AN IDLE THREAT:
EPIPHENOMENALISM EXPOSED

by

Paul D. Raymont

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy,
Graduate Department of Philosophy,
University of Toronto

Paul David Raymont
Department of Philosophy
University of Toronto
Ph.D. 1999

Abstract
An Idle Threat: Epiphenomenalism Exposed

This thesis involves a consideration, and rejection, of the claim that recent varieties of non-reductive physicalism, particularly Donald Davidson's anomalous monism, are committed to a new kind of epiphenomenalism. Non-reductive physicalists identify each mental event with a physical event, and are thus entitled to the belief that mental events are causes, since the physical events with which they are held to be identical are causes. However, Jaegwon Kim, Ernest Sosa and others have argued that if we follow the non-reductive physicalist in denying that mental features can be reduced to physical properties, then we must regard mental properties as being causally irrelevant to their bearers' effects. In short, the non-reductive physicalist is said to be committed to the belief that while there are mental causes, they do not cause their effects in virtue of being the types of mental state that they are. It is in this sense that non-reductive physicalists are thought to represent a new form of epiphenomenalism. After a brief survey of the history of epiphenomenalism, and its mutation into the contemporary strain that is believed to afflict non-reductive physicalism, I argue against the counterfactual criterion of the sort of causal relevance that we take mental features to enjoy. I then criticize the "trope" response to the epiphenomenalist threat, and conclude that much of the current debate on this topic is premissed on the mistaken belief that there is some variety of causal relevance that is not simply a brand of explanatory relevance. Once this is seen, it will seem much less plausible that mental properties are excluded from relevance to the phenomena of which we typically take them to be explanatory.

to my parents

Robert and Nancy Raymont

and my wife

Victoria Burke

Acknowledgements

I would like to thank my supervisor, William Seager, and my advisers, Sonia Sedivy and Ronald de Sousa, for their advice and incisive comments on earlier drafts of this material. I would also like to thank Calvin Normore for his stimulating suggestions during the early stages of this project, and Peter Hess, my first philosophy teacher. I owe a tremendous debt of gratitude (and money) to my parents, Robert and Nancy Raymont. Finally, I would like to acknowledge the support and encouragement I have received from my wife, Victoria Burke, both emotionally and intellectually.

Table of Contents

1. Introduction

> We laugh at him who steps out of his room at the moment when the sun steps out of its room, and then says: "*I will* that the sun shall rise"; and at him who cannot stop a wheel, and says: "*I will* that it shall roll"; and at him who is thrown down in wrestling, and says: "here I lie, but *I will* lie here!" But, all laughter aside, are we ourselves ever acting any differently whenever we employ the expression: "*I will*"?
>
> Nietzsche[1]

> Our mental conditions are simply the symbols in consciousness of the changes which take place automatically in the organism; . . . to take an extreme illustration, the feeling we call volition is not the cause of a voluntary act, but the symbol of that state of the brain which is the immediate cause of that act. We are conscious automata.
>
> T. H. Huxley[2]

Credit for the doctrine of epiphenomenalism must go to Shadworth Hodgson (or perhaps to the organism on which he supervened). Hodgson's presentation of the view antedates Huxley's by four years, having first appeared in Hodgson's *The Theory of Practice* in 1870.[3] The view, stripped to its essentials, is that mental phenomena are caused by physical events, but in turn cause nothing.

Epiphenomenalism obviously concerns many of the issues that have traditionally arisen in discussions of `the will'.

---

[1] Friedrich Nietzsche, *Daybreak*, trans. R.J. Hollingdale, ed. Maudemarie Clark and Brian Leiter (Cambridge: Cambridge University Press, 1997) bk. II, sec. 124.

[2] T.H. Huxley, "On the Hypothesis That Animals Are Automata," in *The Philosophy of Mind*, ed. Brian Beakley and Peter Ludlow (Cambridge, MA: the MIT Press, 1992), 133-6 (at 136); originally published in 1874.

[3] William James identifies Hodgson as the first proponent of epiphenomenalism in William James, *The Principles of Psychology* (New York: Henry Holt and Company, 1890), 1:130. I have been unable to locate a copy of Hodgson's book.

Indeed, in the course of discussing quite different claims about the will, some philosophers have given clear statements of epiphenomenalism without intending to do so. This is probably true of the above epigraph from Nietzsche, and is certainly true of the following quotation from Wittgenstein:

> You sometimes see in a wind a piece of paper blowing about anyhow. Suppose the piece of paper could make the decision: 'Now I want to go this way.' I say: 'Queer, this paper always decides where it is to go, and all the time it is the wind that blows it. I know it is the wind that blows it.' That same force which moves it also in a different way moves its decisions.[4]

Here, Wittgenstein is giving expression to the view that the will is not free, and yet the image he proffers nicely captures one of the central tenets of epiphenomenalism, according to which the same physical forces that engender the bodily motions constitutive of behaviour also "in a different way" produce the decisions that appear (misleadingly) to be the causes of that behaviour. This model enables the epiphenomenalist to account for the manifest regularities that obtain between our decisions and their corresponding behavioural expressions *without* being compelled to say that the former events are causes of the

---

[4] Ludwig Wittgenstein, "Lectures on Freedom of the Will," from notes taken by Yorick Smythies, in *Ludwig Wittgenstein: Philosophical Occasions 1912-1951*, ed. James C. Klagge and Alfred Nordmann (Indianapolis: Hackett Publishing Co., 1993), 434.

latter ones. These regularities can be explained simply by saying that the decisions and their behavioural expressions are joint effects of a common (physical) cause; as Huxley says, the volitions constitute "a collateral product" of the body's workings.[5]

Another image that nicely captures this strategy for making sense of psychophysical regularities without according efficacy to the mental was given by Hugh Elliot in the following story:

> Suppose that Tantalus, his hammer, and his anvil were concealed . . . by a screen . . . and that a light . . . threw the shadow of the hammer and anvil upon a wall. . . . Every time the shadow of the hammer descended upon the shadow of the anvil, the sound of the percussion is heard. . . . What is the inevitable effect upon the observer's mind? . . . He cannot escape the conclusion that the cause of each sound is the blow which the shadow of the hammer strikes upon the shadow of the anvil. . . . States of consciousness are shadows of cerebral functioning; . . . the cause of action lies in the cerebral functioning and not in the shadows which accompany it.[6]

---

[5] T.H. Huxley, "On the Hypothesis That Animals Are Automata," 135.

