

## Article

# Dretske on the Causal Efficacy of Meaning

---

MANUEL GARCIA-CARPINTERO

The object of this paper is to discuss some issues raised by Fred Dretske's account of the causal efficacy of content, as given in his book *Explaining Behavior*.<sup>1</sup> To warrant the causal efficacy of folk-psychological properties while staying within a naturalistic framework, Fred Dretske contends that these properties are causes of a peculiar variety, which he calls *structuring causes*. Yet structuring causes are not postulated *ad hoc*, to somehow account for the causal efficacy of content. Dretske claims that we independently need this dualism of causes to make sense of some of our causal explanations—not necessarily involving mental states.

For several reasons, it seems preferable to establish that content proper-

---

Among the many people who have helped me in articulating the ideas in this paper, I would like to express my gratitude to Ned Block, Manuel Campos, Josep Corbí, Fred Dretske, Anne Mackor, Begoña Navarrete, Daniel Quesada, Robert Stalnaker, Pepa Toribio and two anonymous referees for this journal. I would specially like to thank the invaluable philosophical and stylistic advice by Martin Davies. The paper has greatly improved both in truthfulness and in clarity because of their comments, criticisms and suggestions; any remaining falsehood or confusion should most probably be blamed on my stubbornness. The first version of the paper was completed while I was a visiting scholar to the Center for the Study of Language and Information, Stanford University. I thank the CSLI for the warm and encouraging atmosphere I found there, and the Spanish DGICYT for the grant that made possible my stay at the CSLI. Versions of the paper were given at the fifth SIUCC, held at the University of Valencia, and at the second Karlovy Vary conference; I benefitted from reactions of the audience there. Research for the paper has been funded by the Spanish Government's DGICYT, as part of the project PB90-0701-C03-03. Address for correspondence: Departamento de Lógica, Historia y Filosofía de la Ciencia, Universidad de Barcelona, Baldiri Reixach, s/n 08028, Barcelona, Spain. Email: garcia@cerber.ub.es.

<sup>1</sup> Dretske, 1988. The paper originated in discussions during and subsequent to a seminar which Fred Dretske and Michael Bratman held at Stanford throughout the academic year 1990-91.

ties are causally efficacious in the same sense that physical properties are. The main point I will try to make in this paper is that Dretske's *structuring causes* can be accommodated in a wholly ordinary metaphysical framework. I will try to show that we can make sense of the phenomena that structuring causes are introduced to account for in more conservative terms, having only recourse to the material we must anyway posit to understand ordinary causal explanations.

The issue is of the greatest importance for the philosophy of mind because, if I am right, Dretske's validation of the causal efficacy of content can be cashed out as a very interesting answer to the most commonly invoked type of argument against intentional realism—when intentional properties are understood as 'broad' or 'extrinsic'—namely, the various Twin-Earth considerations.

### 1. *The Problem*

Let me start with a reminder of the problem addressed by Dretske. Our folk-psychological concepts, concepts like *having a desire to drink beer* and *believing that there is beer in the refrigerator* have, as one of the most fundamental roles, that of entering the causal explanation of behavior. Let us suppose that I raise my arm, that the rising of my arm is not the result of something like a reflex, but something I intentionally do. We would say that my intention to raise my arm is a causally efficacious event in the bringing about the rising of my arm, and that it is explanatorily relevant, too. And we would say also that the causal efficacy of the event has somehow to do with its instantiating the type *intention to raise my arm*; that the event is not causing the rising of my arm just 'in virtue of' its physical properties. The familiar contrast is with the soprano's singing in Dretske's inspired example—which seems to be causing the shattering of the glass merely in virtue of its physical properties, and not at all in virtue of its property of being the singing of a word meaning *Help*.<sup>2</sup>

One of the metaphysical problems that may seem to be bothering Dretske is related to Kim's *explanatory exclusion* problem:<sup>3</sup> why do we need the macro-explanation, given that there must be a micro-one? There must be events in my brain that undoubtedly have also to do, from a causal point of view, with the rising of my arm. Moreover, *their* claim to causal efficacy

<sup>2</sup> I am taking for granted a standard picture of causation, generally in agreement with Dretske's. Causation relates particular events relative to some of the types of events they exemplify. I indulge in referring to a type of event that a particular event exemplifies as a 'property' of the event.

<sup>3</sup> See, for instance, Kim, 1987 and 1989. The 'supervenience answer' to the problem of explanatory exclusion that I will be assuming throughout the paper is inspired by Kim's own writings; see, for instance, Kim, 1984a. However, it is clear from those writings he himself does not consider it a good answer; this is also the main thrust of Kim, 1992. I cannot properly discuss the issue here.

seems to be more secure: for one thing, the laws used in the lower-level explanations are less charged with the vagueness of *ceteris paribus* clauses. What can the mental aspects add? Perhaps they add something psychologically or even epistemologically valuable, but what are they adding to the explanation in a stronger, metaphysical sense of 'explanation'?

However, this should not be seen as the problem addressed by Dretske. For, on reasonable assumptions, this is a fairly general problem, affecting the causal efficacy and explanatory value of every 'macro-property', whatever scientific theory it belongs to. But Dretske's solution, to the extent that it is a solution, does not solve this general problem. Even though the solution is not specific to the psychological properties, it has a narrower range. What is the problem he should be viewed as addressing, then?

It is the *supervenience problem*. It will cause no harm, for the purposes of this paper, to assume that the 'explanatory exclusion' problem is averted if the allegedly efficacious but apparently redundant macro-types can be shown to *strongly* supervene on the micro-types doing the 'fundamental' causation; at least, that this is a necessary condition for its solution.<sup>4</sup> Mental types however seem not to supervene on the relevant physical, let us say neurophysiological, types. There are Davidson's and Dennett's considerations to that effect.<sup>5</sup> But we will be considering here only the type of argument which is believed strongest nowadays. Hilary Putnam, and then Tyler Burge, marshalled Twin-Earth arguments and other related considerations to show that contents 'are not in the head'—that they are conceived of as extrinsically individuated.<sup>6</sup> These same considerations have been paradoxically used to show that for contents to be causally efficacious, there must be contents which *are* 'in the head'; for relationally individuated contents cannot be causally efficacious. The argument proceeds from the assumption that contents are environment-dependent, 'extrinsic' or 'broad', to conclude that mental types do not supervene on physical types—for there could be (in Twin-Earth) someone physically like me, but lacking my mental states. Being physically like me, though, my twin would do exactly what I do: his arm would rise. Mental states, then, provided that they are broadly conceived, do not, it seems, supervene on physical events. Their causal efficacy and their explanatory relevance are therefore threatened, 'screened off' by other, intrinsic properties of

---

<sup>4</sup> For the concept of *strong supervenience*, see Kim, 1984b. The properties in family *M* are said to be *strongly supervenient* on the properties in family *N* if any event of type *M<sub>i</sub>*, for some *i*, is also an event of type *P<sub>j</sub>*, for some *j*, and necessarily, any event of type *P<sub>j</sub>* is also an event of type *M<sub>i</sub>*. There is an important and largely overlooked problem related to how the 'necessarily' in the definition should be understood. It will acquire great momentum as the paper proceeds.

