily, however. Strictly speaking, on Dretske’s account of representation, it is *properties* that representational states or systems represent. (Recall: “a system, S, represents a property, F, if and only if . . . .”) And in the traditional literature on intentionality, that which is intended or represented in an intentional state (that to which the state is “directed”) is called the “object” of that state. So, in the traditional terminology we should say that, for Dretske, the property represented in a representational state is the *object* of the representation; and this property, since it plays the role of represented object, should be called the “reference” of the representation. Dretske, however, calls the represented property the “sense” of the representation. (Here is a first indication that Dretske’s terms, as he uses them, do not always mean what phenomenologists would most likely take them to mean.)

For Dretske, then, representational systems represent properties; and the property a system represents is called its “sense.” But under certain conditions, there is a derivative sense in which the object that has the represented property may also be said to be “represented.” A speedometer, for example, represents speed, and when appropriately installed in a vehicle it represents the speed of that vehicle. Installed in my automobile, a speedometer whose needle is pointing to 50 represents *my automobile* as traveling at 50 kph. In that same representational state, but installed in your automobile, it would represent *your automobile* as traveling at that speed. Thus, Dretske says, just which object (if any at all) a representational system represents is purely a contextual matter. The “sense” of a representational system – the property the system represents an object as having – is determined by the representational character of the system; but its “reference” – the object it represents as having that property – is determined by “a certain external causal or contextual relation” (p. 24 ).

On Dretske’s account, then, the “sense” of a representation is *representationally* related to the representation: it is the property that the representation *represents*. The “reference” of a representation, on the other hand, is only *contextually*, and not *representationally*, related to it: in particular, the reference of a representation is not *represented* (in the proper sense of the term, as Dretske uses it).

Dretske’s account of representation has the virtue of preserving the possibility of misrepresentation and referential failure. Given the functional account of representation, a

speedometer reading of 50 represents a speed of 50 kph simply because of the speedometer's design and calibration. If something goes wrong and the speed of the vehicle in which the speedometer is installed is actually 60 kph, then the speedometer's reading has *misrepresented* the vehicle's speed. And since reference, for Dretske, is strictly a matter of context, a speedometer can represent speed even though there is no vehicle connected to it in the appropriate causal and contextual way.

## 2. DRETSKEAN REPRESENTATION VIS-À-VIS HUSSERLIAN INTENTIONALITY

Dretske's notion of representation exhibits, at least on the surface, some striking similarities to Husserl's notion of intentionality. "Intentionality," of course, is just Husserl's term for the representational character of thoughts and experiences – their characteristic feature of being "of" or "about" things other than themselves. Like Dretske, Husserl distinguishes the *sense* of an intentional thought or experience (let me use Husserl's term "act" for short) from the *object* (or reference) that the act represents or intends: an act intends or represents its object "as" such-and-such. As in Dretske, an act (or representational state) can have a sense without having a reference or can have a sense that *misrepresents* the reference.

For Husserl, the intentionality of an experience is a *phenomenological* feature of the experience; indeed, he sees it as the key feature that must be explicated if we are to achieve an understanding of experience from the first-person, phenomenological point of view. Thus, given the apparent similarities between Husserlian intentionality and Dretskean representation, one might think that Dretske's naturalization project (to the extent it succeeds) can shed light on how to understand the *phenomenological* features of experience naturalistically. Unfortunately, however, these features do not play a role in Dretske's project. To begin to see that this is so, let us turn briefly to Husserl's version of the sense/reference distinction.

Husserl's distinction is motivated, in large part, by the fact that one and the same object can be represented in various "ways." Oedipus's complex relationship with his mother dramatically demonstrate this fact: that Oedipus desires to marry the queen does not mean that he desires to marry his mother, for example, despite the fact that the queen *is* his mother. As Husserl explains it, these desires intend or represent the same object, Jocasta (or a state of affairs involving Jocasta), but they represent that object with a different sense ("the queen" in the one case, "my [Oedipus's] mother" in the other).