[6] Hugh S. R. Elliot, *Modern Science and the Illusions of Professor Bergson* (London: Longmans, Green, and Co., 1912), 185-7 (quoted from Paul Edwards, ed., *Immortality* [Amherst, NY: Prometheus Books, 1997], 184-5). Elliot gives a particularly strident, scientistic defense of epiphenomenalism. It should be noted that the doctrine did not win the support of all late Victorian, scientistic authors. Herbert Spencer felt compelled to augment the sixth edition of his *First Principles* with a denunciation of Huxley's views (Herbert Spencer, *First Principles*, 6th ed. [New York: D. Appleton and Company, 1900], sec. 71b, 198-202). He there offers the interesting observation that the epiphenomenalist owes us an explanation of why it is that the volitions that accompany an action when it is first performed cease to occur after the action has been repeated many times and become habitual.

Most writers (including Wittgenstein) rightly present this view as an example of determinism. Huxley himself took epiphenomenalism to be consistent with the freedom of the will, on the grounds that it does not preclude one's acting in accordance with one's desires.[7] Clearly, though, freedom requires not just that one's actions accord with one's desires, but also that they be products of those desires; however, in the world depicted by epiphenomenalism, while an agent can act as she wants, she cannot act *because* of what she wants.[8]

It is this lack of efficacy that has led many to regard epiphenomenalism as a particularly pointless brand of dualism. At least the interactionist takes there to be something the mind does, something that cannot be fully understood without appealing to a mind. By contrast, epiphenomenalists ask us to affirm the existence of non-physical mental states while in the same breath maintaining that all of the effects commonly imputed to those states are in fact produced by the brain alone. But, one may ask, if a physical thing does everything that we thought was done by the mental, why not simply identify it with the mental? As the neuropsychologist, D.O. Hebb, said in his critique of epiphenomenalism, "A brain that functions in every respect

---

[7] T.H. Huxley, "On the Hypothesis That Animals Are Automata," 135.
[8] Determinism does not require epiphenomenalism, since our volitions could be real causes while yet being themselves determined.

like a mind *is* a mind."[9] In short, epiphenomenalism appears to be merely an anachronism in the evolution from interactionist dualism to physicalism, arising from the failure to see that once the causal powers associated with the mind have been appropriated by physical entities, there remains no reason for continuing to believe in distinct, non-physical mental events. Epiphenomenalism disappears as a relevant concern if we simply become physicalists.

It is not clear, however, that epiphenomenalism really does require a metaphysical setting in which the identity of mental and physical events is denied. The world might conform to the spirit, if not the letter, of Huxley's outlook even if all events are physical, for it might be the case that while mental events are physical events, their being mental in no way contributes to their having the causal powers that they have. This manifestation of epiphenomenalism within the context of a physicalist metaphysics was limned by C.D. Broad in 1925 when he said,

> Epiphenomenalism may be taken to assert . . . that certain events which have physiological characteristics have *also* mental characteristics, . . . and that an event which has mental characteristics never causes another event in virtue of its mental characteristics, but only in

[9] D.O. Hebb, *Essay on Mind* (Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers, 1980), 39. Hebb's own argument for this claim is pragmatist. He says, "To paraphrase C.S. Peirce's principle of pragmatism: Consider what practical effects such a distinction of mind from brain may have; then our conception of those effects is the whole conception of the distinction. The practical effects are null, the distinction is merely verbal" (D.O. Hebb, *Essay on Mind*, 40).

virtue of its physiological characteristics.[10]

Contrary to popular belief, Broad was not the only, or even the first, early twentieth-century philosopher to consider this new brand of epiphenomenalism. He seems to be regarded as such by Brian McLaughlin.[11] In fact, however, George Santayana, as early as 1906, had endorsed a position that bears remarkable similarities to the kind of epiphenomenalism that is described by Broad.[12] Santayana's manner of expression, being often poetic (and sometimes bombastic), is not nearly as clear as Broad's. Nevertheless, it appears that Santayana wished to deny efficacy to thought "in its ideal capacity"[13] or when it is "taken as a psychological existence",[14] while at the same time allowing thought to have real causal force "through the natural

---

[10] C.D. Broad, *The Mind and Its Place in Nature* (London: Routledge & Kegan Paul Ltd., 1925), 472.

[11] Brian P. McLaughlin, "Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical," *Philosophical Perspectives* 3 (1989): 109-35 (at 109). See also John Heil, *The Nature of True Minds* (Cambridge: Cambridge University Press, 1992), 121; and Stephen Yablo, "Mental Causation," *The Philosophical Review* 101 (1992): 245-80 (at 248 n. 8). McLaughlin implies that Broad himself was vexed by this brand of epiphenomenalism and struggled vainly to overcome it (Brian P. McLaughlin, "On Davidson's Response to the Charge of Epiphenomenalism," in *Mental Causation*, ed. John Heil and Alfred Mele [Oxford: Clarendon Press, 1993], 27-40 [at 28 n. 3]). However, as William C. Kneale points out after quoting this same passage from Broad (William C. Kneale, "Broad on Mental Events and Epiphenomenalism," in *The Philosophy of C.D. Broad*, ed. Paul Arthur Schilpp [New York: Tudor Publishing Company, 1959], 437-55 [at 442]), although Broad gave voice to this new kind of epiphenomenalism, he in fact happily endorsed an old-fashioned, dualistic epiphenomenalism. This interpretation is confirmed by Broad in his reply to Kneale in *The Philosophy of C.D. Broad*, ed. Schilpp, 791-4.

[12] George Santayana, "The Efficacy of Thought," *The Journal of Philosophy* 3 (1906): 410-12.

[13] Santayana, "The Efficacy of Thought," 410.

