Accidental Associations, Local Potency, and a Dilemma for Dretske

PAUL NOORDHOF

Abstract: I argue that Fred Dretske's account of the causal relevance of content only works if another account works better, that put forward by Gabriel Segal and Elliot Sober. Dretske needs to appeal to it to deal with two problems he faces: one arising because he accepts that the mere association between indicators and indicated is causally relevant to the recruitment of indicators in causing behaviour, the other from the need to explain how a present token of a certain type of content is causally relevant. For this and other reasons their approach has clear advantages over Dretske's.

How can mental content be a cause of behaviour when it does not supervene on a subject's internal state? Fred Dretske has recently confirmed that it is this question that he sought to answer (Dretske, 1994, p. 206). His story is by now pretty familiar, but for those who have missed it, it runs as follows. Let N stand for neural events of a particular type, M for bodily movements of a particular type,[1] E for the presence of something in the environment, and let n, m, and e be particular instances of these types of thing. His basic idea is that n's possession of content is efficacious because it is the *indication* of E by Ns that *caused* the causal relationship between Ns and Ms to hold. For Dretske, a piece of behaviour is n's causing of m, not m. So, if indications of E by prior ns were the cause of the causal relationship being present, they would be the cause of behaviour. Dretske calls this the *structuring* cause of the behaviour to distinguish it from the *triggering* cause, the thing that set the causal process between n and m going. Ns indicate E if and only if there is a dependency between them such that $P(e/n) = 1$. Dretske takes indication

---

[1] It is worth remarking here that although I shall be restricting the discussion to bodily movements, Dretske takes behaviour to include causal processes which end in non-movements and, indeed, with external events like the spilling of a glass of wine.

in this sense to be a type of meaning, *natural meaning*. Prior ns' indication of E caused the causal relationship between Ns and Ms because Ns were recruited to be causes of these bodily movements *in virtue of such indication*. He thinks that this fact is enough to show that *content is causally relevant*, that *a state's possession of content is efficacious*. I use these formulations interchangeably. The recruitment mentioned is a piece of learning resulting from reinforcement. Good things happened when an N gave rise to a bodily movement in the past, so the connection between Ns and that type of bodily movement was reinforced. As a result of Ns' recruitment, Dretske suggests that Ns have the *natural function* of indicating E and hence are *representations* of E (Dretske, 1988, p. 56). They have the distinctive feature of representations, namely that they may be involved in misrepresentation. Ns that have the natural function of indicating E have the *non-natural meaning* or *content* that E is present.

My claim will be that this story only works as an account of how content can be causally relevant to behaviour if another account also works, an account that works better. This is unfortunate for Dretske as part of the motivation he offers for his own account is that it alone can explain how content can be causally relevant (Dretske, 1988, p. 36, and 1990, pp. 9–10). I begin by outlining an alternative account put forward by Gabriel Segal and Elliot Sober. I then go on to isolate two problems with Dretske's account. The first stems from the fact that he allows that a mere association between Ns and Es is efficacious. It is the fact that whenever Ns are present Es are that is a cause of the recruitment of Ns. The second problem arises when one considers whether Dretske can explain how a present token of a certain type of content is causally relevant. I claim that the answer to these two difficulties is to adopt the Segal and Sober account of efficacy. If this is so, their account has clear advantages over Dretske's in so far as it applies independently of his apparatus and allows that a state's possession of content is a triggering cause of a bodily movement. I close by considering another way of resolving the second problem Dretske faces, suggested by Jaegwon Kim. I argue that there is a reason to prefer the Segal and Sober approach.

