

Privileged Access to Information: Dretske's Accounts of Self-Knowledge

Kai-Yuan Cheng, Chia-Yi, Taiwan

I

Fred Dretske (1981, 1988) has been noted for his vigorous project of naturalizing the mind. In his version of representational theory, a mental state is characterized as a physical state of a person, which has a certain function of carrying information about aspects of the environment, to which the person is causally related. So construed, a mental state is an objective item in the world, which shall be accessible to both its possessor and external observers in a similar way. This, however, does not seem right. It does not seem to do justice to our observation that we appear to have a privileged (not easily challenged by others) and immediate (not mediated by inference or observation) access to our mental states, an access which no other people share. How could this special first-person epistemology and phenomenology be explained within a naturalistic framework? Dretske (1995, 1999, 2000) has taken up the challenge to solve this daunting problem, by offering some ingenious and illuminating accounts of introspection. This paper aims to discuss how, and whether, those accounts work, and to the extent that they do not, whether they can be amended.

Although Dretske does not make it explicit, there are two main accounts that Dretske has offered to explain the special characters of self-knowledge (c.f. Lycan, 2003). In one account, Dretske (1995) exploits the notion of displaced perception to provide a model of introspection. In another account, Dretske (1999, 2000) first distinguishes between three different kinds of awareness, and then shows how making this distinction can offer a "neat and satisfying" explanation of the special characters of self-knowledge. The core idea of these two accounts is the same, but the ways in which the idea is implemented are different. I will make three main claims in this paper: 1) the first account has a difficulty explaining the psychological immediacy of self-knowledge; 2) the second account can avoid this difficulty, but it would encounter a problem of its own; 3) An alternative model is proposed to account for the directness of self-knowledge that shall be congenial to the two accounts offered by Dretske. The following three sections are devoted to illustrate these three claims respectively.

II.

Dretske (1995, Ch.2) offers a displaced perception model of introspection to explain the special characters of self-knowledge. This model is built on the basis of the conceptual framework in which a cognitive subject is assimilated to a simple measuring instrument like a scale as a representational system. A scale can represent my weight when I stand on it. If it works properly, it delivers a correct piece of information about my weight. A scale can also *misrepresent* my weight if it breaks. In that case, it does not carry correct information about me. Whether a scale misrepresents or not, there is always an objective representational fact about the scale, a fact that is automatically and necessarily there whenever the scale stands in a causal relation to something. I, as a cognitive subject, can also represent certain aspects of the world to which I am causally

related. By standing in front of a red flower, I can represent what it looks like by having a certain visual experience of it. I may misrepresent the flower as yellow, when something goes wrong. In this case, there remains a representational fact about me, about how I (mis)represent an external object. Other people cannot occupy my position to have my (mis)representational state; it belongs only to me. They can nonetheless obtain the informational content of my representational state, but only in a more indirect way.

Now, could a scale know, or have access to, the informational content of its representational state when I stand on it? Obviously not! The reason is that it could not conceptually represent what a scale and an object like me are, and what the relation between the two may be. As an external observer, in contrast, I can know the informational content of the scale's representational state, because I know that the scale would not have a pointer position like that unless I had a certain weight. This is an instance of *displaced perception*, which has the following form:

I know that k is F by perceiving that h is G, plus a connecting belief about k and h, namely, h would not be G unless k were (probably) F. (c.f. Dretske, 1995: p.42)

Dretske then applies the notion of displaced perception to explain how we know our own mental states. To the question: "How do I know what my experience is like when I see a red flower?", Dretske's answer is that I do not "look inward" to find out facts about my experience. All I have to do is to look at the red flower in front of me, see what its color is like, and infer that my experience is of a certain kind, based on the connecting belief that if my experience were not of this kind, the flower would not appear in a certain way to me. Introspective knowledge is thus a form of displaced knowledge, which is obtained by perception of external objects, plus certain connecting beliefs.

Dretske holds that the nature of perceptual displacement in introspection is the "source of the first-person authority" (1995: p. 53). A person does not need to look at herself, in order to know some state of her own, e.g., to know what kind of experience she has. She only has to look at a red flower, and think to herself that her experience must be an experience of a red flower, for otherwise the flower would not look red to her. She may misrepresent the flower as yellow. But in that case, she would again know that her experience is of a yellow one, for otherwise the flower would not look yellow to her. There is thus a strong sense in which a person's introspective knowledge is *infallible*. On the other hand, the explanation of the psychological immediacy of self-knowledge is not so straightforward on the displaced perception model. On this model, the obtainment of introspective knowledge is mediated by a connecting belief. This obviously renders introspective knowledge inferential. Dretske has noticed this defect in his model, and tries to neutralize its negative impact by pointing out that the form of inference involved in introspective knowledge on the displaced perception model is very unusual. It differs from other species of displaced knowledge in two aspects. One is that introspective knowledge does not require that one's representation of the world be veridical, while other kinds of displaced knowledge do. Another is that one's connecting belief does not

have to be true in order for introspective knowledge to be obtained, whereas the same thing does not apply to other kinds of displaced knowledge. Dretske claims that this is the "source of the 'directness' and 'immediacy' of introspective knowledge" (1995: p. 62)