[14] Santayana, "The Efficacy of Thought," 411.

efficacy of the creature whose life it expressed."[15] Apparently, then, thought may be considered under both psychological (or "ideal") and "natural" aspects,[16] and it is only in virtue of the latter, physical properties that it is efficacious. In Santayana's words,

> Events in nature are never wholly mental, and it is on their material side, through their substance and physical tensions, that they are derived from previous events and help to shape the events which follow.[17]

Santayana sought to emphasise the novelty of his position in his reply to Eliseo Vivas, who had claimed that Santayana's epiphenomenalism committed him to dualism.[18] Santayana rejected the charge, maintaining that, "Sensation, passion and thought are therefore efficacious materially *in so far as* they are material, but not *in so far as* they are spiritual,"[19] the implication being that since questions

---

[15] Santayana, "The Efficacy of Thought," 411. Note that the very title ascribes efficacy to thought, and that in the sentence from which this quote is taken, it is *thought* itself that is said to be active in the creature's natural efficacy. The sentence reads as follows: "*Thought* might still be called efficacious in the only sense, not magical, in which *its* efficacy would be at all congruous with its intent; namely, through the natural efficacy of the creature whose life it expressed." (Emphasis added)

[16] Santayana uses the term "aspects" in connection with his version of epiphenomenalism when he says that our appreciation of this doctrine "is obstructed by superficial empiricism, which associates the better-known aspects of [mental] events directly together, without considering what mechanical bonds may secretly unite them" (George Santayana, "Reason in Common Sense," in his *The Life of Reason* [New York: Charles Scribner's Sons, 1905], 223-4).

[17] George Santayana, *Realms of Being* (New York: Charles Scribner's Sons, 1940), 315.

[18] Eliseo Vivas, "From *The Life of Reason* to *The Last Puritan*," in *The Philosophy of George Santayana*, ed. Paul Arthur Schilpp (New York: Tudor Publishing Company, 1940), 313-50 (at 319).

[19] George Santayana, "*Apologia Pro Mente Sua*," in *The Philosophy of George Santayana*, ed. Schilpp, 497-605 (at 542). (Emphasis added)

about the efficacy of mental phenomena are to be answered by considering those *same* phenomena under both "spiritual" *and* material features, the epiphenomenalism endorsed by Santayana involves no commitment to dualism.

It is interesting to note that Santayana articulates his epiphenomenalism in the context of a materialism that does not require the reduction of the mental to the physical.[20] Similarly, in more recent debates about epiphenomenalism, the kinds of physicalism that are thought to be susceptible to this variety of epiphenomenalism are those that eschew the identification of mental characteristics with physical features. It is easy to see why. After all, if an event causes its effects in virtue of its physiological properties, and if those properties just are its mental properties, then the effects were caused in virtue of those mental characteristics. Thus, the sort of epiphenomenalism that Broad and Santayana describe afflicts only those versions of physicalism that do not countenance the reduction of mental properties to physical features.

Although in this passage Santayana qualifies the sort of efficacy under consideration as efficacy *with respect to material effects*, he elsewhere denies that mental phenomena *qua* spiritual are efficacious with respect to other mental states (Santayana, "The Efficacy of Thought," 411).

[20] As Vivas says, "He [Santayana] is a materialist, but he does not believe that we can reduce mind to matter" (Vivas, "From *The Life of Reason* to *The Last Puritan*," 319). Vivas bases his claim on what Santayana says in Santayana, "Reason in Common Sense," 205-7. The interpretation appears to have been accepted by Santayana, who remarks that Vivas "sees what my principles are, . . . I am condemned without being misrepresented" (Santayana, "*Apologia Pro Mente Sua*," 541).

Unfortunately, the current most popular brands of physicalism are of just this sort. Physicalists today tend to allow that although each mental token is in fact a physical token, each such particular is a token of two distinct types, one mental and the other physical; the mental and physical types cannot be identified with each other. The main reasons for endorsing this view have been doubts about the existence of the psycho-physical laws that would be needed to support the reduction of mental properties to physical features,[21] along with the belief that any given type of mental state is realisable by many different kinds of physical state, and is therefore identical with none of them.[22]

It is thought that once the physicalist thus sunders mental and physical properties from each other--that is, once she becomes a non-reductive physicalist--she thereby opens the door to a host of worries about epiphenomenalism. However, in order to get these worries off the ground we must first articulate a principle that is presupposed (but not explicitly stated) by both Broad and Santayana in their descriptions of the new kind of epiphenomenalism that is now widely thought to bedevil non-reductive physicalism. This

---

[21] The most influential rendering of these doubts derives from Donald Davidson's anomalous monism (Donald Davidson, "Mental Events," in his *Essays on Actions and Events* [Oxford: Clarendon Press, 1980], 207-25).
[22] Hilary Putnam, "The Nature of Mental States," in his *Mind, Language and Reality* (Cambridge: Cambridge University Press, 1975), 429-40.

is, namely, the principle that when one event causes another, only some of the cause's properties need be relevant to the production of the effect.[23] For example, when a red brick is thrown at a window, the window breaks because a brick of that mass and moving at that velocity struck it, and not because it was struck by a brick that was red. The brick's mass and velocity are causally relevant features of the cause, while its redness is not. More generally, for any cause, we must distinguish between those of its properties that are *causally relevant* and those that are inert. The worry that besets non-reductive physicalists is that mental features will fall on the inert side of this divide.

There are four main sources of the fear that mental features are causally irrelevant. The first is a worry about the efficacy of content and derives from the view that all content is broad.[24] To say that content is broad is to say that it is not local. The content of a belief, for example, is not contained within one's head but is instead a much more expansive social matter, being partly constituted by the practices of one's community of language users. But if

---

[23] In the above quotations from Broad and Santayana, this relativisation of efficacy to the cause's properties is captured by Broad in the locution "in virtue of", and by Santayana in the phrase "in so far as".
[24] Hilary Putnam, "The Meaning of `Meaning'," in his *Mind, Language and Reality*, 215-71; and Tyler Burge, "Individualism and the Mental," *Midwest Studies in Philosophy* 4 (1979): 73-121.

this is so, then how can content be efficacious? We like to think that what happens here and now is an effect of causes that were present and exercising their influence in the here-and-now (or the here-and-recent-past). But if content exceeds these boundaries, if it really is the broad social being that many now think it is, then it would seem to be too remote to influence such local phenomena as my raising my arm.[25]

Problems concerning the efficacy of broad content are not peculiar to non-reductive physicalism. As Tim Crane has noted, even if content properties could be reduced to purely physical environmental properties and relations, it would still not be clear how such broad features could be locally efficacious.[26] For this reason, questions arising from the broadness of content will not be a central focus in what follows (although the results that are reached may well have some bearing on those questions).

