

Naturalizing the Mind, by Fred Dretske. Cambridge, MA: MIT Press, 1995. Pp. 208.

In this important book Fred Dretske defends a version of externalism which he calls representational naturalism. The goal of representational naturalism is to explain much of what is puzzling about sensory experience and the contents of sensory experience, qualia. Many of the arguments for representational naturalism come from promoting what Dretske calls the Representational Thesis, particularly as it applies to sensory experience. The Representational Thesis states that all mental facts are representational facts and all representational facts are facts about informational functions (p. xiii). The first four chapters of the book are the 1994 Jean Nicod Lectures. In these chapters, Dretske sets out the relations and differences between cognitive and experiential representations; he argues that introspection is a species of displaced perception; he explains qualia from a representational, naturalistic perspective; and he analyzes consciousness within this perspective. The fifth and final chapter defends externalism of sensory experience from internalist objections.

Dretske begins by laying out the relationship between representation and function: X represents property F if and only if X has the function of indicating F (p. 2). The senses have biological functions of providing information. These indicator functions give rise to natural (as opposed to conventional) representations in the organism. There are two flavors of indicator functions and natural representations: *systemic* and *acquired*. *Systemic* indicator functions are inherited, or phylogenetic, and give rise to *systemic* representations, denoted by the subscript "s" as representations_s. *Acquired* indicator functions are ontogenetic, acquired through learning, and give rise to *acquired* representations, denoted by the subscript "a" as representations_a. Dretske points out there are also two sorts of mental representations: experience, which is nonconceptual representation, and thought (beliefs, judgements, knowledge, etc.) which is conceptual

representation. Experience and thought are paired with systemic and acquired representations:

...experiences are to be identified with states whose representational properties are *systemic*. Thoughts (conceptual states in general), on the other hand, are states whose representational properties are *acquired*.(p. 15)

Imagine a child seeing a tree for the first time. The child would have the same representations_s as an adult looking at the tree from the same angle -- her *experience* would be the same. The adult, however, would have the belief, or representation_a that the object was a tree. We could say the same for a color, such as red. People are hardwired, phylogenetically, in such a way to represent_s sensual properties, to have experiences; and we can learn, ontogenetically, to represent_a properties and so come to have thoughts. It is this phylogenetic vs. ontogenetic difference which distinguishes between experience and thought.

In this view of representation, learning is like the calibration of an instrument. Consider a speedometer dial with no markings. It indicates the speed of the car, because the dial position varies nomically with car speed. This can be thought of as a representations_s of car speed. Now add markings to the dial. The speedometer has now been calibrated; the representations_s has been calibrated to be a representation_a -- a representation that, like thought, can guide behavior executively. Empty dials may carry very precise information, but not in a form that makes it useful for a system. Dretske uses the example of color calibration. Most people can experience hundreds of different colors, but have only a few dozen categories conceptually. In using this notion of calibration Dretske joins company with Gareth Evans in saying that *only* representations_s that can be made into representations_a can qualify as experiences. Other representations_s cannot be experiences. Dretske is surely right about this point. For example, our visual

system may have zero-crossing detectors in Marr's (*Vision*, Freeman, 1982) sense. We thus have, phylogenetically, representations_s of zero-crossings in every visual experience. Yet these representations_s are inaccessible -- they cannot be calibrated into beliefs, or representations_a. They are too low a level, cannot be accessed cognitively, and so cannot be learned. *I* certainly do not experience zero crossings--there is no quality of how they look phenomenally to me.

According to Dretske, representational systems represent properties, not the objects exemplifying those properties. For a given representation, what determines the object exemplifying the property is an external relation he calls C, to suggest context. He says, "There is nothing in the content of the representation, nothing the representation *says*, which makes it about this object rather than that object or no object at all."(p. 24,25) A state S has the function of indicating some property F of objects (under conditions C). This is a representational fact. But it is not a representational fact (though it may be a fact) that the object in question stands in relation C to the state S. Dretske tells us that this is why introspection cannot tell me the object, if there indeed is one, that I experience. This seems right. I don't experience the relations I have to the objects I am experiencing or representing. I experience the properties. Introspection brings us back to the properties we are experiencing, not further -- not to the object itself that exemplifies those properties.

