The Journal of Mind and Behavior
Summer 2001, Volume 22, Number 3
Pages 249–262
ISSN 0271–0137

Structural Causation and Psychological Explanation

Jeffrey Hershfield

Wichita State University

A key test of any philosophical account of the mind is its treatment of mental causation. Proponents of the token-identity theory point to its strengths in both "demystifying" mental causation — by identifying mental causes with the physical causal mechanisms underlying bodily movements — and in avoiding commitment to dubious forms of causal overdetermination. I argue against this account of mental causation, pointing out that it mistakenly identifies actions with bodily movements. I suggest instead treating action explanations as explanations of redundant causalities in behavior, and the mental causes cited in such explanations as structural causes.

Given what we know about the world, it is eminently reasonable to assume that there are gapless chains of events (states or processes) in the brain and central nervous system that causally underlie bodily movements. It also seems reasonable to suppose, as Burge (1993) puts it, that "[c]ausal implications are built into our intentional concepts and intentional modes of explanation" (p. 118). That this assumption is reasonable, I would argue, is revealed through judicious and philosophically unbiased reflection on our explanatory practices. "Much of the interest of psychological explanation," Burge contends in that same paper, " . . . lies in helping us understand ourselves as agents . . . . Our agency consists in our wants, willings, thoughts, values as such (under these 'aspects') having some sort of efficacy in the world" (pp. 118–119). Thus human behavior would appear to admit of two types of causal explanation: one neurological and the other intentional. How then are these two accounts of behavior related? This is the central question of this paper.

One well-known attempt to answer this question is the token-identity theory. Here, the two explanations pick out the same etiologies: proposi-

Requests for reprints should be sent to Jeffrey Hershfield, Ph.D., Department of Philosophy, Wichita State University, Wichita, Kansas 67260–0074.

tional attitudes are held to be token identical with neural states (events or processes), and actions are to be token identical with bodily movements.[1] Not only does such a view manage to allay Ockhamian worries about causal overdetermination, but it also claims to "demystify" mental causation. Mental causation turns out not to be some special sort of causation over and above whatever types of physical causal mechanisms are to be found in the world. Rather, mental causation is shown to be a form of physical causation. Defenders of the token-identity theory are quick to point out that mentalistic explanations are nonetheless not replaceable by ones couched entirely in a neurological vocabulary, at least not without loss of explanatory power. The central idea is that the mentalistic vocabulary is couched at a level of generality that cannot be captured in a nonmentalistic, nonintentional idiom. Thus, when we explain behavior in terms of an agent's propositional attitudes, we show the behavior to fit into a pattern of regularities, or to be subsumed by generalizations, that would otherwise be lost to view were we to switch to a nonintentional idiom.[2]

We can illustrate how this is all supposed to work with the following example. While play is going on in a football game, the referee tosses a yellow flag onto the field. The yellow flag is a penalty flag, and by tossing the flag onto the field the referee is calling a penalty. The mentalistic explanation of why the referee threw the flag — the explanation of the action — is relatively straightforward: he had certain beliefs — that an infraction had occurred, that he had an obligation to penalize the team that committed the foul, that tossing the yellow flag was the appropriate way to call a penalty — and a desire to fulfill his professional obligations (and no stronger competing desires). Of course, in throwing the penalty flag the referee's grip exerted a certain amount of pressure on the flag, the flag was positioned just so in his hand, his arm moved in a certain way, the flag was released from his hand at a particular point in the throw, and so on. Now it is assumed by the token-identity theory that there is some story, presumably quite complex, about why the referee's throwing motion took just this form — a story to be told in terms of electro-chemical activity in the brain and nerves running along the referee's arm. We may not at present be able to tell the story in full detail,

[1]The reason for the parenthetical additions is that it is a matter of controversy as to what sorts of "tokens" participate in causal transactions. My account is neutral as between events, states, or processes. For brevity's sake, henceforth I shall use "states," although nothing crucial turns on this choice.

[2]For various ways in which this idea can be spelled out, see Davidson, 1970, 1973, 1974, 1993; Fodor, 1974, 1980, 1987, 1989, 1991; and Schiffer, 1987, 1990. Although he disavows the token-identity theory, Dennett has long advocated the idea that intentional explanations capture a set of patterns or regularities that can *only* be described by adopting that explanatory stance — what he calls the "intentional stance." See Dennett, 1971, 1981a, 1981b, 1981c, 1987a, 1987b, 1987c, and 1991.

but that is thought to be a function of our ignorance, and not the way the world is (or is not).