A second reason for doubting the causal relevance of mental properties is grounded in one type of non-reductive physicalism, namely, functionalism.[27] Functionalists

---

[25] Excellent discussions of the efficacy of broad content can be found in John Heil, "The Legacy of Cartesianism," chap. 2 in *The Nature of True Minds;* and Tyler Burge, "Individuation and Causation in Psychology," *Pacific Philosophical Quarterly* 70 (1989): 303-22.
[26] Tim Crane, "The Mental Causation Debate," *Proceedings of the Aristotelian Society* (Suppl.) 69 (1995): 211-36 (at 224-5).
[27] This problem is presented in Frank Jackson and Philip Pettit, "Functionalism and Broad Content," *Mind* 97 (1988): 381-400; and Frank Jackson and Philip Pettit, "Program Explanation: a General Perspective," *Analysis* 50 (1990): 107-17; and Ned Block, "Can the Mind Change the

conceive of mental features as abstract, higher-order properties. More specifically, each mental property is thought to be the property of having certain first-order properties that causally interact in such a way as to realise the typical causal role that is characteristic of that type of mental state. The problem is that the first-order realising physical properties appear to be doing all the causal work. Causal relevance accrues only to them. The higher-order functional properties attach to their bearers only in consequence of the causal connections having already been fixed by the first-order implementing states.

While this problem does not confront varieties of non-reductive physicalism other than functionalism, it will be addressed in the following chapters, both as a direct challenge to the efficacy of mental properties and as a source of counterexamples to purported tests of causal relevance.

A third challenge to the efficacy of mental properties is posed by the apparent causal closure of the physical realm. According to the principle of closure, only physical events and properties belong to any causal series that results in a physical event. If this is true, then it seems that mental properties are not causally relevant to any

---

World?" in *Meaning and Method: Essays in Honor of Hilary Putnam*, ed. George Boolos (Cambridge: Cambridge University Press, 1990), 137-70.

action that produces physical effects; for, according to the principle of closure, only physical features contribute causally to the production of these effects, and, if non-reductive physicalism is true, mental properties are not physical properties.

It should be noted that we do not strictly need the principle of closure in order to generate this problem. All that is needed is the claim that each of our actions can be accounted for in the language of a physical science. Thus, even if we remain unsure about something as comprehensive as the principle of closure, we might feel confident that the cause of my opening the fridge can be characterised in purely physical terms, by talking about the neurophysiological events that sent signals through my nervous system and into my muscles, causing them to contract and relax in ways that culminated in the fridge door's being opened. If mental features are not physical properties, then it is hard to see how my desire for juice can be fit into this causal sequence.[28]

---

[28] Interestingly, many of the old-fashioned dualistic epiphenomenalists were motivated by similarly modest considerations. They refrained from making sweeping metaphysical claims about the nature of causation or the closure of the physical realm. For example, Huxley briefly considers the worry that mental events are too unlike physical states to act on them, but quickly dismisses it as "superfluous" (T.H. Huxley, "On the Hypothesis That Animals Are Automata," 135). He clearly regards his epiphenomenalism as an empirical hypothesis supported by evidence about reflex actions and the behaviour of unconscious frogs (T.H. Huxley, "On the Hypothesis That Animals Are Automata," 133-4). Broad goes so far as actually to defend the coherence of interactionism and its compatibility with the conservation laws (C.D. Broad, *The Mind and Its Place in Nature*, 95-133), but, in the end, opts for epiphenomenalism on the basis

Returning to the principle of closure, it may be felt that this principle is too strong, and should be replaced by the more modest claim that for every physical event, there are physical events and properties that were sufficient to produce it, or at least to fix the probability of its occurrence; there may indeed be other, non-physical factors in its causal history, but they were overdetermining causes that did not bring about any result (or yield any probability of an outcome) that was not already fixed by the purely physical elements of the causal chain. Unfortunately, this possibility is not a promising basis for an account of mental causation. For even if overdetermination is possible, it is surely not as pervasive as it would need to be in order for every human action to be an effect both of physical and mental antecedents. Thus we still face the problem of according causal potency to the mental in a world in which all the causal work has apparently already been done by physical events and features.[29]

Problems arising from the alleged causal closure (or at least completeness) of the physical realm confront all forms of non-reductive physicalism. Perhaps that explains why, of

---

of the apparent sufficiency of physical explanations to account for our behaviour, together with considerations of "economy" (C.D. Broad, *The Mind and Its Place in Nature*, 475-7).

[29] My statement of this problem of mental causation follows closely Jaegwon Kim's presentation of it, particularly in Jaegwon Kim, "The Myth of Nonreductive Physicalism," in his *Supervenience and Mind* (Cambridge: Cambridge University Press, 1993), 265-84 (esp. 279-84).

the four problems of mental causation here canvassed, this one has received the most attention in the literature. Although the focus in what follows will not at first be on this problem of mental causation, we will have occasion to revisit this set of issues later when it becomes apparent that some of the attempted solutions to this problem have also been advanced as solutions to the fourth problem of mental causation (e.g., the putative `trope' solution). Moreover, much of what we say in coming to grips with the fourth problem (particularly about the very concept of causally relevant properties) will also have application to the closure worries.

The fourth difficulty for mental causation arises in the framework of Donald Davidson's version of non-reductive physicalism, anomalous monism. Davidson is the non-reductive physicalist who has had most often to contend with the charge of being an unwitting epiphenomenalist. Indeed, much of the contemporary debate surrounding epiphenomenalism originated with criticisms of his philosophy of mind. The claim that the principles of anomalous monism support epiphenomenalism seems first to have been made by Frederick Stoutland, Peter Hess and Ted Honderich.[30] Contemporary

---

[30] Frederick Stoutland, "The Causation of Behavior," in *Essays on Wittgenstein in Honor of G.H. von Wright* (*Acta Philosophica Fennica*, XXVIII [Amsterdam: North-Holland, 1976]), 286-326 (at 307); Peter Hess, "Actions, Reasons, and Humean Causes," *Analysis* XLI (1981): 77-81; and

discussions often include references to Stoutland and Honderich as the first to have hit upon this set of worries.[31] Hess's work, by comparison, has been neglected. This is strange, since it was Hess's paper that began the discussion in *Analysis* to which Honderich's paper was a contribution.