Introspective knowledge, on this view, is metarepresentational: a representation of a representation *as* a representation. It is a form of displaced perception. A surgeon 'sees' organs inside a patient using another (displaced) object, an ultrasound device, as a representational vehicle. Introspection is similar; we gain knowledge of our experiences through perceiving a displaced object. This object is not the experience itself, but the object of the experience.(p. 44) Introspection also requires a connecting belief that experiences of F are generally reliable indicators of F. Given an experience of red, and the connecting belief that such experiences are reliable, we can gain introspective

knowledge that we are experiencing red. Introspective knowledge is not a systemic representation, or representations_s, but a cognitive or conceptual representation -- a representation_a. Given my experience of red, I don't directly introspect *red*, this is not the content of my introspective representation, rather, I have a representation *that* this is an experience of red. Dretske tells us that there is no phenomenology to introspection beyond the phenomenology of the experience one is introspecting. He uses a gustatory example to make the point: when asked to describe the experience of tasting a particular wine, the sampler has nothing to attend to beyond the wine itself. Because there is no extra phenomenology beyond what is being experienced (the wine), this counts as evidence against having an internal "scanner" with its own phenomenology. The conclusion is that there is no internal scanner, or extra sense, for introspection.

All this might sound unsatisfying to the what-it-is-like-to-have-an-experience-of-X crowd. I can imagine them giving this sort of reply, on Dretske's own terms: If introspective representations are cognitive representations, or representations_a, then, according to Dretske's theory they are learned. As representations_a, they have been serviced by representations_s. But representations_s that can support representations_a are systemic, natural indicator functions with their own phenomenology. Other representations_s that support representations_a have a phenomenology, so shouldn't those representations_s that support *introspective* representations_a also have a phenomenology? Perhaps these representations even serve as internal scanners with their own phenomenologies. I can introspect my tasting of the wine. My representing this experience through introspection is different from my simply tasting the wine. A different phenomenology, or feel, can accompany this metarepresentational state, because the metarepresentation issues from a different representation_s than the experience.

Dretske's counter to this objection would, I think, be twofold. First, recall that introspection is a species of displaced perception. I introspectively perceive my experience through the object I am experiencing. The content of my introspective

representation is the experience of F, but the vehicle or instrument through which I gain this content is the object of the experience, viz., the object which is F. This object plays a role analogous to an altimeter which informs the pilot in an aircraft, or an ultrasound device which informs the surgeon. The pilot or surgeon perceives an external object is F (altitude or location), by looking at another object (gauge or screen). Since the property associated with our introspective instrument is the property F, the very property that we are *experiencing*, there is no other phenomenology to be found to introspection beyond that of the experience, and consequently no internal scanner either. Second, Dretske tells us that “Representational systems have the function of supplying information about the world, not about themselves.”(57) The same goes for introspection. Introspection tells the system about how external objects are being represented by the system -- as F’s for example. It doesn’t inform about the *internal* objects or properties of the system doing the representing.

A defender of internal phenomenology could counter the second point by saying that the natural function of introspective representations is to indicate other mental representations. It is not to indicate external objects and their properties. That is the job of the senses. We don’t need two sets of representational abilities, senses and introspection, to tell us about properties of the external world. Further, states with different representational contents have different functions, since they indicate different things. The introspective content 'experience of F' is different from the experiential content 'F'. Thus introspection and experience have different functions. Why shouldn't the two types of states have different phenomenal characters, since their contents and functions are different? Simply denying any phenomenology to introspection beyond that given by the content of the experience introspected, as in the first counter available to Dretske, just ignores the relation between function and phenomenology. From what we have seen so far, natural indicators with different functions have different phenomenologies. This suggests that it is the function of a representation that determines

its phenomenology. Perhaps only sensory systems (the five senses plus proprioception) have a phenomenology, and no other internal representations, like introspection, do. But if this is the case Dretske must provide further explanation for holding that introspection is an exception in not having a phenomenology of its own.

Next for Dretske is an analysis of qualia. He identifies qualia with those properties that experience represents_s objects as having. The color blue is the property my experience represents_s something as being; it is the content or quale of my experience of blue. The Representational Thesis is a materialist thesis, so these properties are knowable objectively and the subject of a particular experience does not have privileged access to them. Qualia are objectively determinable because they are properties that the senses have natural functions to provide information about. These functions are as objectively determinable as the biological functions of bodily organs.