The sense in which mentalistic explanations advert to different regularities or patterns from those adverted to by neurophysiological ones turns on the idea that their respective causal-explanatory constructs *cross-classify* one another. Consider the obvious fact that the referee did not have to move his arm just so in order to have performed the act of throwing a penalty flag (thereby calling a penalty). He could have thrown the flag any number of different ways and still succeeded in calling a penalty. We can describe this by saying that the mentalistic explanation displays a certain degree of *stability* in counterfactual circumstances. Below I shall argue that this feature of stability reveals the mentalistic explanation of throwing the penalty flag to have a different explanandum from the physicalistic explanation of the relevant bodily movement: *pace* the token-identity theorist, the relationship between the action and the bodily movement that realizes it is not one of token identity. But if the two explanations don't share explananda, then it is unlikely that they share explanantia: the brain events/states that causally underlie bodily movements are *not* token identical with the mental events/states that causally underlie actions. In what follows I shall situate these ideas within the context of a general theory of explanation, one which will reveal the nature of "structural" causal explanations.

Explanatory Relativity and Higher-Order Causation

Garfinkel (1981) argues that explanations have complex underlying structures. Differences between explanations in their underlying structure are often masked by similarities in their "surface structure." Garfinkel suggests that there is an important philosophical lesson to be learned from the famous story of Willie Sutton the bankrobber: when asked why he robs banks, Sutton replied with the quip, "That's where the money is." What makes Sutton's cheeky response clever is that he has managed to avoid answering the question he was asked by answering another question that superficially resembles the original. He has managed to explain why he robs banks without satisfying his interlocuor's request to explain why he robs banks. This is possible because there is no univocal explanandum that can be characterized simply as "Sutton's robbing of banks." Sutton explains why he robs banks as opposed to other possible targets, such as movie theaters or elementary schools. His interlocutor wants to know why he robs banks as opposed to not robbing anything at all. Garfinkel argues that this example nicely illustrates a pervasive feature of explanations which he refers to as *explanatory relativity*. The idea is that explanations invoke, either implicitly or explicitly, a contrast space or space of alternatives. Sutton contrasts his robbing of banks

with the robbing of other establishments. His interlocutor contrasts his robbing of banks with not robbing anything at all. The contrast space or space of alternatives serves to *individuate* the explanandum. The contrast space delineates those differences that "matter" for the purposes of explanation; in effect, it indicates the variations that would constitute a change in the explanandum. On the other hand, variations that don't show up in the contrast space are inessential for the purposes of explanation; they are differences that make no difference. "These irrelevant perturbations," Garfinkel argues, "determine an equivalence relation, 'differs inessentially from,' and the real object of explanation is an equivalence class under this relation" (1981, pp. 30–31). Sutton's interlocutor is seeking an explanation of why he chose robbery as a way of supporting himself. As it turns out, Sutton chose to rob banks; however, for these explanatory purposes the difference between robbing banks and robbing other venues is negligible. In Garfinkel's terms, these would be "irrelevant perturbations" that "differ inessentially." This is why Sutton's response to his interlocutor is inapposite. It treats the difference of venues as salient and in need of explanation. Garfinkel's analysis of the situation is that Sutton has gotten hold of the wrong explanandum (intentionally so, of course). Sutton's proffered explanation is aimed at *too specific* an explanandum. The explanandum he ought to have targeted has somewhat more abstract criteria of individuation which is more or less captured by the expression "Sutton's propensity to rob."

Let's consider another of Garfinkel's examples. He imagines a hypothetical ecosystem consisting of foxes and rabbits whose populations fluctuate relative to one another according to the following pattern: when the rabbit population is high relative to the fox population there is little competition amongst the foxes for food. As a result, the foxes flourish, their numbers increase, and this puts pressure on the rabbit population (the foxes' primary food source). The rabbit population begins to decrease and eventually there come to be too many foxes and too few rabbits. The population of rabbits can no longer support the population of foxes. Competition for food amongst the foxes becomes so intense that the fox population begins to decrease. As the number of predators decreases, the rabbit population makes a comeback until eventually the rabbits come again to greatly outnumber the foxes and the pattern repeats itself.