<sup>5</sup> Of course, Davidson believes in the supervenience of mental states on physical states. But his sense of the word is generally understood as weaker than the one assumed here.

<sup>6</sup> For the use of Twin-Earth considerations to establish the relational (or 'extrinsic') character of content ascriptions, see Putnam, 1975.

the events instantiating them. As people say nowadays, this renders mental properties *epiphenomenal*. Flying in an aircraft carrying a bomb causes a person's death, but the property the flying has of *being a flight on a Friday 13th* is in the intended sense epiphenomenal in the production of the death. The soprano's singing causes the glass to shatter, but the property of *meaning 'Help'* is similarly epiphenomenal in producing the causation of the broken glass.

## 2. Dretske's Solution: Structuring Causes

Dretske's solution to the supervenience problem requires some conceptual stage-setting. First, Dretske asks us to distinguish between *behavior* and the mere *result of behavior* (say, bodily movement). In the first chapter of *Explaining Behavior*, he develops a theory of behavior (or action broadly conceived) which is at odds with other well-known alternatives, like Davidson's. Behavior should not be identified with bodily movements (or other external 'results' of behavior), according to Dretske. Behavior must rather be viewed as a *process*, the *causing* of the bodily movement by some internal state. Next, Dretske distinguishes between two different kinds of causal explanation: the explanation of a (particular) event, like the result of behavior, a bodily movement—something that he calls *explanation by triggering causes*—and the explanation of a process, of *the causing of an event by another*, what he calls *explanation by structuring causes*.

What precisely is the difference between *triggering* and *structuring* causes? The mere difference between *explaining events* and *explaining processes* is not adequate to distinguish triggering from structuring causes: for particular processes could be explained by triggering causes, too. After all, processes are nothing else than complex events. What geologists give as an explanation of the continental drift is an explanation by means of triggering causes. And exactly the same happens when the historian explains particular processes, like the Russian Revolution or Spain's decline in the seventeenth century. To explain a process, in this sense, is typically to explain the first event in the causal chain constituting the process. This is not what Dretske means by a 'structuring cause' explanation.

*Structuring causes* are the causes invoked in correct answers to a certain sort of contrastive demand for explanation which Dretske introduces by means of examples. For instance, a thermostat has been wired to a garage door opener, so that it opens the garage door when the room gets cold. The question which is correctly answered by mentioning a structuring cause here is not: 'why does the garage door open *now?*', which is answered by mentioning the drop in room temperature, but 'why does the thermostat open the garage door, rather than doing something else (say, turning on the furnace?)' This sort of question is correctly answered by describing how, say, through the intervention of an electrician the relevant states in the thermostat (the bending of its bimetallic strip) got 'wired' to the garage

door opening mechanism. Similarly, pressing the key on the keyboard now *triggers* the cursor's movement on the screen, while the programmer's activities which *structured* that process explain why pressing the key has this effect.

A difference between structuring and triggering causes suggested by the examples, which still does not take us to the heart of the matter, but which is important to emphasize now, is this. While explanation by triggering causes is directly explanation of *particular* events, explanation by structuring causes is directly explanation of *types* of processes, and only indirectly an explanation of the concrete processes which instantiate the types. Thus in the thermostat case what we wanted to know was not anything about the particular instance involved: we wanted to know why the thermostat regularly *behaves* in this odd way. Likewise with the computer's case. This is so with every example provided by Dretske (and examples, together with the contrastive difference between asking for triggering causes and asking for structuring causes we exemplified before, is all he provides).

Besides, it would not do to say that structuring causes are causes of a plurality of tokens. What we want to know, in asking the contrastive question typical of structuring causes, is not what caused *C* to cause *M* whenever a process of this type is instantiated, for in some cases a *C* could cause an *M* in a 'wayward' way. What we want to know is what causes *C* to cause *M* in normal conditions, or, as I said, *typically*. This is why the answer will concern the type and not the generality of the tokens. (It will then indirectly concern, of course, every instance of the type instantiated when normal conditions are satisfied.) This is also why structuring causes are to be viewed as causes of types of processes, even when it is in the nature of the process that it can only be instantiated once, if ever. The turning of the car's ignition key *triggered* the explosion; the terrorist's wiring the bomb to the ignition key *structured* it.<sup>7</sup> Even if something interferes and the process never occurs, however, it still makes sense to speak of the terrorist structuring the process-type, i.e. accounting for the explosion when the ignition is turned *in normal circumstances*.

Whatever our worries about the exact nature of structuring causes, it seems that we can grant the existence of a characteristic pattern of explanation in examples like these. Assuming then the existence and differential nature of structuring causes, let us briefly review Dretske's solution to the supervenience problem. We have already noted that, according to him, behavior is a *process*: some internal event, *C*, causing some bodily movement, *M*. Now, he concedes to the friends of Twin-Earth that meanings, and so states with content like beliefs, desires and intentions, cannot *as such* be triggering causes of bodily movements. Regarding the causation of *bodily movements*, mental properties are epiphenomenal. If we think of

<sup>7</sup> The example comes from Dretske, 1991.

my (voluntarily) raising my arm as merely the resultant bodily movement (i.e. the rising of my arm; let us call it *M*), we must agree that it is not in virtue of the mental properties of some internal state, *C*, in virtue of *C* being, say, my intention to move my arm, that *M* happened. It is only the neurophysiological properties of *C* that are causally efficacious in the production of *M*.<sup>8</sup> Nevertheless, Dretske claims that we should not think of my raising my arm as a mere bodily movement (my arm rising). This is not the *behavior*, but the result thereof. My raising my arm is a process, a piece of behavior: *C* causing *M*. And behavior so conceived may well have *structuring causes*. These are the ones that could figure in a correct answer to the appropriate kind of question: something like 'Why is this neurological type causing this bodily movement, rather than some other movement, or something else?' or 'What *structured* or *shaped* this kind of process?'

It is in answering this sort of question, according to Dretske, that meanings in particular and mental properties in general enter: they are *structuring causes* of behavior. The story is told about very simple mental states. Suppose that *C* is some internal state of an organism, instances of which are caused in normal circumstances by an external state, *F*. (*C* is an *indicator* of the presence of *F*.) Suppose that there is some bodily movement of this organism, *M*, instances of which are appropriate for the organism given that external state *F* obtains (i.e. *M* contributes in those circumstances to the satisfaction of general *needs*—food, sex, etc.—of the organism, or more specific *drives*, even *desires*). Suppose further that, through a process of learning, this organism develops the background conditions (Dretske typically refers to these background conditions as the 'wiring') necessary in order for instances of *C* to cause instances of *M*. Through the process of learning, *C* has acquired the disposition to produce movements attuned to the external circumstances, *F*, movements attuned to the external circumstances because their producer, *C*, correctly *indicates* them.