1. The Alternative Account

One answer that Dretske presumably thinks doesn't work is that offered by Gabriel Segal and Elliot Sober (Segal and Sober, 1990). Their idea is that a state's possession of content may be efficacious if only part of the content's supervenience base is efficacious.[2] Although content does not supervene on

---

[2]   'May' rather than 'is' because they also require that content-bearing states figure in *non-strict laws*. Since there is no debate about whether content-bearing states can figure in non-strict laws, I shall suppress this qualification in the text. Also, I shall not go into details about the type of supervenience that they envisage. Suffice to say that they characterize it as mereological supervenience and claim that possession of properties that supervene in this way 'amounts to nothing more than' possession of the super-

internal properties alone—relational properties partly determine what content a subject has—it does supervene on internal properties plus relational properties. They make a comparison. The property of being air is efficacious in combustion because part of its supervenience base is oxygen. The presence of air was a causal factor behind the fire because oxygen was. The thought is that a state's possession of content is efficacious in the same way. It can be a cause of the bodily movement because it partly supervenes on neural properties that are causes of these movements. Obviously, it couldn't be a cause of behaviour understood in Dretske's sense because for him behaviour is a process *including* the mental state with content (Dretske, 1988, p. 38). Something cannot be a cause of a thing of which it is a part. But there is nothing mandatory about this notion of behaviour especially if it is not needed as *the* explanandum for which content-bearing states are the explanans. If one wished to retain the idea of behaviour being a process intimately linked with the causal activity of a mental state one could say that it was a process that starts with the immediate causal consequence of the mental state that is token identical with n, and ends with the bodily movement. The mental state's possession of content would be a triggering cause of this type of behaviour.

Dretske indicates his attitude to the proposal put forward by Segal and Sober in the following passage (Dretske, 1990, p. 9):

> And what has been shown about meaning is that if meaning supervenes, at least in part, on the *extrinsic* properties of an event—historical and relational facts that *need not* be mirrored in the event's current (= the time at which it has its effects) physical constitution or structure—then, if A causes B, the fact, if it is a fact, that A means M will not—indeed, cannot—figure in a causal explanation of B. It cannot because, in similar conditions, an event lacking this meaning, but otherwise the same, will have exactly the same effects. So it isn't A's having the meaning M that explains why B occurred. It is, rather, A's having C, a given set of intrinsic properties, that explains this.

In fact, Dretske doesn't believe that content supervenes even in part on the intrinsic properties of a subject (Dretske, 1990, pp. 6–7). This is surprising if one bears in mind that the association of Ns and Es must surely depend in part on internal properties.[3] However, the crucial point is that Dretske can-

---

venience base. It is not obvious that they capture this type of supervenience by thinking of it as a simultaneous *nomic* relation (see Segal and Sober, 1990, pp. 9–10).

[3] Something that he seems to tacitly acknowledge, see Dretske, 1988, *Explaining Behavior*, ch. 3, p. 56. A footprint doesn't indicate a quail rather than a pheasant, even if nomically dependent on the former, if the pheasant produces the same kind of footprint. This suggests partial supervenience of natural meaning on internal properties.

not afford to take this attitude to the Segal and Sober approach if I am right in thinking that his account faces the two problems outlined below.

2. Two Problems with Dretske's Account that Need One Answer

2.1 Accidental Associations and Learning

The first problem arises from Dretske's response to Carol Slater. She put it to him that the selective process behind learning would not distinguish between an accidental association between all past ns and es and a nomic dependence, given the reasonable assumption that this dependence is a modal one (Slater, 1994, pp. 168–9). That means that the selective process behind learning cannot distinguish between an accidental association and the indication relation. Dretske concedes this. He writes (1994, p. 204):

> It is quite true that if indication is a modal relation, then learning of the sort I will describe will not respect the difference between the effects of contingent and necessary co-occurrence of N with E. (my relettering)

He does not deny that indication is a modal relation. Indeed, he is committed to it being so in his earlier work (Dretske, 1988, pp. 56–7). Instead, he suggests that although the modal relationship is not necessary, it is a sufficient cause of the effects he has in mind. We can infer that it is N's indication of E rather than an accidental association between them which is causally responsible for the recruitment of Ns as an inference to the best explanation. Accidental associations of this sort are just unlikely.