Now imagine that we wish to explain the death of a rabbit R who, as it turns out, is killed by a fox F. As before, surface grammatical structure belies a complex underlying logical structure. On the one hand, we might be interested in explaining the specifics of R's death: why at the hands of F? (We could of course get more specific still.) Such an explanandum would be situated in a contrast space of possibilities in which R is killed by different foxes: the explanandum would be death at the hands of F (rather than F', F", . . .).

On the other hand, we might be interested in explaining why R was killed by a fox rather than not killed at all. Here R's death is being contrasted with possible scenarios in which R does not meet its end in the jaws of a fox. That these constitute distinct explananda requiring distinct explanantia becomes evident when we consider what would have had to be different for the explananda not to have occurred, *ceteris paribus* (or to have been highly unlikely to occur, *ceteris paribus*).[3] Very different factors were causally necessary in each case. In the first explanation, such factors as the time and place of R's departure, and the route taken, are essential to the occurrence of the explanandum: had R taken a different route or started out half an hour earlier or from a point half a kilometer away, then it would not have encountered a hungry F, *ceteris paribus* (or at least this encounter would have been unlikely, *ceteris paribus*). These factors thus figure in the explanans of the first explanation insofar as it aims to identify factors that were causally necessary for the occurrence of the explanandum.

Because there were so many more foxes than rabbits, the details of R's time of departure, route taken, etc., could have differed from what they actually were without significantly decreasing the likelihood that R would end up being killed by a fox. When it comes to explaining why R was killed by a fox (rather than not so-killed), it is the high fox to rabbit ratio — *a structural* feature of the ecosystem — that appears to be the relevant (necessary) factor. Had that structural feature failed to obtain — had there been more rabbits and fewer foxes — R's chances of survival would have greatly improved. In the second explanation what is being explained is R's having a deadly encounter with a fox. This means more than just explaining R's actual encounter with F; it means explaining why such an encounter, if not with F then with some other member of the local fox population, was inevitable (or highly probable). In Garfinkel's terminology, what we seek to explain in the second case is a certain *redundant causality* in the ecosystem.

> Systems which exhibit redundant causality therefore have, for every consequent Q, a bundle of antecedents (P_i) such that
>
> 1. If any one of the P_i is true, so will be Q.
>
> 2. If one P_i should not be the case, some other will. (1981, p. 58)

In the example under discussion, Q is R's being killed by a fox, and the P_i are various possible scenarios involving the initial positions, routes taken, rates of motion, etc., of R and assorted foxes (including, of course, the events that led to R's deadly encounter with F). Although R could not have been killed by a fox without some P_i having occurred — without there having been a set of

[3] I don't want to prejudge the question of whether or not causal relations are deterministic.

circumstances leading to a deadly encounter with a particular fox at a given time and place — it does not follow that the structural property (the high fox-to-rabbit ratio) was causally irrelevant. For unlike any of the individual P_i, it is the structural property whose absence would have insured (or significantly raised the probability) that R was not killed by a fox.

The same points apply to explanations of action. Recall our earlier football example. The explanation of the referee's act of throwing the penalty flag cites the obvious panoply of beliefs and desires. This explanation, we noted, is couched at a level of generality that abstracts from the precise details of the bodily movement M that realizes the action. We put the point by saying that the action explanation exhibits a degree of "stability" across counterfactual situations. The manner in which the referee tossed the flag could have differed from M in any number of ways while still counting as an act of calling a penalty. Thus the bodily movement exhibited by the referee could have differed along several different dimensions without materially affecting the explanation of the action. In Garfinkel's terminology, these various arm movements M', M'', . . . M^n, differ inessentially from M; they are members of the equivalence class that constitutes the explanandum. This means that an explanation of the referee's throwing the penalty flag is an explanation of a redundant causality in behavior. It is an explanation of why, under a variety of perturbations of the initial conditions, an agent is bound to exhibit some form of behavior that belongs to the relevant class of inessential differences; in other words, some form of behavior that realizes the action in question. As we saw in the foxes and rabbits example, it is "higher-order" features of a system — in particular, structural features — that are capable of ensuring that under a variety of perturbations of the initial conditions, some member of the class of inessential differences is bound to occur. Furthermore, if we want to know what would have had to have been different in order for the relevant redundant causalities not to have occurred, we must look to these same structural features. Now, in explaining why an agent acted as she did by citing her reasons, we take ourselves to be identifying states or events of the agent that causally led to the action in question. In many such cases reasons are *also* viewed as causally necessary for the action: had the agent not believed so and so, or desired such and such, then she would not have done A, *ceteris paribus*.[4] We thus have the makings of an account of mental causation: propositional attitudes are capable of causing actions because they are (token) identical with realizations of the structural properties in an agent's brain that causally explain the relevant redundant causalities in the agent's behavior. My account thus proposes to assimilate mental causation to a form