The difficulty in Davidson's outlook is thought to derive from three central principles of anomalous monism. The first principle is that mental events causally interact with physical events. Second, events that causally interact fall under strict laws; if *a* causes *b*, then there is a strict law that relates a property of *a* to a feature of *b*. Third, mental properties are absent from strict laws; there are no strict psychological or psychophysical laws.[32] According to the first principle mental events enter into causal transactions, but, in view of the second claim, they do so only by virtue of falling under strict laws. However, given the third claim, none of their mental features are referred to in those laws, and none of their mental qualities can be reduced to the physical properties that are

---

Ted Honderich, "The Argument for Anomalous Monism," *Analysis* XLII (1982): 59-64.

[31] In his chronicle of the debate, Brian McLaughlin credits Stoutland, Honderich and a host of authors writing in the mid-1980's with the epiphenomenalist criticism of Davidson (McLaughlin, "Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical," 131 n. 2). Ernest LePore and Barry Loewer do likewise (Ernest LePore and Barry Loewer, "Mind Matters," *The Journal of Philosophy* 84 [1987]: 630-42 [at 634 n. 10]).

[32] Donald Davidson, "Mental Events," 208.

cited in the strict laws. As a result, it would seem that mental events enter into causal transactions solely because of their strictly nomic physical features and not because of any mental properties that they possess. Thus, according to Davidson's critics, anomalous monism generates epiphenomenalism: even though it allows events that have mental properties to be causes, it implies that no such event enjoys its causal status in virtue of its mental properties.

These points can be illustrated by means of the following example that was presented by Ernest Sosa.[33] Suppose a gun shot kills Ed. The shot is a loud noise, so a loud noise kills Ed. In a sense, that is true. However, adds Sosa, Ed is only killed by the loud noise *qua* a shot, not by the noise *qua* a loud noise. The shot's loudness is causally irrelevant to its effect (Ed's death). This is shown, according to Sosa, by the truth of the following counterfactual: "had the gun been equipped with a silencer, the shot would have killed the victim just the same."[34] The point of Sosa's criticism is that if anomalous monism is true, then every mental property is like the shot's loudness: although an event that has such a feature may be a cause, it is not its possession of that feature that makes

---

[33] Ernest Sosa, "Mind-Body Interaction and Supervenient Causation," *Midwest Studies in Philosophy* 9 (1984): 271-81 (at 277-8).
[34] Ernest Sosa, "Mind-Body Interaction and Supervenient Causation," 278.

it a cause. Causes count as causes only in virtue of their strictly nomic physical qualities. Since mental properties are not strictly nomic, a mental cause's mental aspect contributes nothing to its being a cause. It is thus causally irrelevant.

The key assumption here is that whenever one event causes another, some of the cause's properties are *causally relevant* to the effect, in the sense that the cause produces that effect only in virtue of having those properties. In Chapter Two, I shall accept this assumption for the sake of argument, and then inquire as to the nature of causal relevance thus conceived. I begin by considering an attempt by Barry Loewer and Ernest LePore to arrive at a precise definition of the causal relevance that we take mental features to enjoy. The conclusion of Chapter Two is that their attempt to do so fails. We typically regard mental properties as having a stronger form of causal relevance than certain other properties (e.g., dispositions). However, while LePore and Loewer allow that there is something more in the way of causal relevance, that is, a stronger relevance relation that some properties enjoy (what they call relevance$_1$), this too is a metaphysical relation, one that (by their own lights) only strictly nomic features have. Mental properties, since they are not strictly nomic,

are deficient in this respect. Thus, if we accept their initial assumptions, then mental properties appear to lack something that basic physical properties have, and that would set mental features apart from dispositional features. Moreover, there does not appear to be any promising way to rescue mental properties from this plight, so long as we continue to regard this project as one of defining a metaphysical relation called "causal relevance", whereby a property makes an event the cause of some effect, and in which mental properties can then be shown to stand.

In Chapter Three, I consider recent attempts to define a more explicitly metaphysical relation of causal relevance that mental properties can be said to have. The attempts involve an appeal to tropes, that is, particular instances of properties that can plausibly be thought to enjoy some real efficacy of the sort that seems to be at issue in discussions of causal relevance. I argue that all such attempts fail, but that their failure is instructive, for it shows the futility of trying to understand causal relevance as a metaphysical relation. The lesson taken from this chapter is that insofar as we have any legitimate conception of causal relevance at all, it is a conception only of an intensional, explanatory relevance relation, a creature of

pragmatics and epistemology rather than of metaphysics. Causal relevance is just causal-explanatory relevance.

In the final chapter, I set out what I take to be the most important implications of this realisation for the mental causation debate. First among them is the thought that there is no clear sense in which merely explanatorily relevant properties can be said to *compete* with each other for, and exclude one another from, relevance to a given effect. This opens the way to an accommodating pluralism in which all properties can be seen to be metaphysically on an equal footing, but in which some may stand out in the context as being particularly salient to the explanatory task at hand. It is a central emphasis of this chapter that in this way, pluralism can be achieved without the desperate expedient of explanatory parallelism (first described by William James), in which it is assumed that properties really do compete for and exclude each other from relevance to the same explanandum, and that we can therefore only give mental features some explanatory work to do by delimiting a sphere of non-physical explananda over which they wield an exclusive proprietary concern.

A second theme of the concluding chapter is that since causal relevance is just a relation of explanatory relevance, it can in no sense be understood as a relation

whereby properties *make* the cause to be a cause of just those effects. Not all of the philosophers who emphasise the merely explanatory role of properties draw this conclusion. For example, Fred Dretske says, "Events are causes, but facts explain, and facts, at least in the case of causal explanations, have to do with the properties of the cause that *make* it a cause."[35] However, to think of explanatory relevance in this way, as any kind of *making* of the world's causal series to follow the trajectories that they do, is thereby to render it a metaphysical relation. It is to implicate properties, mere abstractions, in a strange kind of efficacy. For it is to depict properties and the relevance relations between them as being somehow prior to, and determinative of, the distribution of the natural, extensional relation of causation amongst the concrete particulars that bear those properties. It is to represent abstractions as somehow shaping the aggregate of particulars into a causally ordered series. If we can break free of this habit of thought, we shall find that the temptation to assign basic physical *properties* (as opposed to events) a privileged position, as having an exclusive power to make any given causal history take the shape that it has, loses its allure. Once basic physical features have lost their

---

[35] Fred Dretske, "Reasons and Causes," *Philosophical Perspectives* 3 (1989): 1-15 (at 2). (Emphasis added)

privileged position, other properties (including mental features) will no longer seem to be causally deficient by comparison with them.