Consider Mary, who thoroughly understands dogfish and electric-fields (E-fields). Because she can represent any E-field precisely, she knows what it is like to be a dogfish experiencing an E-field of type T. "If Mary knows what a field of type T is, she knows all there is to know about the quality of experiences of this type."(p.85) Dretske says that the dogfish as *experiencer* of an E-field of type T does stand in a different causal relationship to the field because it is aware that *this* is an E-field of type T. But he argues Mary still knows what the dogfish experiences. He uses this analogy: I experience blue, you do not (you're in the other room), and I tell you I see blue. "There is nothing you can refer to as *this* that is the color quality I happen to be experiencing. This does not show you do not know what the quality of my experience is." The same goes for Mary and the dogfish. Dretske concedes that we cannot know everything about what it is like to be a dogfish but only because we cannot be immersed in the complete mental life of such a creature. But its qualia are in principle knowable.

This brings Dretske to problems of consciousness. He considers objects of awareness, and points out that the properties we experience when we dream or

hallucinate are precisely the properties we would be aware of were we sensing those properties veridically. What we lack in these situations is the contextual relation C which connects us with the objects of veridical representations. The same type of internal state, with the function of representing certain properties, come into play in both cases -- all that is lacking in hallucination is the external object with those properties. This is a case of misrepresentation to be sure, but the quality of the experiences are the same by virtue of the states having the same functionality.

Higher Order Theories (HOT) of consciousness are rejected by Dretske, because they do not recognize lower order states such as experiences as conscious states. These theories hold that consciousness only occurs in those organisms that have a higher 'internal scanning' level in which they are conscious *that* they are having a particular experience (or thought). Dretske gives two examples to show that consciousness also occurs at the 'lower' level of the sensory experience or cognitive state. First, children only acquire metarepresentational abilities (come to know of themselves *as* representers) between the ages of 3 and 4. According to HOT then, they can't be conscious at younger ages. But surely 2 year old children are conscious. Second, there are actual sensory differences when I view you momentarily holding up 7 cards, then 8 cards, in your hand - - even when I *believe* in *both* cases that you have held up 7 cards. This is a real conscious difference at the sensory level which is inaccessible to a HOT theory. Consciousness, therefore does *not* require a scanner. Experience is enough.

In the final chapter, Dretske makes his case for an externalist theory of experience by opposing what he calls the Internalist Intuition. This intuition says that even if externalism is true for cognitive states such as belief, it can't be true for experiences. He argues that awareness of experiential properties is an indirect process that requires a cognitive capacity, namely, the possession and use of concepts. For example, I need a concept (the concept RED) to become aware that I am having the experience that something looks phenomenally red. I can become aware of the color red without this

concept, but I need the concept to become aware *that* I am having an experience of a red quale. I cannot introspect a quality unless I have the concept of the quality. Qualia are therefore inaccessible until the subject has the requisite conceptual resources for being aware of them. Since awareness of experiential states requires cognitive capacities, these states are subject to the same Putnamian twin arguments that other cognitive states (such as beliefs) are. The result is that experiential contents -- qualia-- aren't in the head, and for the very same (externalist) reasons that meanings aren't in the head.

Dretske reminds us that according to the Representational Thesis mental facts -- contents and qualia included -- are facts about representational functions. And functions depend on the *histories* of the states with the function. So, mental facts do not supervene on what is in the head. Rather, contents and qualia supervene on the causal history that installed the agent's cognitive and experiential states. The histories of cognitive states are ontogenetic: the learning histories of the organism. The histories of experiential states are phylogenetic: evolution gives us our systemic representational states. It is to these histories that we look, then, to determine content and qualia within organisms.

Conscious experience is a product of natural selection. Consciousness itself isn't selected for, but systemic indicators are, and since these indicators -- experiential state-types -- have been selected for, they have a natural function to deliver the information they do.

Let me conclude by pointing out that many of the notions used in this book were developed at more length in Dretske's previous work, *Explaining Behavior* (MIT, 1988). Readers not acquainted with Dretske's treatment of indication, representation, and function, and the distinction between behavior and bodily movement might benefit from reading the earlier work. But it is not necessary. *Naturalizing the Mind* is written clearly and in a very readable style. There are many examples which are particularly germane for understanding the difficult topics of experience, qualia, and consciousness. The book should be considered a must-read for internalists and externalists alike.