Following tradition, Husserl uses the term "object" for that which is intended or represented in an intentional state (that to which the state is "directed"). Recall that, for Dretske, *properties* are what representational states represent. For Husserl, however, intended or represented objects are not all of the same kind: intended or represented objects include properties but also physical

objects, events, states of affairs, numbers, persons, and whatever else we can bring before our minds.

Unlike the object of an intentional thought or experience, the sense of an act is not itself intended or represented in the act. Rather, it belongs to the “content” of the act: it is part of the internal structure of the act whereby the act achieves its intentionality and so represents its object in a particular way. Senses, for Husserl, are thus not among the entities that we commonly experience in the natural world. (Husserl takes them to be “ideal” or abstract entities.) In particular, senses are not properties of the objects we intend: when we intend a physical object, we have a sense of it as an individual having certain properties; but the sense belongs to the content of the experience, while the properties belong (or are represented as belonging) to the object.

An act is intentional by virtue of having a sense or content, even if there is no object that “satisfies” this sense. Thus, on Husserl’s account as well as Dretske’s, misrepresentations and referential failures are possible.

So, there are some surface similarities but also some deep differences between Husserl’s account of intentionality and Dretske’s account of representation. For Dretske, a representation represents its sense and does not (in any literal sense of the term) represent its reference: the reference is only contextually, not representationally, related to the representation. For Husserl, a representation represents its reference and does not represent its sense: the sense belongs to the structure of the representation, not to the reference that is represented by its means. For Dretske, senses are properties of the sort that physical objects have. For Husserl, they are abstract “contents” of intentional thoughts and experiences.

Most important, mental states and experiences on Husserl’s account are *intrinsically* intentional: an act is intentional because it has a sense, and its sense is intrinsic to the act. Thus, the intentional or representational character of an act is due to its own internal makeup as the act that it is.<sup>2</sup> For Dretske, however, representational systems (or states) are *not* intrinsically representational. Even if a system’s *indicating* a certain property is due to its internal character, its *representing* something is not. A system represents something only if it has the *function* of indicating it, and systems get their functions from the “outside”: by being designed, in the case of conventional representations; by their evolutionary history, in the case of natural representations. Dretske does not emphasize this aspect of his theory until his final chapter, but there he is most emphatic. Says Dretske:

The Representational Thesis is an externalist theory of the mind. It identifies mental facts with representational facts, and though representations are in the head, the facts that make them representations – and, therefore, the facts that

make them mental – are outside the head. A state of the brain . . . represents the world in a certain way . . . only if it has an appropriate information-carrying function. Since functions . . . have to do with the history of the states and systems having these functions, mental facts do not supervene on what is in the head. What is in heads A and B could be physically indistinguishable and yet, because these pieces of gray matter have had relevantly different histories, one is a representational system, the other is not; one is the seat of thought and experience, the other is not. (pp. 124-25)

### 3. “EXTERNALIZING” QUALIA

As “subjective” features of experience, qualia have been thought to pose especially serious problems for any objective, naturalistic account of experience. Behaviorism, functionalism, and computationalism, for example, have all had little success in dealing with the fact (or apparent fact) that qualitatively distinct mental states can be behavioristically, functionally, and computationally equivalent. Accordingly, Dretske sees the problem of naturalizing qualia as a significant test of his representational version of naturalism, and he hails his own solution to that problem as a major triumph for his theory.