It seems to me that this vitiates Dretske's approach. The problem is this. If we ask whether it is N's *mere* association with E in the way specified or N's indication of E that causally explains why Ns were recruited, the answer is that it is N's association with E. In similar circumstances, an association of N and E that lacks modal force but is otherwise the same will cause the recruitment of Ns. A more precise account of the cause of the causal relationship should mention only the association and not the stronger, indication, relation. So it turns out that the indication relation is not causally explanatory.[4] This conclusion is a consequence of the same type of reasoning as Dretske used to discredit the causal relevance of content in the passage quoted above. If he thinks it valid there, then why does he fail to conclude

---

[4] In fact, bearing in mind Dretske's acknowledgement that not even a universal association during the learning period is necessary for recruitment (whatever conclusion one comes to in concept learning), it is not even clear that it is the universal association that is causally explanatory as opposed to whatever degree of association is necessary for recruitment. See Slater, 1994, pp. 171–8, and Dreske's reply (1994, pp. 204–5).

that the indication relation is not causally explanatory when considering his own theory? We seem no further forward.

This does not mean that the indication relation is not a sufficient cause of the causal relationship between Ns and Ms. In fact, we already have a story available as to how this is so. Indication relations do not supervene on associations of the sort specified, because indication relations are modal, but indication relations do *partly* supervene on these associations. The efficacy of these associations transmits itself upwards to the indication relation itself. If one adopts such an account, one has to distinguish between whether something causes something and whether it is causally explanatory of that thing. Although the latter implies the former, the cause of something need not figure in the causal explanation of that thing if a more precise characterization can be given.

But it must be obvious that this is just an application of the Segal and Sober approach. So it looks as if Dretske needs to appeal to it to make his own account work. Dretske's dilemma is that either he allows that his indication relation is causally irrelevant, or he admits that his account is not the sole way in which content can be causally relevant. A state's possession of content can be a straightforward triggering cause in the way I have already described.

2.2 Present Effects

The second problem needs careful phrasing. One line of criticism to Dretske's approach has been that it appears as if his account can only allow that meaning was causally relevant in the past. Ns' indications of Es caused Ns to be recruited. They are not a cause of this current piece of behaviour. As Dretske points out, this is not so. The present causal relationship between an N and an M occurred because of the background conditions which hold at the time of the causal relationship. These background conditions are a cause of this causal relationship. But the cause of the background conditions is the previous occurrences of Ns' indicating Es. Since a cause of a cause of X is a cause of X that means that the prior occurrences of Ns' indicating Es caused the present piece of behaviour along with the triggering cause (Dretske, 1988, p. 84). The triggering cause set it off, Ns' prior possession of (natural) meaning settled the way things went from there (Dretske, 1991, pp. 113–4).

However, there is another difficulty. By making the present n a part of the behaviour, Dretske seems to rule out the present n's content from being causally relevant to the behaviour in *any* sense (Segal and Sober, 1991, pp. 23–4). The account he offers of the causally explanatory role of contents seems to talk only of the present effects of Ns' *previously* indicating Es, not the present effects of n's *present* possession of content. Dretske's reply to this objection appears to be this (1991, p. 216):

> [it] is wrong ... to conclude that this shows that the non-natural meaning of n, the current token of N, does *not* explain the same

> behavior. For the meaning, the non-natural meaning, of the present
> n *is*, on this view of non-natural meaning, whatever natural meaning
> (information) in past ns explains the present causal arrangements.
> (my relettering)

What Dretske seems to be arguing is that he has shown that the content of
the current n is causally relevant because Ns were recruited as a cause of
Ms as a result of past ns having the corresponding natural meaning. But, at
best, what this shows is that he has established that something with the *same
semantic properties* as the content that is currently tokened is causally rel-
evant. What he has not done is establish the causal relevance of the current
token of content with these semantic properties.[5]

Jaegwon Kim offered Dretske a solution to this problem (Kim, 1991, p. 67).