[4]The reason for the circumspection is just that agents sometimes have more than one reason for acting in a given case.

of causation — structural causation — that also operates outside of the mental sphere. However, it takes seriously the idea that action explanations are couched at a greater degree of generality than explanations of the bodily movements that realize actions. Hence it avoids the mistake made by the token-identity theory of assuming that the two kinds of explanation aim to explain the very same things under different descriptions. In what follows I will discuss two important objections that might be raised against the account.

Whither Structural Causes?

I have suggested that propositional attitudes are (token) identical with realizations of structural properties of brains that causally explain redundant causalities in behavior. Mental causation is to be understood as a type of structural causation. Quite obviously, the account stands or falls on the plausibility of the claim that structural properties can be causally efficacious. An important feature of structural properties is that they are *multiply realizable*. For example, the high fox-to-rabbit ratio that explained the near certainty of R's being killed by a fox was realizable by many different distributions of foxes and rabbits.[5] Now, Kim (1993) argues against the tenability of multiply realizable causal-explanatory properties, focusing his attack on functionalist accounts which treat mental properties as *higher-order* properties: they are properties of properties having a given causal role. Glossed in this way, mental properties turn out to be multiply realizable. In different subjects, one and the same causal role might be played by different lower-order properties. It is not unheard of for functionalists to indulge in speculation about the possibility that each mental property might be realized by a collection of wildly heterogeneous lower-order properties. They point out, for example, that pain sensations may have very different physical realizations in octopi, rats, human beings, silicon-based computers, and non-carbon-based aliens from outer space. If that is indeed the case, then, notwithstanding these differences in implementation, there is a level — the functional level — at which these creatures can be said to share salient properties and to be subject to the same lawlike regularities in behavior. This is what functionalists mean when they assert that the multiple-realizability of mental properties reveals psychology to be an autonomous science with its own irreducible laws.[6]

---

[5]This is to overstate the case slightly. There will be some distributions of foxes and rabbits that realize the relevant ratio but do not result in the same redundant causality. Such would be the situation, for example, were the rabbits to be all bunched together at one end of the ecosystem, the foxes at the other, and were there enough obstructions to keep the populations largely apart. This shows that the causal powers of structural properties are not entirely independent of those of the "substratum elements." I discuss this point in more detail in the text.

[6]See, for example, Putnam, 1967.

Now, my account differs from the sort of functionalist theory that Kim targets in two important respects. First, the kind of multiple-realizability that I am attributing to structural properties is relatively modest. My account of mental causation confines itself to a single substratum, the brain, and proposes a degree of latitude in the ways in which certain structural properties of that substratum can be realized. Second, structural properties are *not* functional properties. Functional properties are properties at level $_{m}$ of properties at level$_{m-1}$ having such and such causal roles. Thus, importantly, functional properties are *defined* in terms of the causal profiles of properties at lower levels. For that reason functional properties fail to be causal explanatory (Block, 1990; Searle, 1992). My account is immune to this criticism, however, since structural properties are identifiable *independently* of their causal roles. For example, in the ecosystem of foxes and rabbits, the causally efficacious structural property is identifiable as a given ratio of foxes to rabbits. Nevertheless, Kim's argument purports to undermine the causal efficacy of all multiple-realizable properties. Hence my account falls within the purview of his argument.