The Representational Thesis is plausible enough for the propositional attitudes [but] less plausible – some would say completely implausible – for sensory affairs, for the phenomenal or qualitative aspects of our mental life. Nonetheless, . . . I concentrate on . . . qualia – that dimension of our conscious life that helps to define what-it-is-like-to-be-us. I focus here because, frankly, this is where progress is most difficult. This, then, is where progress – if there is any – will be most significant. (pp. xiv-xv)

Given that Dretske’s Representational Thesis is the key feature in an *externalist* project for naturalizing the mind, his task of naturalizing qualia amounts to giving an “external” account of them, and this is a task that does seem more than a little implausible, as he puts it. As features that distinguish the “subjective feel” of one experience from another, qualia seem to be “internal” characters of experience par excellence. Dretske himself places them among the “qualitative aspects of our mental life,” and elsewhere he calls them “qualities of experience” that make up the “subjective life of [a] being” (p. 65). But Dretske’s naturalistic account of qualia, as we shall

see, locates qualia *outside* the experiencing organism and its experiences, identifying them with properties of the *objects* that the organism experiences.

We get a clue to Dretske's strategy in a casual remark about intentionality: "Intentionality is real enough," he says, "but it turns out, as Fodor . . . suggests it must, to be really something else" (p. 28). In that Fodorian spirit, Dretske succeeds in naturalizing qualia by construing them as "really something else": not "internal," inherently subjective properties of *experiences*, but objective properties of the *objects experienced*. "In accordance with the Representational Thesis," he says, "I . . . identify qualia with phenomenal properties – those properties that (according to the thesis) an object is sensuously represented . . . as having" (p. 73). Given that identification, says Dretske, there is no problem, in principle, with my knowing the quale of your experience, or a bat's or a parasite's, even though I am incapable of having experiences of those sorts. He offers the following example.

Consider a parasite that attaches to its host if and only if the host's surface temperature is almost precisely 18°C. The parasite then attaches to the host when the parasite has an experience that represents the host as being at 18°C; that is, when, to the parasite, the host seems to be 18°C. Now, says Dretske, the *quale* of the parasite's experience of the host is "how the host seems to the parasite," and "how the host seems to the parasite" is "being 18°C." And so the quale of the parasite's experience is the property of being 18°C. That property, moreover, is not something "inside" the parasite or the parasite's experience. Being 18°C is an objective property – a property of the *host*, if the parasite is representing it accurately. As such, it is something that you and I can know even though we cannot examine the parasite's experience. If we know what 18°C is, then we know the quale of the parasite's experience. Says Dretske:

If you know what it is to be 18°C, you know how the host "feels" to the parasite. You know what the parasite's experience is like as it "senses" the host . . . . All you have to know is what temperature is. If you know enough to know what it is to be at a temperature of 18°C, you know all there is to know about the quality of the parasite's experience. To know what it is like for this parasite, one looks, not in the parasite, but at what the parasite is "looking" at – the host. (p. 83)

Now, Dretske's view here strikes me as strange. For one thing, I am not sure that I do know "what it is to be (or be at) 18°C." Since Dretske's goal is to naturalize qualia, I presume he intends us to understand such properties as "being 18°C" naturalistically and objectively. As an educated person, I have some inkling of what objective temperature is: to be at a temperature of 18°C is, I think, to have a certain mean molecular kinetic energy, but I cannot be more precise

than that. And if I were not an educated person, I would not know even this much about temperature – or color, sound, shape, and so on – considered as an objective, naturalistically characterizable quality of things.

More to the point, though, knowing which physical properties constitute the temperature of an object (even of an object represented or experienced as having that temperature) seems simply irrelevant to knowing the qualia that temperature-experiences have. Even if I do not know what temperature is (in the naturalistic, scientific sense), I know how various temperatures typically feel. These various modes of what it is like to feel temperature are what phenomenologists, and I believe most other philosophers as well, mean by temperature “qualia.” But qualia so understood do not align one-to-one with the properties objects are experienced as having. The same water can feel cool to one hand and warm to another, as Berkeley famously noted. Yet, science tells us, the natural property that the water itself has – the temperature being experienced in these qualitatively different ways – is the same in both cases. But if such qualitatively different experiences are representations of the same objective property, then qualia cannot be identified with “those properties that . . . an object is sensuously represented . . . as having,” and Dretske’s representationalism founders on the same rock as behaviorism, functionalism, and computationalism.