2. Looking for Causal Relevance

Mihi a docto doctore
Domandatur causam et rationem quare
Opium facit dormire.
A quoi respondeo,
Quia est in eo
Vertus dormitiva,
Cujus est natura
Sensus assoupire.

Moliere[36]

I. The LePore and Loewer Solution

Ernest LePore and Barry Loewer have tried to defend anomalous monism from the accusation of epiphenomenalism.[37] They begin by distinguishing between two kinds of causal relevance. Properties *F* and *G* are said to be "causally relevant$_1$" to event *c*'s causing event *e* iff *c* has *F* and *e* has *G*, and there is a strict law of nature to the effect that *F*-type events cause *G*-type events.[38] Note that in this case the explanandum (what *F* and *G* are said to be relevant *to*) is *c*'s causing *e*. That is, we are trying to explain why *c* and *e* count as being causally interrelated in the first place. By contrast, when we set out to find a property that is "causally relevant$_2$" our explanatory project is not so

---

[36] Moliere, *The Imaginary Invalid (Le malade imaginaire)*, in *The Dramatic Works of Moliere*, trans. Charles Heron Wall, vol. III (London: George Bell and Sons, 1908) 465. According to Wall, the Third Interlude, from which this quote is taken, is untranslateable, being a pun-filled mixture of dog-Latin and French. The character who utters these lines is not a doctor, but, for simplicity, I shall refer to him as such anyway. My thanks to William Seager for suggesting the title of this chapter.
[37] LePore and Loewer, "Mind Matters," 630-42.
[38] LePore and Loewer, "Mind Matters," 634-5.

ambitious: instead of explaining why *c* and *e* count as being causally related at all, we are merely trying to explain why the effect (*e*) has one of the properties that it has, and we hope to account for this by appeal to the nature of its cause (i.e., by appeal to the cause's having one of the properties that it has). Thus we speak of *c*'s having *F* being causally relevant$_2$ to *e*'s having *G*, a relation that obtains iff

i. *c* causes *e*,

ii. *Fc* and *Ge*,

iii. if *c* had not possessed *F* then *e* would not have had *G* (or, in LePore's and Loewer's notation, $-Fc > -Ge$),

and

iv. *Fc* and *Ge* are logically and metaphysically independent.[39]

We now have two ways to rebut the claim that any given set of properties is epiphenomenal: we can either show that those properties are cited in strict causal laws and are thus causally relevant$_1$, or we can show that they meet conditions i-iv and are thus causally relevant$_2$. With regard to this second option, condition iii is the most salient for our purposes, for it outlines a counterfactual dependency relation that can indeed plausibly be taken to establish a type of causal relevance. Moreover, properties need not appear in any strict causal laws in order for this

[39] LePore and Loewer, "Mind Matters," 635.

dependency relation to obtain between them. All that is necessary in order for *e*'s having *G* to depend counterfactually on *c*'s having *F* is that there be a causal law, possibly non-strict (i.e., containing a *ceteris paribus* clause), according to which *F*-type events produce *G*-type events. Such a law is sufficient to support the relevant counterfactual. In the light of this, we cannot infer the epiphenomenal status of a property merely from the fact that it is not mentioned in strict causal laws.[40]

According to LePore and Loewer, the psychological and behavioural features of some states counterfactually depend upon the psychological properties of antecedent events, with the result that the latter features are causally relevant$_2$ to the instantiation of the former properties by those states (assuming that conditions i, ii and iv are also met). For example, when I open the fridge it is true (according to LePore and Loewer) that if I had not wanted a Coke and believed that I could get one from the fridge, then I would not have opened it. That is, if my brain state had not had those psychological features, it would not have been followed by that sort of behaviour. LePore and Loewer interpret this counterfactual as follows: in the possible

---

[40] Others who appeal to counterfactual dependency relations in order to ward off the bogey of epiphenomenalism are John Heil and Alfred Mele, "Mental Causes," *American Philosophical Quarterly* 28 (1991): 61-71; and Terence Horgan, "Mental Quausation," *Philosophical Perspectives* 3 (1989): 47-76.

worlds that are the most similar to the actual world but in which I do not want a Coke and believe that I can get one from the fridge, I do not open the fridge. This counterfactual is true, they maintain, and is supported by a *ceteris paribus* law to the effect that thoughts with those contents typically produce that sort of behaviour.[41]

II. A Problem for LePore and Loewer

Content thus seems to have been exonerated of the charge of epiphenomenalism. Unfortunately, this putative vindication of content's efficacy is merely apparent, for LePore's and Loewer's causal relevance$_2$ comes too cheaply. This is evident from the fact that properties that do not seem to have the sort of causal relevance that we regard mental features as having[42] can nonetheless pass the test laid out in conditions i-iv. The properties that I have in mind are akin to the *virtus dormitiva* that was famously pilloried by Moliere in *Le Malade Imaginaire*. In Moliere's tale, a doctor, when asked why opium induces sleep, answers that it does so because it has a power to induce sleep (a *virtus dormitiva*). We find this answer comical because it is so clearly vacuous: rather than explaining why opium puts people to sleep, the doctor's appeal to its power to induce

---

[41] LePore and Loewer, "Mind Matters," 641.
[42] To which I shall refer in the remainder of this chapter simply as "causal relevance".

sleep merely re-describes the phenomenon for which we had requested an explanation.