Now, the fact that *C* had this property of indicating properties of the environment, *F*, partly *explains* (or *causes*) its 'wiring' with *M* during the learning process through a process of operant conditioning, the creation of the standing conditions under which, in the future, the instantiation of *C* would cause the instantiation of *M*. It is thus that mental properties are structuring causes of behavior; a true answer to the question: 'Why is *C* causing *M*, rather than something else?' is 'Because the fact that *C* indicates *F structured* or *shaped* the causing of *M* by *C*'.

It goes without saying that this story needs fine-tuning before we can apply it to ordinary, full-fledged beliefs. But Dretske's account is ingenious enough to suggest that we will be able to extend it to more complex mental states. Moreover, structuring causes being causes—even from a

<sup>8</sup> Dretske, 1990b, p. 12.

stubbornly naturalistic point of view—it seems then that Dretske has solved our problem. But Dretske's solution, as I will try to show, does not depend on the detour through behavior as process and structuring causes. The conceptual apparatus already needed to understand ordinary causal explanations allows us to translate structuring causes into more familiar jargon.

### 3. An Account of Structuring Causes

Let us go back to the nature of structuring causes. It is well known that what we ordinarily call *the cause* of something is no more than a *causal factor*. The throwing of the match caused the fire. But without the presence of oxygen, or inflammable materials, there would have been no fire. Usually we isolate as *the cause* those causal factors that are less ordinary or predictable, or more *salient* to us in some way. Anyhow, all the *background* conditions without which the effect would not have taken place must be counted as causal factors in any philosophical account of causation.

Now, consider Dretske's example. The sudden drop in the room's temperature (call this event *F*) causes the movement of the bimetallic strip inside the thermostat (call this *C*) which, in its turn, causes the garage door to open (call this *M*). Usually, we would choose either *F* or *C* as causes. But neither of them would have caused the garage door to open without the presence of certain background conditions, chief among them, the bimetallic strip being wired to the garage door. If *T* is a type of causal process, Dretske's *structuring causes* are, in a first approximation, ordinary (namely, 'triggering') causes of some background causal factor necessary for some instance of *T* to take place (here, the wiring of the bimetallic strip to the garage door).<sup>9</sup>

However, not every cause of a background factor could be happily used to answer one of the contrastive questions Dretske mentions to introduce the notion of *structuring cause*. Suppose, the driver's suddenly swerving pulled the car off the road, thereby killing Mary, who was just walking by. Suppose further that, even in the presence of the swerve, the car would not have slid off the road had there not been grease on the road at that very point. And suppose finally that the grease was there because a truck spilled it some time before. The grease on the road is then a background causal factor, and the truck's spilling of it is one of its causes. It is nonetheless doubtful that we want to call the spilling of the grease a structuring cause of the swerve causing the death of Mary; i.e. it is doubtful that we would want to answer the question 'why did the swerve kill Mary, rather than causing anything else instead?' by pointing to the spilling of the grease. As a matter of fact, it is rather doubtful that we

<sup>9</sup> As he himself claims in the book, and even more clearly in Dretske, 1991.

want to give *any answer at all* to this kind of question in this case. 'Nothing structured or shaped that process', we would say; 'it just happened by chance, it was a fatal accident'. Structuring causes are then something more than just triggering causes of background causal factors. Part of what was missing in this example is the *generality* of structuring causes we mentioned previously: structuring causes explain *types* of processes; types of process which, in certain 'normal' conditions are instantiated in a regular way.<sup>10</sup>

A convenient way to present my own explanation of *structuring causes* consists in analyzing an example provided by Dretske himself of merely apparent structuring causes. Dretske is eager to point out that there are seeming structuring causes (entities mentioned in seeming appropriate answers to 'structuring causes' questions) that are not causes at all, structuring or otherwise. According to him, 'contents' acquired not through an individual process of learning but through the collective process of natural selection are examples of this. More generally, Dretske does not think that natural selection provides structuring causes. Like many others, I disagree with him about this point;<sup>11</sup> my reasons will be clear by the end of the paper. For the moment, let us concentrate on the analysis of an example he offers to hammer home his point:

Suppose Clyde, an eccentric, collects clocks that lose ten minutes a day. He has scoured the world for them. These are the only clocks he will have in his house. His personal favorite, an old pendulum clock, is named Ben. Do we know why Ben runs slow? Of course not. Though we know why (selectional explanation) all Clyde clocks runs slow, and though we know that Ben is one of Clyde's clocks, we do not know why Ben runs slow. If the truth be known, Ben runs slow because its escapement is bent. The fact that Clyde collects clocks that run slow has nothing to do with Ben's behavior.<sup>12</sup>

Compare the thermostat's example. The triggering cause of that instance of *M* (the opening of the garage's door) is the particular instance of *C*, the bending of the bimetallic strip—triggered in its turn by the drop in

<sup>10</sup> The features of that particular spilling of the grease could well explain why the swerve killed Mary, instead of some other pedestrian. But the content of the contrastive question which so receives an answer in terms of features of that particular spilling is different from that of those asking for structuring causes. The contrastive question in Dretske's thermostat example asks why the drop in temperature would open the garage's door, *whenever* it occurred in normal circumstances. By contrast, the question in the example is a request for information about that particular spilling.

<sup>11</sup> See the papers by Horgan, Cummins and Dennett in McLaughlin, 1991, and the ones by Millikan and Stampe in the 'Book Symposium' on *Explaining Behavior in Philosophy and Phenomenological Research*, 50 (1990).

<sup>12</sup> Dretske, 1990a, p. 827. See also Dretske, 1988, pp. 89–95.



temperature. In this case, it also makes sense to ask for a structuring cause of the type of process, *C* causing *M*; i.e. why the thermostat causes the door to open, instead of causing something else. This question is answered by mentioning the activities of the electrician who brought about the necessary 'wiring'. In Clyde's example the triggering cause intended to correspond to *C* is Ben's escapement being bent, and the intended instance of *M* is Ben's running slow. However, it seems clear that the fact that Ben is one of the clocks belonging to Clyde, who collects only clocks which run slow, does not provide us with a structuring cause of the type of process *Ben's escapement being bent causing Ben to run slow*. It is even doubtful that this type of process has any structuring cause. But why?