Kim (1993) cites the example of jade. It turns out that jade is not a mineral kind; it actually comprises two distinct mineral kinds, jadeite and nephrite. Though they overlap in the observable properties that we associate with jade — the properties that underlie our application of the predicate "jade" — jadeite and nephrite differ in their molecular structures in ways relevant to our classificatory scheme for minerals. Hence the property of being jade is multiply realizable: it can be realized or instantiated in heterogeneous substrata. Because of this, Kim argues, the predicate "jade" is ill-suited to the formulation of laws. Generalizations containing "jade," even if true, lack such essential features of laws as *projectibility:* their degree of confirmation is not proportional to the number of observed positive instances. There is not the sort of "accretion" of confirmation as a function of the observation of increasing numbers of positive instances that we expect to find with genuine laws. Kim offers the following candidate generalization by way of illustration.

(L) Jade is green

He points out that (L) is true only if (L_1) and (L_2) are both true.

(L_1) Jadeite is green
(L_2) Nephrite is green

His argument continues as follows.

> [W]e can imagine this: on re-examining the records of past observations, we find, to our dismay, that all the positive instances of (L), that is, all the millions of observed samples of green jade, turn out to have been samples of jadeite, and none of nephrite!

> If this should happen, we clearly would not, and should not, continue to think of (L) as well confirmed. All we have is evidence strongly confirming (L_1), and none having anything to do with (L_2) . . . . But all the millions of green jadeite samples *are* positive instances of (L): they satisfy both the antecedent and the consequent of (L). (p. 320; emphasis in original)

Exactly the same considerations would appear to apply, *mutatis mutandis*, to mental properties insofar as they are multiply realizable. Generalizations statable in terms of such properties, even if true, will be conjunctions of heterogeneous laws that subsume distinct microstructures. Hence such generalizations will fail to be projectible, and thus fail to qualify as laws. But the fact that mental properties fall short of being nomic kinds means that they fail to qualify as *causal* kinds (on the plausible assumption that the capacity to appear in laws is a necessary condition of being a causally efficacious property).

It is important to remember that Kim has in mind functionalist theories that allow for the possibility of wildly heterogeneous realizations of mental properties. It is precisely this feature of functional properties upon which the analogy to the property of being jade is based. However, the situation is very different when we compare the propery of being jade with the kinds of structural properties that figure in my account. What is supposed to make the property of being jade unsuitable for laws (*mutatis mutandis* for functional properties) is the fact that it is realized in distinct microstructures, each having its own causal powers. The reason why jadeite is green may have little to do with the reason why nephrite is green. Thus any generalization statable in terms of the predicate "jade," even if true, fails to be projectible, hence fails to be lawlike. But when we turn to the structural properties posited in my account we find the situation to be very different. My account assumes the multiple realizability of structural properties within a single substratum, the brain. When we imagine different possible realizations of a given structural property, we are not imagining heterogeneous microstructures with differing causal powers. Rather, we are imagining different possible ways in which a single neural substratum might realize a structural property (for example, via different patterns of neural interconnections). There is thus not the compelling reason, as there is in Kim's jade case, for thinking that generalizations stated in terms of structural properties of the brain would fail to be projectible, hence fail to be lawlike.

Nonetheless, some concessions must be made to Kim. An important assumption underlying his argument is the notion that the direction of confirmation percolates upward from the microlevel to the macrolevel. Facts at the microlevel (or substratum level) are held to *explain* facts at the macrolevel. This is why (L)'s truth is supposed to be explained by the joint truth of (L_1) and (L_2). Kim (1993) claims that this is rooted in the metaphys-

ical thesis that "macrophysical properties of substances are determined by microstructure," a thesis that serves to "regulate our inductive and explanatory practices" (p. 322). He argues that no version of materialism worthy of the name could deny this thesis.

Insofar as this thesis denies the possibility of structural causes, it ought to be rejected. However, the principle is correct at least to the extent that the causal efficacy of structural causes is not entirely independent of the causal potentials of the substratum elements in which they are realized. To illustrate the point, let's return to Garfinkel's example of the ecosystem of foxes and rabbits. I argued that it is a structural property of the ecosystem that is causally responsible for R's death at the hands of a fox (this is tantamount to the explanation of a certain redundant causality in the ecosystem). The structural property is a certain ratio of foxes to rabbits. It is crucial to the causal efficacy of this structural property that its realization involves members of the *fox* population. The same ratio of blades of grass to rabbits, or mice to rabbits, would not have resulted in the redundant causality in question. Thus the higher-level explanation cannot be so easily severed from the lower-level properties; in this case, from the hunting skills of individual foxes.