Dispositions generally are causally irrelevant to their manifestations, and yet appear to satisfy conditions i-iv. Consider, for example, the claim that Mort fell asleep because he took a sleeping pill. The cause here has the property of dormitivity: it is a taking of a dormitive pill.[43] Is this property causally relevant$_2$ to Mort's falling asleep? Suppose he really did fall asleep as a result of taking the pill, so that conditions i and ii are satisfied; that is, he both takes the pill and falls asleep, and the former event causes the latter one. Moreover, if the cause had not been a taking of a dormitive agent, Mort would not have fallen asleep. So condition iii is met.

It might appear that we have a violation of condition iv, since the property that we are assessing for causal relevance$_2$ (viz., dormitivity) is defined in terms of its effects, and thus would appear not to be "logically and metaphysically independent" of them. Notice, however, that

[43] I should note briefly that there inevitably arises in these contexts a minor glitch: causes and effects are events; dormitivity and other dispositions are properties of things (`continuants'), such as pills, rather than of events; and yet I am treating dispositions as though they were properties of the cause, and thus of events. I will continue to speak of dispositions as properties of causes, where this means that they are properties either of the event that is the cause or of some thing that is implicated in that event. This difficulty is not peculiar to dispositions. Shapes, e.g., are not, strictly speaking, properties of events, and yet it is surely legitimate to speak of squareness as a causally relevant feature of the cause when we, say, try to fit a square peg into a round hole.

iv only precludes a very strong sort of conceptual connection between the properties of the cause and the effect. As LePore and Loewer put it,

> *c*'s being *F* and *e*'s being *G* are metaphysically independent iff there is a possible world in which *c* (or a counterpart of *c*) is *F* but *e* (or a counterpart of *e*) fails to occur or fails to be *G* and vice versa.[44]

In view of this, condition iv bars *c*'s property *F* from being causally relevant$_2$ to *e*'s having *G* only if *F* is defined in such a way as to entail that *that* particular event, *e*, has *G*. It thus ensures, for example, that the explosion's property of "being what destroyed the bridge over the River Kwai" is not causally relevant$_2$ to the bridge's destruction. More generally, clause iv bars any property of the cause that is defined in such a way as to presuppose that the effect *token*, the one that was actually brought about by that cause, has certain of the properties that it has, and it bars that property of the cause from being relevant$_2$ to just those properties of the effect. It does not bar properties of the cause that are defined as dispositions or tendencies generally to produce a certain *type* of effect (but not to produce any effect in particular). For instance, given that dormitivity is a disposition generally to induce sleep, the cause's "being a taking of a dormitive agent by

---

[44] LePore and Loewer, "Mind Matters," 635 n. 13.

Mort" is metaphysically independent (in view of the above definition) of the effect's "being a falling asleep of Mort", since there is a possible world in which the former state of affairs obtains but in which the effect is not a falling asleep of Mort. Sleeping pills don't always work. Thus, dormitivity does not transgress condition iv.

It might be thought that LePore's and Loewer's understanding of logical and metaphysical independence is needlessly strong, and that we should simply amend it to bar those properties of the cause that bear weaker conceptual connections to properties of the effect from being causally relevant to them.[45] Unfortunately, it is hard to see how to do this without inadvertently barring mental features from being causally relevant to their behavioural effects. As A.I. Melden emphasised, mental causes bear an internal connection to their effects in virtue of their content.[46] Indeed, some have plausibly maintained that it is part of the very structure of an intention's content that it represent not only the effect but its own efficacy with respect to it.[47] Moreover, according to one of the most influential contemporary theories of the mind, namely, functionalism, mental features are themselves definable in

---

[45] My thanks to William Seager for pressing this objection.
[46] A.I. Melden, *Free Action* (London: Routledge and Kegan Paul, 1961).
[47] John Searle, "Intention and Action," chap. 3 in *Intentionality* (Cambridge: Cambridge University Press, 1983).

terms of their typical causes and effects, and thus in part as general tendencies to produce certain types of effect. It is thus hard to see how we could change condition iv in such a way as to bar dispositional features from causal relevance to the effects in terms of which they are defined without also catching mental properties in our net.

It would seem, then, that the property of the cause that we are considering (viz., being a taking of a dormitive agent by Mort) passes the test laid out in conditions i-iv and is therefore causally relevant$_2$ to the effect's having the property of being a falling asleep of Mort.

LePore and Loewer may accept all of this, but deny that it poses difficulties for their account, since dispositions really are, in some perfectly respectable sense, causally relevant to their manifestations. After all, they may continue, we find the doctor's explanation in Moliere's example to be comical only insofar as we take his explanandum to be the fact that each person who takes opium falls asleep. In trying to account for this, he is attempting to explain why one event causes another; more accurately, he is trying to explain why, in any given case, someone's taking of opium causes an event that can be characterised as "falling asleep". This task can only be carried out by mentioning properties that are causally

relevant$_1$, for it requires explaining *why one event (c) causes another event (e)*. The fact that dispositions are not up to this task only indicates that they are not causally relevant$_1$, and we have as yet no reason to reject the claim (offered in defense of LePore and Loewer) that dispositions possess a kind of causal relevance that is every bit as strong as that enjoyed by mental features.

Purveyors of causal relevance are likely to find this unsatisfactory, since it would appear that dispositions are standardly irrelevant not only to such facts as that *c* caused *e*, but also to facts of the form "*e* has *G*"; that is, they are causally irrelevant to precisely the sort of explanandum that is supposed to be the focus of causal relevance$_2$. To put it in terms of my example, dormitivity is not causally relevant to Mort's falling asleep (i.e., to the effect's having the property of being a falling asleep of Mort). My reason for saying this derives from Ned Block's discussion of "the standard causal inertness of the second-order."[48] Dormitivity is a second-order property, namely, the property of possessing some or other first-order (usually chemical) property that tends to induce sleep in human beings. It is only in virtue of its having this first-order property that the pill counts as a dormitive agent at all, for it is only because of the first-order property that

[48] Block, "Can the Mind Change the World?" 163.

the pill tends to cause sleep in humans.[49] Dormitivity is thus conceptually after the fact (so to speak) as concerns the causation of sleep: it cannot be what is responsible for producing sleep, since it is itself a *consequence* of the pill's having some other, first-order property that *is* responsible for producing this effect.[50] Putting it very roughly, the pill's power to induce sleep is already "in place" by the time dormitivity arrives on the scene. In fact, dormitivity only arrives on the scene at all because the causal stage has already been set.[51]

LePore and Loewer might object that this reasoning only illustrates, once again, the causal irrelevance$_1$ of dispositions, since it at best only shows that dispositional properties obtain because of, and thus cannot account for, the causal relations of their bearers; and this only amounts to the claim that dispositions cannot be called upon to

---

[49] Strictly speaking, the lower-level property that is causally relevant to the production of sleep need not be first-order. It need only be a property of a lower level than dormitivity.