Dretske's analysis is that Clyde's selectional activities apply not only to the particular processes going on in Ben, but to many other clocks in sweeping generality. That a clock has been selected by Clyde in virtue of its running slow does not say anything about the specificity of the processes explaining that fact in that particular clock. Dretske intends this as an analogy which can be used to argue against the causal efficacy of characteristics secured by natural selection. I do not accept Dretske's analysis of this example or his conclusion by analogy; I will briefly discuss the issue of natural selection by the end of the paper. Here is my alternative analysis, though: if there is no structuring cause provided by Clyde's selectional activities it is not because selection processes are too general; it is because there is no *lawful* or *nomic* connection between a clock having the property of belonging to Clyde, who collects clocks running slow, and the existence, proliferation and preservation of internal mechanisms in the clocks causing them to run slow. How could there be, given that the 'shaping' of the internal mechanism (the bending of the escapement, and other explanatory links between this state of affairs and the running slow of the clock) took place well before the selection by Clyde and was not sustained in any way by Clyde's selection? If Clyde is powerful enough, his selectional activities may help understanding a coincidence, namely, that the proportion of clocks running slow is so big. But there still will not be a nomic relation between Clyde's selectional activities and the existence of clocks with certain characteristics, and thus no structuring cause; such a nomic relation would require that those activities lawfully help the sustained reproduction of traits forcing the clocks to run slow. (There is not such a nomic relation the way the story has been told; if it is told in such a way that Clyde's activities do account for the reproduction of traits making the clocks to run slow, then there will be a structuring cause of the process, constituted by those very activities of Clyde's.) Clyde's selectional activities might causally depend on the existence of clocks with internal states that cause them to run slow; but the existence of clocks with internal states that cause them to run slow does certainly not causally depend on Clyde's selectional activities.

This analysis gives us a clue to what is really going on with Dretske's new type of cause. As I said before, structuring causes are ordinary causes

of background causal factors, but not every ordinary cause of a background causal factor counts as a structuring cause. We ask about structuring causes when we have a type of process, *C* causing *M*, and we want to know why *C* normally causes *M* instead of something else. This, in a nutshell, is my proposal: structuring causes are in order *only* when *C* has some *function* having to do with the causation of *M*. The function could be just the production of (an instance of) *M*, or (as it will be the case with meaning-functions), the production of something appropriate when environmental condition *F* obtain, etc. And structuring causes ultimately are (partial) supervenience explanations of these functional properties: they (partially) explain why the lower level property *C* 'realizes' or constitutes the higher level functional property. Let me explain.

Let us agree on a convenient understanding of 'function'. It is essential for my purposes to take 'function' not merely in the dispositional, forward-looking sense of the term, but in the biological, both forward- and backwards-looking sense of it. I will assume Larry Wright's analysis of 'function': *C* has the function *F* if and only if (a) *C* is there because it causes *F*, and (b) *C* has *F* as a causal result in normal conditions. The 'because' in (a) can be spelled out in different terms: natural selection, learning, conscious design, etc. In all of them it has an etiological, or causal-historical, content: actual instances of *C* caused *F* in the past, and that is causally related to the existence of present instances of *C*. The 'is there' typically mean: 'states like *C* keep being instantiated', or 'are preserved', etc.<sup>13</sup>

This is where structuring causes enter. Whenever we intuitively feel that a structuring cause explanation is not in order (Mary's accident, Ben's running slow), the reason is that what happened was not 'designed' or 'supposed' to happen: it was not a function of *C* to bring about *M*. Whenever we do have a function, asking for a structuring cause is in order. And, in those cases only, structuring causes are partial explanations of the processes underlying part (a) of what any functional ascription involves, according to the analysis provided before. For a (type of) state *C* to have function *F*, it must cause states of type *F*, and there must be an etiological account of how this causing of *F*'s by *C*'s got established. The function of the heart is to pump blood. This means, according to the foregoing analysis, that there is some heart-event, call it *C*, which in normal conditions causes the pumping of blood in the organism, call this *M* (part (b)); and also that the heart, and *C* with it, 'is there' because it pumps blood (part (a)). A detailed justification of this last part is an account of how the process-type *C* causing *M* 'got structured'; it is an account of the creation of the standing conditions (Dretske's 'wiring') under which *C* is normally a cause of *M*. Darwin's idea of natural selection

<sup>13</sup> See Wright, 1973. See also Millikan, 1989. Millikan claims that Wright's concept of function, in opposition to hers, is a dispositional one, merely looking forward, not backwards. I do not think so, but I cannot dwell on the issue here.

allows us to do this job for biological functions in a naturalistic framework; naturalistic accounts of learning, let us hope, will provide something similar for typically psychological functions. If I am right, then, we can translate explanations by structuring causes into a more familiar jargon: to give a structuring cause explanation of the type of process *events of type C causing events of type M* is to give a (perhaps partial) explanation of how *C*-structures have the function of producing *M*-events.

Dretske does not give a theoretical definition of structuring cause; he just gives several examples and the intuitions regarding differences in contrastive questions about them. As I stressed before, a definition of structuring causes as triggering causes of background causal factors in any process is incoherent with the intuitions exploited in the examples.<sup>14</sup> Lacking such a definition, I am constrained to defend my explanation on indirect grounds. I will show that the account fits the examples and our intuitions about them; that it explains the facts about structuring causes we have uncovered before; and that it allows us to see clearly the precise nature of Dretske's solution to the supervenience problem. Let us see first that the account fits Dretske's examples well. If we feel as if a structuring cause explanation is in order in the thermostat case, it is because we surmise that the thermostat's weird behavior has been intentionally designed; the structuring cause explanation accounts precisely for how the thermostat has acquired the function of opening the garage's door through the electrician's wiring activities. On the other hand, if the link between the thermostat and the garage's door had occurred only by chance, by some bizarre accident, then there would be no structuring cause of the process. Similarly with the computer and the bomb examples.

No wonder that the present account helps us to better understand why Clyde's selectional activities do not give a structuring cause of some internal process in the clocks causing them to run slow—for it was suggested by it. The fact is, we do not have here any reason to think that the internal process (as it happens, the escapement being bent) has the *function* of causing the clocks to run slow. This is so because the bending of the escapement is not there because it caused this clock (or any other clock from which this could have inherited the bending of the escapement)

---

<sup>14</sup> Dretske suggests in Dretske, 1991 such a general definition—although in some other places he acknowledges that structuring causes are typically in order with biological and psychological processes, and in Dretske, 1990b p. 10 he finds natural that a structuring cause of the process *C causing F* is in order when *C* 'is supposed' to bring about (an instantiation of) *M*. I think a possible explanation for this mistake is his habitual conflation of the 'explanatory exclusion' with the 'supervenience' problem, common to many philosophers. As I said before, the explanatory exclusion problem is a fairly general one, affecting every special science; if you think you are looking for a solution to it, and also that structuring causes provide one, you will be prone to think that structuring causes are generally available. There are many texts, however, in which Dretske clearly shows that he is confronting the supervenience difficulty. See, for instance, Dretske, 1991, Section 2.

to run slow in the past. Therefore, part (a) is not here satisfied. At any rate, Clyde's selectional activities do not refute this claim. On the contrary, the reasonable bet is that Clyde's activities and the existence of clocks whose bent escapement causes them to run slow are causally independent (or maybe they are causally related, but in the opposite direction to the one we need here: the bending of the clock caused it to be selected by Clyde, but not the other way around). This is another way of making the same point we made before: there is no *lawful* or *nomic* connection between the internal structure (the escapement being bent) and any function it could have (like causing the clock to run slow) which could be accounted for by Clyde's selectional activities.