> Given that the occurrence of this content property supervenes on
> N, it is appealing to think of the causal role of this content property
> as itself supervenient on the causal powers of the subvenient prop-
> erty N. Thus, whatever causes a given token N-state also causes it
> to be a token R-state, and whatever is caused by this state's having
> R is caused by its having N. In S, the causal powers of R are super-
> venient on the causal powers of N. [R is a content property and S
> a creature with the appropriate learning history.]

And Dretske appears to accept it, but when he gives his version of it, a part
of which I recorded above, it is different (Dretske, 1991, p. 216). The crucial
point is that on Kim's proposal the present token of content is causally rel-
evant because it supervenes on a neural property that is causally relevant
to the occurrence of m, whereas Dretske's version does not explain how the
present token of content is causally relevant.

We have two proposals before us concerning how the present token of a
content is causally relevant, that offered by Segal and Sober and that offered
by Kim. Is there any way we can decide between them? The difference
between these accounts is that whereas Segal and Sober suggest that content
is causally relevant because *part* of its supervenience base is, Kim suggests
that if some creature has an appropriate learning history, then contents are
wholly supervenient on something which is efficacious. A consideration that
appears to favour the Segal and Sober approach is this. One of a content's
causal powers should be its distinctive effect on content-detectors. Segal and
Sober would agree, but apparently Kim in cahoots with Dretske would not.
According to Kim's proposal, a detector of the presence of contents could
only detect them by being able to discern the presence of the relevant learn-

---

[5] Segal and Sober (1991), pp. 25–7, establish this point by a different route. They note that
since the current n is not a structuring cause of the behaviour, its properties cannot be.
But their reason for this is just that the recruitment of N occurred earlier.

ing history. I see no reason for depriving contents of this distinctive causal contribution by claiming that they wholly supervene on the presently instantiated neural properties of a certain kind of creature.

One thing that recommends the Segal and Sober approach is that it explains how a state's possession of content may be a triggering cause. Another is that there need be no debate about whether Dretske has done enough to show that content is causally relevant by showing that Ns' indication of E is causally relevant. Indication lacks an essential feature of representations, indication does not allow for misrepresentation. This has been an oft rehearsed line of criticism of Dretske's approach (in press, Baker, 1991, p. 109). Some support for Dretske's position is derived from the fact that it has seemed to be the only game in town. All other accounts make contents causally irrelevant. But since it appears not to be the only game in town and, indeed, relies upon the existence of another game, it seems reasonable to conclude that although Dretske has pointed to an interesting collection of causal facts, it is far from clear that these display the central causal relevance of content.

Department of Philosophy
University of Nottingham

References

Baker, L. Rudder 1991: Dretske on the Explanatory Role of Belief. *Philosophical Studies*, 63, pp. 99–111.
Dretske, F. 1988: *Explaining Behavior*. Cambridge, MA.: MIT Press.
Dretske, F. 1990: Does Meaning Matter? In E. Villanueva (ed.). *Information, Semantics, and Epistemology*. Oxford: Basil Blackwell.
Dretske, F. 1991: How Beliefs Explain: Reply to Baker. *Philosophical Studies*, 63, pp. 113–7.
Dretske, F. 1991: Dretske's replies. In B. McLaughlin (ed.), *Dretske and His Critics*. Oxford: Basil Blackwell.
Dretske, F. 1994: Reply to Slater and García-Carpintero. *Mind and Language*, 9, 2, pp. 203–8.
Kim, J. 1991: Dretske on How Reasons Explain Behavior. In B. McLaughlin (ed.), *Dretske and his critics*. Oxford: Basil Blackwell.
Segal, G. and Sober, E. 1991: The Causal Efficacy of Content. *Philosophical Studies*, 63, 1, pp. 1–30.
Slater, C. 1994: Discrimination Without Indication: Why Dretske Can't Lean on Learning. *Mind and Language*, 9, 2, pp. 163–80.