Similar points apply to my account of action explanations. Even though I am assuming that the relevant structural properties can be multiply realized within the brain, it can't be the mere presence of these structural features that accounts for redundant causalities in behavior. The nature of the substratum in which these structural properties are realized is also crucial; it is essential that they are realized in neurons and not hair follicles. This is because neurons and not hair follicles have properties which are such that, when the relevant structural properties are realized — for example, by various arrays of neural interconnections — the results are redundant causalities in behavior. Thus intentional psychology cannot develop in isolation from neurophysiology, neurobiology, and so on.

Garfinkel acknowledges the importance of the substratum for higher-level structural explanations. He suggests that such explanations typically make presuppositions that are not themselves components of the explanations. These presuppositions play an important role in delineating both the explananda and explanantia of such explanations. He offers several examples to illustrate this idea, such as the following taken from biochemistry.

> If the explanation for why a leaf is green is that it contains chlorophyll, then the causal form:
>
> contains chlorophyll causes green color
>
> holds generally in a wide variety of circumstances. But not in all circumstances: *jars* which contain chlorophyll are not caused thereby to be green. So obviously, there is a presupposition working here, to the effect that this case is the *kind* of

> case for which "contains chlorophyll" explains "greenness." (1981, p. 33; emphasis in original)

As a further illustration, Garfinkel points out that if one wanted to explain why something moved as it did by citing Newton's laws of falling bodies, then one is presupposing that the explanandum is a physical object. (If it were a shadow, then such an explanation would not be appropriate.) These same points can be applied to the explanation of redundant causalities in behavior: the fact that the structural explanation only works if the underlying substratum has certain causal possibilities is then explained as part of the presupposition of the explanation. Given that we are dealing with organisms capable of producing a wide range of bodily movements under a variety of circumstances, what are the features of the mechanism(s) responsible for these bodily movements that account for the existence of redundant causalities? If my account is correct, then the answer will advert to structural properties of the neural mechanisms that causally underlie bodily movements. Recognition of the role of the substratum in higher-level structural explanations does not threaten to undermine the coherence or significance of the distinction that Garfinkel draws between such explanations and explanations couched at lower levels.

Having argued for the importance of the substratum to the success of explanations that advert to structural properties, I could be thought guilty of having thrown the baby out with the bath water.[7] Crucial to my account is the assumption that such explanations display greater degrees of stability across counterfactual situations than is displayed by explanations of lower-order phenomena. That increased stability is reflective of the robust nature of higher-order explananda. In other words, the greater degree of stability is indicative of the fact that the phenomena which count as inessentially different, hence included in the equivalence class in terms of which a given higher-order explanandum is individuated, are otherwise contrasted with one another in lower-order explanations. Nevertheless, my account of structural causation ties the causal efficacy of structural properties to the causal efficacy of the substratum elements; specifically, a structural property causally explains a higher-order phenomenon only if appropriate causal relations obtain between relevant lower-order phenomena, both actual and counterfactual. Thus recall, one last time, Garfinkel's example of the foxes and rabbits. In arguing for the hautonomy of the higher-order explanation of R's death, I cited both the fact that this explanation prescinds from the details of R's actual fatal encounter with F, and that a structural property of the ecosystem — viz., the high fox-to-rabbit ratio — causally explains the higher-order

[7]As was suggested by an anonymous referee.

explanandum. In justifying the latter claim I contended that it is the structural property which ensures that, in both the actual world, and in a wide range of counterfactual circumstances, a fatal encounter between R and a fox was inevitable or highly probable. *Ex hypothesi*, each such deadly encounter is amenable to a lower-order explanation. But then why not dispense with the higher-order explanation altogether, thereby minimizing one's ontological commitments (in deference to Ockham's Razor)? If each member of the class of phenomena covered by a higher-order explanation admits of a lower-order explanation, then is not the higher-order explanation simply otiose?

Such reasoning derives its specious appeal from a failure to grasp one of Garfinkel's principal points, which is that lower-order explanations are couched at the wrong level of abstraction to explain redundant causalities; they operate at the wrong level of specificity. To be compelling this argument would have to be supplemented by a rejection of redundant causalities as legitimate explananda. But the prospects for such a maneuver appear dim. For if Garfinkel is right, then both the physical and social sciences are rife with examples of explanations of redundant causalities.