The current proposal explains an aspect of Dretske's account I have already pointed out, namely, that structuring causes are *only indirectly* explanations of *particular* events. Directly, they are explanations of *types* of event. On my interpretation, that is easy to understand. There are two very different sorts of explanation. The first consists of explanations of events, 'particular matters of local fact', by means of (ordinary) 'horizontal' laws, either macro-laws or micro-laws: he got lung cancer because he was a heavy smoker. Dretske's triggering causes belong in this group. The second is the explanation of macro-laws by means of micro-laws: heavy smoking causing lung cancer is explained by such-and-such biochemical processes. Dretske's structuring causes, according to my account of them, belong in this second group. They partially explain how the events of a certain lower-level type (say, neurological structures) instantiate also a higher-level, functional type, by giving the details of some of the etiological facts that must obtain for any structure to have teleo-functional properties. Hence, it is of necessity that structuring causes explain types; and also, that the explanation applies to particular cases instantiating the type only when conditions are *normal*. 'Normal' here is to be understood not statistically, but teleologically.<sup>15</sup>

'Horizontal' explanation concerns the explanation of individuals, 'vertical' explanation the explanation of kinds. We commonly put them together. For, according to a certain physicalistic intuition (assumed as correct in giving the partial answer to the explanatory exclusion problem I put forward before, namely, that strongly supervenient properties are indeed causally efficacious), for a macro-law to explain a particular causal process, there must be (whether or not we know it) an explanation of the same process by means of micro-laws. Thus, one way of showing that a macro-event (his being a heavy smoker) caused something is to justify that the macro-type the event exemplifies can be explained by micro-types it exemplifies too, and that we can explain why the causing took place precisely in virtue of the event exemplifying the micro-types: he got lung cancer because he instantiated such-and-such biochemical process, starting

<sup>15</sup> See Millikan, 1984, Chs 1 and 2, and Millikan, 1989.

with such-and-such biochemical property, which is what being a heavy smoker came to be in this case. This same 'putting together' takes place when Dretske claims that structuring causes indirectly cause particular events, by causing the types of process they exemplify (and, we may add, circumstances are normal).

#### 4. *The Causal Efficacy of Meaning*

I said before (in Section 1) that Dretske's solution cannot be understood as answering to the *explanatory exclusion* difficulty, for this is a quite general problem, affecting every property posited as causally efficacious in any of the 'special sciences', while Dretske's solution addresses a more specific problem. Now we see why this is so; the problem of explanatory exclusion affects meteorological properties as much as psychological ones, while there is no hope of solving it for meteorology or geology by having recourse to 'structuring cause' explanations. All along I have been assuming that the explanatory exclusion problem should be answered by establishing that strongly supervenient properties are indeed causally efficacious, and I have taken for granted that it could be solved in that way. Dretske's proposal, I said, was to be taken as aiming to solve a problem specific to psychology; and this was the failure of supervenience allegedly proved by means of Twin-Earth thought-experiments.

How does the appeal to structuring causes, understood accordingly to my proposal, solve this problem? For a macro-type, in general, to be causally efficacious in a given causal transaction, it has to be *explained* by physical types operating in the same causal transaction which account for its causal powers. To prove that this 'vertical' relation of explanation obtains is to establish the supervenience of the macro-types on the micro-types; or so I submit. The supervenience relationships we are considering here are not *conceptual*, in contrast with the supervenience of evaluative properties on descriptive properties. It is my view that supervenience relations of the a posteriori variety we are concerned with here *reduce* to explanatory links between macro-types and micro-types. I cannot argue for this view here. I will restrict myself to stating that it is to me a basic metaphysical contention that supervenience facts cannot be primitive; they must be conceptually reducible to—or at least conceptually supervenient on—other facts: facts about meaning, in the case of the supervenience of evaluative on descriptive properties; facts about explanatory 'vertical' relations, in nomic cases like the one we are discussing.

Structuring causes, according to the present proposal, are but particular cases of these vertical laws: partial supervenience explanations for functional properties. To establish that there is a structuring cause of the heart-event *C* causing the pumping of blood is to explain in lower-level terms the having of the functional property by the heart, and thus to show

how the functional property—even though relational—supervenes, in the relevant sense, on lower-level properties.

Consider now mental causes as structuring causes. How do meanings fit in a teleological account? What kind of functions are meanings? Dretske's first idea for the naturalization of content was the *indicator* account, extensively developed in Dretske (1981). Very roughly, a structure *S* has the content *F* if and only if instances of the structure are caused only by instances of *F*. However, Fodor criticized forcefully these non-teleological causal theories of content—'indicator' or 'informational' theories. Fodor (1984) convincingly argued that they had what has come to be known as the 'disjunction' or 'error' problem. If you explain the fact that 'cow' means *cow* in terms of the existence of a nomic or causal dependence between instances of the term 'cow' and instances of *cow* (that is, cows), and you want to say that this instantiation of 'cow' counts as a mistake, because it has been produced by a horse in the dusk instead of by a cow, then you have to explain why the term 'cow', according to *your own* account, does not mean *cow or horse-in-the-dusk* instead of just *cow*. For otherwise the present instantiation of 'cow' would count, after all, as correct. The idea now is to distinguish 'normal', meaning-giving situations, in which the nomic relations between the sign and the world determine its content, from other situations in which things may go wrong and instances of the sign might be caused by abnormal events. Teleology provides naturally for these situations; they are the ones we need, according to the etiological explanation of 'function', to account for the 'being there' of the structure carrying the function.