Dretske seems to think that he can avoid this problem by collapsing the distinction between the objective properties of experienced things and the subjective ways things are experienced. “The Representational Thesis identifies the qualities of experience – qualia – with the properties objects are [represented] as having,” he says. “Subjectivity becomes part of the objective order” (p. 65). Dretske’s remarks, however, seem sometimes to affirm this identification and sometimes to deny it. “If you know what it is to be 18°C, you know . . . what the parasite’s experience is like as it ‘senses’ the host,” he says on page 83; but on the very same page he admits that “knowing what temperature is will *not* [my emphasis] tell one what it is like . . . to *feel* a temperature of this kind.” Again on the same page, Dretske apparently *distinguishes* knowing what a phenomenal property is from knowing what it is like to experience that property: “Deaf people can know what sound waves are without knowing what it is like to *hear* sound waves.” Later, however, he seems to make the very opposite claim about vision: “A blind person may know what it is like to visually experience [= see?] movement. If he knows what movement is, that is enough” (p. 94). Evidently, there are some tensions, if not outright contradictions, in how Dretske understands the relation between *qualities of experiences (of objects)* and (*experienced qualities of objects*) and the relation of both of these notions to that of *qualia*.

Dretske’s blurring of these notions is especially prominent in an argument he offers for identifying qualia with represented properties. The argument is straightforward, and Dretske challenges dissenters to find some flaw in it. He says: “I am merely drawing out the conse-

quences of facts that almost everyone accepts – facts that are quite independent of the representational point of view being defended in these lectures. . . . If [the] result is absurd, then one of the . . . facts that led to it – not the Representational Thesis – is to blame” (pp. 83-84). We saw Dretske’s argument at work in the “parasite” example above. Its main premise, although Dretske calls it a “fact,” is actually presented as a definition: “The first fact,” he says, “is that qualia are supposed to be the way things seem or appear in the sense modality in question” (p. 83). The argument Dretske builds from this “fact” can be reconstructed as a series of identities:

1. Qualia are the way things seem or appear.
2. The way a thing *seems* or appears in an experience is the way the thing would *be* (that is, the property it would actually *have*) if the experience were veridical.
3. The property a thing would have if the experience of it were veridical is the property that the thing is represented as having.
4. Therefore, qualia are properties that things are represented as having.

The problem with this argument, I believe, lies in that “first fact,” Dretske’s definition or characterization of qualia as “the way things seem or appear.” By “ways of appearing,” Dretske clearly means properties that *things* or *objects* appear to have, or are represented as having, in an experience. So understood, however, the first premise of the argument, “qualia are the way things seem or appear,” just *means* “qualia are properties that things are represented as having.” Taken in this sense, then, Dretske’s first premise is not a definition of “qualia” but a stylistic variant of his conclusion. The argument begs the question.

What makes the argument beguiling is that almost anyone, including those who would resist Dretske’s identification of qualia with represented properties, could accept his characterization of qualia as “the way things seem or appear.” There must, then, be a different way of understanding the phrase “the way things seem or appear.” Phrases such as this are in fact ambiguous. Compare, for example: (1) the way things seem or appear to *S* (= the way things are experienced by *S*) and (2) the way women were depicted by Degas. In what “way” or “ways” did Degas depict women? *How* did he depict them? The range of appropriate answers depends on just what we take the question to be about. If taken as a question about *women*, as Degas depicted them, one appropriate answer is “as dancers.” But if taken as a question about Degas’s *depictions* of women, a different type of answer is appropriate: “impressionistically,” for example. Thus, taken in one sense, “ways of depicting women” are descriptions that apply to women: “as dancers,” “as workers,” “as mother figures,” “as sex objects,” and so on. Taken in another sense, “ways of depicting women” are modes of depicting: “impressionistically,” “realistically,” “pointillistically,” “cubistically,” and so on.