References

Block, N. (1990). Can the mind change the world? In G. Boolos (Ed.), *Meaning and method: Essays in honor of Hilary Putnam* (pp. 137–170). Cambridge, England: Cambridge University Press.

Burge, T. (1993). Mind–body causation and explanatory practice. In J. Heil and A. Mele (Eds.), *Mental causation* (pp. 97–120). Oxford: Clarendon Press.

Davidson, D. (1970). Mental events. In L. Foster and J. Swanson (Eds.), *Experience and theory* (pp. 79–101). Amherst: University of Massachusetts Press.

Davidson, D. (1973). The material mind. In P. Suppes, L. Henkin, G. Moisil, and A. Joja, (Eds.), *Proceedings of the Fourth International Congress for Logic, Methodology, and Philosophy of Science IV* (pp. 709–722). Amsterdam: North Holland.

Davidson, D. (1974). Psychology as philosophy. In S. Brown (Ed.), *Philosophy of psychology* (pp. 41–52). New York: Macmillan.

Davidson, D. (1986). Knowing one's own mind. *Proceedings of the American Philosophical Association, 60*, 441–58.

Davidson, D. (1993). Thinking causes. In J. Heil and A. Mele (Eds.), *Mental causation* (pp. 3–17). Oxford: Clarendon Press.

Dennett, D. (1971). Intentional systems. *Journal of Philosophy, 68*, 87–106.

Dennett, D. (1981a). True believers: The intentional strategy and why it works. In A.F. Heath (Ed.), *Scientific explanation* (pp. 53–75). Oxford: Oxford University Press.

Dennett, D. (1981b). Three kinds of intentional psychology. In R. Healy (Ed.), *Reduction, time, and reality: Studies in the philosophy of the natural sciences* (pp. 37–60). Cambridge, England: Cambridge University Press.

Dennett, D. (1981c). Making sense of ourselves. *Philosophical Topics, 12*, 63–81.

Dennett, D. (1987a). Reflections: Real patterns, deeper facts, and empty questions. In D. Dennett (Ed.), *The intentional stance* (pp. 37–42). Cambridge, Massachusetts: Bradford Books/MIT Press.

Dennett, D. (1987b). Reflections: Instrumentalism reconsidered. In D. Dennett (Ed.), *The intentional stance* (pp. 69–81). Cambridge, Massachusetts: Bradford Books/MIT Press.

Dennett, D. (1987c). Reflections: When frogs (and others) make mistakes. In D. Dennett (Ed.), *The intentional stance* (pp. 103–116). Cambridge, Massachusetts: Bradford Books/MIT Press.
Dennett, D. (1991). Real patterns. *Journal of Philosophy, 88*, 27–51.
Fodor , J. (1974). Special sciences. *Synthese, 21*, 97–115.
Fodor, J. (1980). Methodological solipsism considered as a research strategy in cognitive psychology. *Behavioral and Brain Sciences, 3*, 63–110.
Fodor, J. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind*. Cambridge, Massachusetts: Bradford Books/MIT Press.
Fodor, J. (1989). Making mind matter more. *Philosophical Topics, 17*, 59–79.
Fodor, J. (1991). Reply to Stich. In B. Loewer and G. Rey (Eds.), *Meaning in mind: Fodor and his critics* (pp. 310–312). Oxford: Basil Blackwell.
Garfinkel, A. (1981). *Forms of explanation*. New Haven: Yale University Press.
Kim, J. (1993). *Supervenience and mind: Selected philosophical essays*. Cambridge, England: Cambridge University Press.
Putnam, H. (1967). Psychological predicates. In W.H. Capitan and D.D. Merrill (Eds.), *Art, mind, and religion* (pp. 37–48). Pittsburgh: University of Pittsburgh Press.
Schiffer, S. (1987). *Remnants of meaning*. Cambridge, Massachusetts: Bradford Books/MIT Press.
Schiffer, S. (1990). Physicalism. In J. Tomberlin (Ed.), *Philosophical perspectives, 4: Action theory and philosophy of mind 1990* (pp. 153–185). Atascadero, California: Ridgeview Publishing Company.
Searle, J. (1992). *The rediscovery of the mind*. Cambridge, Massachusetts: Bradford Books/MIT Press.