However, there is a mismatch in the direction of causality between the structure which has the function and the function *germane to the definition of 'function' we have given*, and the direction to be expected from a 'teleologized' indicator account. As I have emphasized before, the function is something the structure carrying it *does*; while, if content is the function, and we want to account for it by resorting to an indicator account, the function *will cause* the structure carrying it. Dretske's teleological definition is this (Dretske, 1986, p. 22): a belief-like structure *C* has content *F* if *C* has the function of carrying the information *F* (or: *C* has the function of being an indicator of *F*); as before, '*C* carries the information *F*' (or '*C* indicates that *F*') is ultimately explained in terms of causal connections *from F to C*. But what is it for a structure *to have the function of carrying information*? What is it that these structures *do* in virtue of having the properties accounting for their carrying the function? The 'structuring cause' account makes it all come down to this: first, structure *C* indicates *F*; then, because of that, it is 'recruited' or 'wired' in the appropriate background to produce behavior 'good' to the organism in which it is placed *when F*. Hence, *C has the function of indicating F* if there is a historical account of the proliferation and preservation of instances of *C* which mentions the fact that there were conditions ('normal conditions') such that *C*, being an indicator of *F*, was instantiated only if *F* was instantiated

in them. The function of *C*, then, is to produce bodily movements *appropriate when external condition F obtains*; in the very simple cases we consider, the bodily movements belong to a single type, *M*. In more complex cases, we should drop any reference to a specific *M* and consider the joint causation of bodily movements by beliefs and desires. We would obtain thereby a teleologized version of the old, non-teleological functionalist account of mentality. Mental states are states which have certain causal relations among them, with inputs and outputs, and which 'are there' because they enter those causal relations.

Let us, however, keep to Dretske's simple examples to see what we get when we read Dretske's structuring causes account of the causal efficacy of content according to the proposal in the former section. For the content *that F* ('food here') to be causally efficacious, there must be an account of the sustained proliferation (the 'being there') of instances of a structure, *C* (some neurological state), in terms of the creation of the background (neurological, etc.) conditions such that *C*, being caused by *F*, produces in its turn bodily movements (*M*) appropriate for the organism in which *C* happens when *F* obtains. This is what the 'operant conditioning' account outlined before is supposed to fill in. As I warned before, it is obviously too narrow an account to be directly applicable to provide a naturalistic theory of full-fledged intentional states, but we may hope to be able to proceed from here.<sup>16</sup> The point is that we have in it the makings for an account of the relevant supervenience of a meaning-property (being in a state with the content *that F*) on lower-level properties (being in *C*). We saw before how the 'structuring cause' idea solves the problem of the causal efficacy of meaning, in these simple cases: the fact that *C* had the property of indicating properties of the environment, *F*, partly *explains* its 'wiring' with *M* during the learning process, the creation of the standing conditions under which, in the future (and in normal conditions), the instantiation of *C* will cause the instantiation of *M*. This, according to my explanation of 'structuring causes', is just to provide a partial account of how the event having the higher-level property of being a belief that *F* *supervenes on* lower-level properties of it, like its instantiating the neurophysiological property *C*—insofar as we take into account a teleological analysis of *being a (proto-)belief that F*, i.e. once we understand by this *having the function of producing bodily movements*—like *M*—*which would satisfy such-and-such needs of the organism when F*.

It might not be immediately obvious how to get from this an answer to the Twin-Earth conundrum. It is agreed that meanings are a particular kind of teleo-functions, and that Dretske's vindication of their causal efficacy through the structuring cause idea has somehow to do with it. He thinks he has provided for the causal efficacy of meaning by finding,

---

<sup>16</sup> See Proust, 1993 for convincing criticism of the narrowness of the operant conditioning account.

over and above ordinary triggering causes, a new variety of 'horizontal' causation. I claim that structuring causes are just cases of 'vertical' causation, when the macro-property to be accounted for in lower-level terms is a function, and that this secures the causal efficacy of the macro-level property (the function, in particular the contentful state) by nailing down its supervenience. But, it might be wondered, how can this be so? Is it not even clearer now that content properties are 'extrinsic'? How can they then supervene on intrinsic properties of the organism?

This is the clue to the issue: it is not *conceptual* supervenience that must be proved (or, more sensibly, that must be shown not to have been disproved) before we can claim that a macro-type is causally efficacious. This is too stringent a condition. To put in jeopardy the relevant kind of supervenience of the mental on the physical, it is not enough to show that it is conceivable that someone physically like me could be mentally different from me. A thought experiment is not enough. What must be shown is that the possible world in question is a *physically* or *metaphysically* possible relevant one. And that is not so easy. A mere *epistemic possibility* will not do.<sup>17</sup>

Do the common Twin-Earth thought-experiments establish more than epistemic possibilities? What Dretske's story does is to show how the neurological properties of a particular state could *explain*, and so *constitute*, its mental properties. He does this in two parts: first, by means of a conceptual analysis of mental properties in teleological terms; second, by giving a plausible account (the structuring cause story) of how the state has the mental-functional property just in virtue of its having the neurological property. For the structuring cause story tells us that by instantiating the neurological property the state has also what happens to be the mental property; i.e. the story tells us (tell us *if* and *when* the details have been filled up, and insofar as it is true) that a state with those characteristics 'is there' just because it disposes the organism in which it takes place to behavior appropriate to certain external circumstances. And this is all that is required to claim that, even if we can think of (epistemically) possible circumstances in which a neurological structure like that lacks the functional property, *this* structure (or its counterpart) has the mental properties in every physically possible world which it inhabits. (Also, given plausible further premises for which I have not the space to argue here, in every *metaphysically possible* world.)

<sup>17</sup> There is much confusion in the literature about this point regarding supervenience. For instance, in his defense of the explanatory value of 'extrinsic' or 'broad' intentional properties in Burge, 1986, Tyler Burge claims that supervenience is not necessary for explanatoriness. But he obviously has in mind *conceptual* supervenience, for he observes in several places along the paper that the possible worlds establishing the failure of supervenience are probably not physically possible. On the other hand, when Fodor argues against the explanatory value of non-supervenient properties (for instance, in Fodor, 1987, ch. 2) he obviously has in mind nomic supervenience.



Consider again the particular causal transaction in the usual example, *c*, my intention to quench my thirst by drinking this glass of water, causing *e*, my reaching-behavior; and the claim that a content property of *c*, *F* (being an intention to quench my thirst by drinking from the glass of water in front of me), is causally efficacious in bring about *e*. What Dretske tells us is that, insofar as there is a 'vertical' explanation of how the neurological property *C* of *c*, whose causal efficacy is not under dispute, realizes the meaning property *F*, any situation we might conceive of in which *c* has *C* without having *F* is just irrelevant to establish the truth of the counterfactual *c* could have had the same intrinsic properties while lacking *F*. For the truth of such counterfactuals, taking the modality in the relevant sense, is disproved just by the existence of the vertical explanatory relations. In the presence of such explanations the Twin Earth-like situation is just an epistemic possibility, not a nomic one.

If this answer does not sound right, it is because Dretske's (or Millikan's, for that matter) theory cannot be directly applied to very complex states, like any water-belief. The story applies clearly only to very basic belief-like states; to extend it to the ones under consideration here, we should first spell out their conceptual connections to the basic ones. I do claim that the point applies also to water-thoughts, or even to elm-thoughts and arthritis-thoughts, provided that we have already established that it applies to more basic thoughts and we take into account the necessary conceptual connections of those other more complex thoughts to the more basic ones (their 'conceptual roles'). And we do not have more than rough ideas about what a theory like that would be like. By way of example, and without any commitment about the correctness of the details, let us consider instead more basic intentional states, to appreciate the point.