In my view, the explanation of the epistemic privilege of self-knowledge on the displaced perception model has more plausibility than that of psychological immediacy. Despite the fact that introspective knowledge is achieved *via* a very peculiar form of inference, it *remains* inferential in nature. Our phenomenology of introspection is not like that, however. We know what we think, and what our experience is like, in a saliently immediate way. We do not normally go through an inference that *h* looks *G* to me, so my experience *k* must be of the *F* kind, in order to obtain introspective knowledge. There must be something wrong on the displaced perception model.

III.

Dretske (1999; page numbers refer to 2000) has further, and independently, offered an account to solve a vexing problem concerning phenomenal experience. This problem has to do with the follow two gripping statements:

- (1) Conscious perceptual experiences exist inside a person (probably somewhere in the brain"
- (2) Nothing existing inside a person has (or needs to have) the properties one is aware of in having these experiences. (Dretske, 2000: p. 158)

It seems right that my experience of seeing a red flower is inside me. It also seems right that nothing inside me has the (redness) property that I am aware of when I have this experience. The puzzle then arises: "How, then, can I be aware of what my perceptual experiences are like...if none of the properties I am aware of when I have these experiences are properties of the experience?" (Dretske, 2000: p. 159) Dretske tries to solve this puzzle by drawing our attention to the distinction between the following three kinds of awareness: object-awareness (o-awareness), property-awareness (p-awareness), and fact-awareness (f-awareness). Dretske puts great effort to show that one could have one kind of awareness without having another two. For the sake of the main concern in this paper, we may put aside how Dretske demonstrates this and whether the demonstrations are convincing. The point is to note that the result of showing the distinctness of these three kinds of awareness has important implications.

Suppose *e* is my perceptual experience of a red flower, and *r* is a property of *e*. Dretske argues that I become aware of what my experience of a red flower is like, not by having an awareness of either *e* or *r*, but by having an awareness of the fact that I have an experience of a red flower. In other words, my awareness of an experience *e* is an f-awareness that *e* is *r*, rather than an o-awareness of *e*, or a p-awareness of *r*. The puzzle can then be solved. Both (1) and (2) can be true, because the experience *e* is an internal state of mine, and it does not have the property of redness *r*, since it is simply a brain state. All I should, and can, do, in order to become aware that *e* is *r*, is to be o-aware and p-aware of a red flower which is externally related to me. The mistakes of the Lockean "inner sense" theories of introspection lies, according to Dretske, in failing to make distinct the three kinds of awareness, and in holding that one's f-awareness that *e* is *r* is directly achieved by having an o-awareness of *e* and p-awareness of *r*.

Dretske claims that this account of the mind's awareness of itself can give a neat and satisfying account of the special characters of psychological immediacy and epistemic privilege of self-knowledge. Dretske admits that one's awareness of one's own experience is indirect, since it consists in having a f-awareness that *e* is *r*, rather than in having a direct o-awareness of *e* or p-awareness of *r*. Nonetheless, the externally represented property of an object directly "reveals to the person having the experience exactly what property [i.e., *r*] it is that his or her experience has" (Dretske, 2000: p. 170). Dretske is making the point that one's f-awareness that *e* is *r* is so directly given by a p-awareness of the property of an external object that one's introspective knowledge seems direct and immediate. Dretske adds that one's introspective knowledge can be made more direct and immediate by confusing the property of an external object one is p-aware of and the property of one's experience which one is not p-aware of. As to explaining the character of epistemic privilege, Dretske's story does not differ much from what he gives in the previous model, that a person having an experience is the only person who can possess this experience. No one else can occupy that privileged position.