Dretske's argument fails because of this ambiguity, I believe. His definition of qualia as "the way things seem or appear" is uncontroversial only if these "ways" are modes of *appearings* – qualities of *experiences* of objects – rather than properties of appearing *things*. But if the first premise must be given this reading rather than Dretske's, the argument fails. And if the first premise *cannot* be given this reading, then the argument seems not to be about *qualia* at all.

While Dretske's argument depends on blurring the distinction between qualities of experiences of objects and qualities of objects as experienced, much of his defense of his account of qualia depends on our recognizing that distinction. The highlight of Dretske's identification of qualia with represented properties is that it entails that anyone can know the quale of a supposedly subjective experience simply by knowing what objective property the experience represents. But here one wants to object that this cannot be correct: to know what property an experience is an experience "of" is not at all the same as knowing what it is like to experience that property. But, as we saw earlier, Dretske simply grants this: "Surely knowing what temperature is will not tell one what it is like (if it is like anything) for a parasite (or even another human being) to *feel* a temperature of this kind. . . . I do not wish to deny this. I am *not* denying it" (p. 83). To understand this response I believe we must make the very sort of distinction that Dretske's argument suppresses.

Let us return to Dretske's parasite, which is sensing its host as being 18°C. Dretske has said

1. "If you know what it is to be 18°C, you know how the host 'feels' to the parasite. . . . You know all there is to know about the quality of the parasite's experience."
2. "Knowing what temperature is will not tell one what it is like . . . to *feel* a temperature of this kind."

In affirming (2), I think Dretske must also mean to affirm

3. Knowing what 18°C is will not tell one what it is like to *feel* a temperature of 18°C.

So Dretske is making our same subtle distinction here. Knowing "how *the host* feels to the parasite when the parasite senses the host as being 18°C" is not the same as knowing "how it feels to *the parasite* to sense the host as being 18°C." We can understand this distinction and also see why Dretske identifies the first, but not the second, with the property of being 18°C. To ask how *the host* feels to the parasite is to ask which property the parasite senses *the host* as having, and in the present case that is the property of being 18°C. To ask how it feels to the parasite to sense the host as being 18°C is to ask, not how *the host* seems, but how the parasite's *sensing* of the host feels to the parasite. Thus, it makes sense to say that knowing what 18°C is suffices for our knowing how the host feels to the parasite but not for our knowing what the parasite's experience of the host feels like to the parasite.

However, if we allow Dretske to distinguish in this way between "how the host feels to the parasite (when the host feels 18°C)" and "what it is like to feel a temperature of 18°C," then he

loses the bigger game. “How the host feels to the parasite,” on this reading, is objectively characterizable, characterizable in terms of the temperature the host is represented as having. It is, after all, a quality of the *host* (or at any rate, a quality of the host if the parasite’s experience is veridical). But that means that this property is trivially “natural” – natural in the very same way that all intersubjectively observable properties are natural.

Dretske’s naturalization of qualia would be noteworthy if the “qualities of experience” that he naturalizes were indeed phenomenological features of experience. But the quale of the parasite’s experience, in this phenomenological sense, is what having that experience is like for the parasite, and that remains unexamined in Dretske’s externalist version of naturalism. One does not naturalize the mind by naturalizing the *objects* that minds represent.

---

\* This paper is derived from an earlier version called “Naturalizing Phenomenology: The Very Idea,” presented at the international conference of philosophy “Actualité cognitive de la phénoménologie: Les Défis de la naturalisation,” Bordeaux, France, October 19-21, 1995.

1. Fred Dretske, *Naturalizing the Mind* (MIT Press, Cambridge, 1995). Page numbers in the text will refer to this work.

2. I oversimplify Husserl’s version of “internalism” here. The intentionality of a particular act is also affected by its relation to background beliefs and other items internal to the experiencing subject. See David W. Smith and Ronald McIntyre, *Husserl and Intentionality: A Study of Mind, Meaning, and Language* (D. Reidel, Dordrecht, 1982), pp. 246-56.