Just for the sake of the example, let us assume, with David Hilbert (see Hilbert, 1987), that color-words refer to classes of *reflectances*, and that color-sensations are externally individuated (in part) by means of them: they are states whose (biological) function is to cause behavior appropriate to the presence of a surface with a reflectance among the ones in such-and-such a class, by being themselves caused by the presence of such a surface. It might well be that there is a neurophysiological state, *C*, that explains, for all (normal) human beings, what it is to have a sensation of green, the term 'sensation of green' understood as I have just said. An explanation like that would tell us how *C* is caused (in normal circumstances) by the presence of a reflectance in the class *green*, and how *C* was 'wired' to cause bodily movements appropriate to the presence of a reflectance in that class. Now, in the presence of an explanation like that, the fact that we can easily think of possible worlds such that a structure which has *C* is a sensation of red—i.e. one caused by surfaces in the 'red' reflectance-class and producing behavior appropriate to *them*; just think of 'inverted-spectrum'-like situations—or even the taste of Vegemite, or one that lacks any meaning whatsoever, is just irrelevant. Relative to the question of whether or not this structure having the property of

being a sensation of green supervenes on some of its physical properties, those possible worlds are completely beside the point. Sheer intuition cannot warrant a negative response, because what is in question is the existence of a contingent nomic link, only discoverable by empirical means. (Remember that I take it that nomic supervenience conceptually reduces to 'vertical' explanation.)

To summarize: assume *c* and *e* are particular events, and that allegedly *c* causes *e* in virtue of its instantiating  $C_{\text{macro}}$ , a relational property of events. Let  $C_{\text{micro}}$  be an intrinsic, microphysical property of *c* accounting in the most basic terms for the particular causal transaction. Then (the 'explanatory exclusion' problem in abeyance) the existence of an explanatory connection from  $C_{\text{micro}}$  to  $C_{\text{macro}}$  (together with other facts about explanatory relevance that we have not considered here) averts any alleged 'screening off' of  $C_{\text{macro}}$  by  $C_{\text{micro}}$ , for it follows from this that in any physically possible world, *c*, in virtue of its instantiating  $C_{\text{micro}}$ , would instantiate  $C_{\text{macro}}$ . Whether or not this explanatory connection obtains is a contingent matter, to be solved by empirical investigation. A thought experiment, on the other hand, can only establish an epistemic possibility, and therefore, when intended as an argument to the opposite, begs the question against the believer in the causal efficacy of externally individuated content properties. On the other hand, Dretske's appeal to 'structuring causes' is not more of the same, the attempt to prove the existence of a contingent link by philosophical reasoning, but just a way of anticipating science, by giving a plausible outline showing how microstates could *explain* extrinsically individuated content properties: the selection of instances of the microstates by means of learning processes and their preservation would thereby constitute their instantiating those content properties too. Thus, Dretske's appeal to learning processes is one way of anticipating a detailed account of how precisely the Twin-Earth contender may be begging the question.

Of course, it seems not only *physically*, but perhaps even *technically* possible to build a physical replica of a human being. And it could be argued that, lacking an appropriate history, this thing would lack the mental states of its duplicate. Whatever the correct stand on the contentious issue whether or not it follows from a teleological account that this imagined ultimate creation of science would lack mental states, I would like to point out that nothing I have said is intended to contradict its sheer possibility. *c* having  $C_{\text{micro}}$  could well explain its having  $C_{\text{macro}}$ , even if some other particular event might have  $C_{\text{micro}}$  and lack  $C_{\text{macro}}$ . (By virtue of technical improvement or just by accident, say, a thunderbolt striking a rock miraculously turns it into an instance of  $C_{\text{micro}}$ .) For the macro-laws in which the macro-types appear hold only *ceteris paribus*; therefore, the micro-explanations of them are only supposed to work for instantiations *in normal conditions*, when everything is equal.<sup>18</sup> (Of course, they

<sup>18</sup> Stalnaker, 1989 makes what I take to be a similar point.

should also explain how in cases in which *cetera* are not *paria* things can go wrong, etc.) What matters to determine whether my present intention to raise my arm is causally efficacious in the causation of my arm rising is only whether the neurological state causing the rising of my arm explains, and so constitutes, an intention to raise my arm *everything else being equal*. If something like Dretske's story is true of my intention, and those of my fellow human beings, etc., we are entitled to contend that the required explanatory relation obtains, and that a situation in which the same neurological state has been produced Frankensteinwise is not covered by the *ceteris paribus* clause. What matters is *nomic ceteris paribus supervenience*.

Part of the opposition to Dretske's account has come from people correctly resisting his contention that, say, my intention of raising my arm is not a causal factor of my arm movement *in the same sense* in which the hammer hitting the walnut is a causal factor of the walnut cracking into pieces.<sup>19</sup> If we see Dretske's proposals the way I have analyzed them here, we do not need to say that structuring causes are causes of *processes*. A structuring cause of a process, *C causing M*, is a partial explanation of why a certain higher level property supervenes on *C*. But if we have that explanation, we are thereby entitled to claim that a particular instance of the higher level property *itself* is a *triggering* cause of *M in virtue of* instantiating the higher order property. The fact that the event being *C* causally explains the instantiation of an *M* does not preclude the causal role of the higher level property of that same event. On the contrary, the supervenience of the higher level property on the lower level property *C guarantees its own causal efficacy*. (Assuming, of course, that we have a solution to the more general 'explanatory exclusion' problem.) Dretske's account would therefore allow for beliefs and desires to be taken as ordinary causes of bodily movements (and other more complex events), as they are commonly taken to be.

The point I have belabored here is independent of the mechanism accounting for the historical part of the functional ascription being natural selection, learning, or any other, like convention. I do agree with Dretske that learning processes are responsible for the most interesting content properties. However, I think he is wrong in rejecting the possibility of causally efficacious innate contents and, in general, causally efficacious innate functions. I also think that his arguments to that effect do not establish the point. The gist of Dretske's justification can be gleaned from his own analysis of the Clyde example discussed before. His own reason to reject that there is any structuring cause there of the process the *escapement being bent causing the clock to run slow* is that alleged structuring causes are not real structuring causes when they are somehow *general*, when, as he puts it, they do not explain appealing to the properties 'of

<sup>19</sup> See, for instance, the papers by Horgan and Adams in McLaughlin, 1991.