This account of self-knowledge is similar to a displaced perception model of introspection discussed earlier in one basic aspect: both accounts regard introspective knowledge as a product of an indirect process, from perceiving things outside to being aware of something inside. However, there lies a crucial difference between these two accounts. In the second account, no connecting belief is mentioned or needed for the attainment of introspective knowledge. Making this move is exactly what Dretske needs to do, because, as we have seen earlier, only by doing so can the non-inferential phenomenology of introspective knowledge be faithfully captured. However, a problem remains. How does this account explain that a person shifts from a p-awareness of properties of an external object to an f-awareness about her inner mental state? A young child can have a same experience as an adult when they both see a red flower; they are both p-aware of the redness property of the flower. But the young child is not f-aware that *e* is *r*, whereas the adult are normally f-aware that *e* is *r*. Why does this difference exist? A reasonable explanation is that an adult has suitable conceptual representations of what a red flower and an experience are, and of what the relation between the two may be. Seeing a red flower, plus having a connecting belief that I would not see a red flower as red unless I had an f-awareness that *e* is *r*, are responsible for an adult to possess an f-awareness that *e* is *r*, and are lacking in a young child. This explanation is clearly a displaced-perception model explanation. Dretske's theory of introspective-knowledge thus seems to face a dilemma. On one horn of the dilemma, to better explain the psychological immediacy of introspective knowledge, Dretske has to give up the idea of a connecting belief in a displaced-perception model when he offers the account of introspective knowledge as a form of f-awareness directly revealed by p-awareness of external objects. The problem is that this account cannot explain how one moves from p-awareness of external objects to f-awareness. What would be worse is that when attempting to look for an explanation, this account turns to the displaced-perception model which it has rejected. On another horn of the dilemma, Dretske retains a displaced-perception model, but this would leave the problem of explaining the psychological immediacy of introspective knowledge unresolved.

IV.

Dretske's theory of self-knowledge thus seems unstable when it comes to deal with the character of psychological immediacy of self-knowledge. He offers two accounts of introspective knowledge, but one is unsuitable, and another appears inadequate. An internal tension between the two accounts seems to resist any easy fusion of them. My proposal is that Dretske's two accounts of introspective knowledge can be made compatible and unified within an expressivist view of mental states and self-knowledge, such as the one offered by Bar-On (2004). On this position, a mental state is of such nature that it can be expressed by its possessor either in a linguistic or non-linguistic form. When a normal young child sees a box of ice cream, she may reach for it. This physical movement directly expresses her desire for ice cream. As the child gradually grows up, she begins to pick up certain words, and may utter "Ice Cream!" to express her desire for ice cream. As a person's linguistic capacities and social interactions get more sophisticated, she may say "I want ice cream" to make her desire explicit. The third type of expression is a mental self-ascription in a sentence form. It can be either true or false, depending on whether the person who utters it has the self-ascribed mental state.

Now, to characterize the same phenomena with Dretske's terminology, a young child who has not developed any linguistic ability will be said to be merely o-aware and p-aware of a box of ice cream. She has a desire for ice cream, but is not f-aware of it. She only becomes aware of it after she has been taught certain concepts like "desire" and "ice cream", and has learned how to relate the two concepts in a variety of contexts. A person's f-awareness of her desire d for something x is not possible without the person's having acquired elaborate linguistic capacities to form certain connecting beliefs that I would not want x unless I had d . The process of developing relevant concepts and establishing connected beliefs is a necessary condition for a person to possess introspective knowledge, in the form of having an f-awareness that I have a mental state as such and so. However, after a person matures and has mastered relevant practices, the person does not need to entertain a conscious connecting belief in order for her to obtain introspective knowledge. A person, upon seeing a red flower, can simply and directly express her experience in a linguistically self-ascribed form like "I have an experience of seeing a red flower". My conclusion is, therefore, that Dretske's two accounts of introspective knowledge can more adequately and coherently explain the psychological immediacy character of self-knowledge, if some form of expressivism about the mind is supplemented.

Literature

Bar-On, Dorit 2004 *Speaking My Mind: Expressivism and Self-Knowledge*, Oxford University Press.

Dretske, Fred 1981 *Knowledge and the Flow of Information*, Cambridge, MA: MIT Press, Bradford Books.

Dretske, Fred 1988 *Explaining Behavior*, Cambridge: Cambridge University Press.

Dretske, Fred 1995 *Naturalizing the Mind*, MA: MIT Press, Bradford Books.

Dretske, Fred 1999 "The Mind's Awareness of Itself", *Philosophical Studies*, 95, 103-24.

Dretske, Fred 2000 *Knowledge, Perception, and Belief: Selected Essays*, Cambridge: Cambridge University Press.

Lycan, William 2003 "Dretske's Ways of Introspecting", in *Privileged Access: Philosophical accounts of Self-Knowledge*, edited by Brie Gertler, Ashgate.

Email: kcheng@phil.ccu.edu.tw