*this indicator*', but to the properties of a *class* of indicators instead. He thinks this is what is common to Clyde's example and to alleged explanations by natural selection. In Clyde's example, we are given a property common to several clocks; in selectional explanations, a property common to all the individuals involved.<sup>20</sup>

However, this is a mistake, a mistake for which the sloppiness throughout the book on the type/token distinction that I am not the first in pointing out—and Dretske himself now acknowledges—is to be held responsible. The only 'generality' that can be found in natural selection explanations is also present in explanations by means of structuring causes which are acceptable by Dretske's standards. Even Dretske's real structuring causes are general, because as we have seen they are explanations of process-types, and so they explain every (normal) instance of the type. But it is only in this same sense that explanations by natural selection are general. There is indeed a difference between natural selection explanations and learning explanations, namely, that in the latter the processes typically are instantiated in full detail only in one individual, while in the former they are instantiated in several; but this is wholly immaterial, for in the learning cases the explained processes can also be (and usually are) instantiated *at different times* in the same individual. The problem in Clyde's example does not lie in generality *per se*. The problem lies in the generality (the general link between the existence and proliferation of clocks with internal states causing them to run slow and their running slow provided by Clyde's selection activities) being merely *accidental* in Clyde's example, in opposition to what must happen whenever some event actually has a function. But this *essential* generality can be given both by learning and by natural selection. Natural selection, too, provides a nomic, counterfactually supporting link (while Clyde's activities do not) between the proliferation of instances of the structure having the function and the performance of this function.

Dretske does not present his views as I have done. For one thing, he contends in several places that to accept externalism about content (as he does) is to commit oneself to failure of supervenience. But the reasons he offers for that claim make me think that he, too, is mistakenly speaking of *conceptual* supervenience in those places. Furthermore, there are the obvious differences in the phrasing of the two views, his and the one I have espoused here. But no matter how much our views on supervenience and related issues differ, there is a fundamental agreement between us. I think I have learnt from him that the clue to the conceptual dispute about the causal efficacy of mental properties lies in the existence of explanatory links between those mental properties and the properties whose causal efficacy in the production of bodily movements is not in dispute. Our intuition that mental properties are causally efficacious comes down to

<sup>20</sup> See Dretske, 1988, pp. 124–25.

the belief that those explanatory links exist. Dretske's story in terms of operant conditioning is a way of deepening that intuition. A good conceptual theory of explanatory relevance and causal efficacy should allow for the causal efficacy of properties, when those explanatory links obtain. And a good conceptual theory of mental properties should indicate where to look for them. Dretske and I disagree about the extent to which this can be done within the same conceptual framework we need to account for the causal efficacy of, say, meteorological properties. I am not sure whether there are facts to solve this disagreement. Our intuitions concerning mental properties are opposing in such a way that no matter how we choose to talk, as Dretske says in a different connection, we are probably bound to talk funny. I think we do not need to talk as funny as we would do if we adopted Dretske's full account when we translate it the way I have proposed, and that the reasons for thinking otherwise (the compelling intuition that externally individuated mental properties do not supervene on intrinsic features) are misguided.<sup>21</sup>

*Departamento de Lógica, Historia y Filosofía de la Ciencia  
Universidad de Barcelona  
08028 Barcelona  
Spain*

### References

- Burge, T. 1986: Individualism and Psychology. *Philosophical Review*, 95, 3–46.
- Dretske, F. 1981: *Knowledge and the Flow of Information*. Cambridge, MA.: MIT Press.
- Dretske, F. 1986: Misrepresentation. In R. Bogdan (ed.), *Belief*. Oxford University Press.
- Dretske, F. 1988: *Explaining Behavior*. Cambridge, MA.: MIT Press.
- Dretske, F. 1990a: Reply to Reviewers. In *Philosophy and Phenomenological Research*, 50, 819–39.
- Dretske, F. 1990b: Does Meaning Matter? In *Information, Semantics and Epistemology*. E. Villanueva (ed.). Oxford: Basil Blackwell, 5–17.

<sup>21</sup> The first version of this paper was completed before McLaughlin, 1991 appeared. After reading an earlier draft, someone drew my attention to the paper by Kim included in it (Kim, 1991) and I discovered that he makes a similar point about structuring causes. Having developed independently, though, the arguments, the links with other problems (like the Twin-Earth argument), and several nuances are different. Also, I recently listened (in a seminar in Ciudad de Mexico, August 1992) to Jerry Fodor defending similar views on the right answer to the Twin-Earth conundrum to the ones proposed here; my views on this issue developed independently of his, too. My views are certainly not independent of Kim's and Fodor's contributions to the field, without which I might not have any views on these matters.

- Dretske, F. 1991: Mental Events as Structuring Causes of Behaviour. In J. Heil and A. Mele (eds), *Mental Causation*. Oxford University Press.
- Fodor, J. 1984: Semantics, Wisconsin Style. *Synthese*, 59, 231–50.
- Fodor, J. 1987: *Psychosemantics*. Cambridge, MA.: MIT Press.
- Hilbert, D. 1987: *Color and Color Perception*. Stanford: CSLI.
- Kim, J. 1984a: Epiphenomenal and Supervenient Causation. *Midwest Studies in Philosophy*, 9, 257–70.
- Kim, J. 1984b: Concepts of Supervenience. *Philosophy and Phenomenological Research*, 45, 153–76.
- Kim, J. 1987: Explanatory Realism, Causal Realism, and Explanatory Exclusion. *Midwest Studies in Philosophy*, 12, 225–39.
- Kim, J. 1989: Mechanism, Purpose, and Explanatory Exclusion. In J. Tomberlin (ed.), *Philosophical Perspectives*, 3: *Philosophy of Mind and Action Theory*. Atascadero, CA.: Ridgeview Publishing Company.
- Kim, J. 1991: Dretske on How Reasons Explain Behaviour. In B. McLaughlin (ed.), *Dretske and His Critics*. Oxford: Basil Blackwell.
- Kim, J. 1992: Multiple Realization and the Metaphysics of Reduction. *Philosophy and Phenomenological Research*, 52, 1–27.
- McLaughlin, B. (ed.) 1991: *Dretske and his Critics*. Oxford: Basil Blackwell.
- Millikan, R. 1984: *Language, Thought and Other Biological Categories*. Cambridge, MA.: MIT Press.
- Millikan, R. 1989: In Defense of Proper Functions. *Philosophy of Science*, 56, 288–302.
- Proust, J. 1993: Naturalizing Intentionality through Learning Theory. Forthcoming in the *Proceedings of the XVIth International Wittgenstein Symposium*.
- Putnam, H. 1975: The Meaning of 'Meaning'. In H. Putnam, *Mind, Language and Reality: Philosophical Papers, Volume 2*. Cambridge University Press.
- Stalnaker, R. 1989: On What is in the Head. In J. Tomberlin (ed.), *Philosophical Perspectives*, 3: *Philosophy of Mind and Action Theory*. Atascadero, CA.: Ridgeview Publishing Company.
- Wright, L. 1973: Functions. *Philosophical Review*, 82, 139–68.