[50] This sort of reasoning is given by Frank Jackson: "Which of the two properties actually causes the breaking of a fragile object . . . when it is dropped--its fragility, or the categorical basis of its fragility? . . . The fragility of a glass is a matter of its having a nature that would cause the glass to break on dropping; but if the nature would do the causing, then, by Modus Ponens, it does do the causing. . . . But the nature is the categorical basis. Ergo, the categorical basis and not the disposition causes the breaking" (Frank Jackson, "Essentialism, Mental Properties and Causation," *Proceedings of the Aristotelian Society* 95 [1995]: 253-68 [at 256]). Jackson takes a similar line regarding dispositions in Frank Jackson, "Mental Causation," *Mind* 105 (1996): 377-413 (at 393-4); in David Braddon-Mitchell and Frank Jackson, *The Philosophy of Mind and Cognition* (Oxford: Blackwell Publishers, 1996), 264-5; and in Frank Jackson, "The Primary Quality View of Color," *Philosophical Perspectives* 10 (1996): 199-219 (at 202-4).

[51] I am using temporal metaphors in an effort to reflect the order of conceptual priority.

account for the fact that events *c* and *e* are causally related to each other. But this objection misses the point, for when we consider a particular case, such as the one in which Mort takes a sleeping pill, it is clear that the cause can be said to be dormitive only in consequence of the fact that the pill tends to produce a certain *kind* of effect in people. Thus, since dormitivity is a feature of the cause only *because* the cause is likely to put people to sleep, dormitivity cannot itself be called upon to explain why the effect of taking the pill is a falling-asleep. Here, the explanandum (the one that dormitivity cannot be called upon to explain) is not simply the fact that events *c* and *e* are causally interrelated, but is rather the fact that the effect has the property of being a falling-asleep (or, in LePore's and Loewer's schema, it is the fact that *e* has *G*). Dormitivity, then, is a category into which items may be classed only retrospectively, in view of the fact that the causal connection to a particular type of effect (viz., sleep) has already been fixed. To put this in terms of my earlier metaphor, by the time dormitivity arrives on the scene the *whole* causal stage has already been set, and dormitivity is causally relevant neither to the fact that the causal relation obtains nor to the fact that the effect has the property of being a falling-asleep. The fact that

dormitivity and other dispositions meet conditions i-iv, and are thus causally relevant$_2$, only goes to show that causal relevance$_2$ is not a form of genuine causal relevance at all, or, at least, it is not the kind of relevance that we take mental properties to have.[52]

III. Causal-Explanatory Relevance

Whatever view some may take concerning the causal relevance (or lack thereof) of dormitivity and other dispositions, we must acknowledge that dispositions are not entirely devoid of explanatory significance. This comes out most clearly when we consider the debate surrounding one such feature, namely, evolutionary fitness. We often explain the survival of a species by appeal to its fitness. Some have maintained that all such explanations are tautologies, on the grounds that a species counts as being `fit' only if it actually survives; so that the supposed explanation boils down to the trivial truth that the species survived because it survived.[53] This presupposes that fitness is to be defined in terms of an actual token effect, namely, the actual survival of the very species that is said to be fit.

---

[52] Brian McLaughlin also denies that causal relevance$_2$ is genuine causal relevance, in the sense that is at issue in debates about epiphenomenalism. He concludes that LePore and Loewer "should be viewed as attempting to mitigate the severity of the charge of [epiphenomenalism] rather than as attempting to refute it" (McLaughlin, "Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical," 124).

[53] Karl Popper makes this sort of complaint in his "Autobiography," in *The Philosophy of Karl Popper*, ed. Paul Arthur Schillp (La Salle: Open Court, 1974), 3-181 (at 137).

Explanations in terms of fitness are thus assimilated to the explanation of the bridge's collapse by appeal to the cause's property of "being what destroyed the bridge over the River Kwai". Those who have risen to the defence of fitness explanations reject this assimilation. They deny that a species is fit only if it *actually* survives. Instead (they maintain), fitness is a matter of being *disposed* to survive.[54] Fitness is thus rendered a second-order property. As Peter Godfrey-Smith says, it is "identified with the property of having *some* structural property or properties that make the organism likely to have a lot of offspring."[55] Intuitively, this response to the critics of Darwin does restore some explanatory content to appeals to fitness.

The case of fitness illustrates our tendency to accord properties explanatory significance even if they are second-order, dispositional features. This might appear to pose difficulties for the claim that such second-order properties are typically not causally relevant to their bearers' effects. However, it should be noted that the explanatory significance that is secured for evolutionary fitness by the above approach is of a relatively low grade. By construing evolutionary fitness as a disposition, we accord it a degree

---

[54] Where this disposition is understood in terms of probabilities construed as propensities.
[55] Peter Godfrey-Smith, *Complexity and the Function of Mind in Nature* (Cambridge: Cambridge University Press, 1996), 190.

of explanatory relevance, but (again quoting Godfrey-Smith), "Only in the weak sense in which the existence of a disposition explains its manifestation."[56] This weak form of explanatory significance does not bring with it causal relevance. More specifically, even though second-order dispositional properties enjoy a weak variety of explanatory relevance, they seem to remain causally irrelevant and thus epiphenomenal.

One may ask how a feature can have this grade of explanatory relevance, which is a kind of causal-explanatory relevance (i.e., relevance to a causal explanation), and yet not be causally relevant *simpliciter*. Block has suggested how this might come to pass.[57] According to Block, the claim that I fell asleep because I took a dormitive pill is causal-explanatory, "Because it rules out alternative causal explanations of my falling asleep."[58] For instance, it rules out saying that I fell asleep because I had had no sleep the night before. The appeal to the dormitivity of the sleeping pill rules out alternative explanations (and is thus causal-explanatorily relevant) because it "'brings in'" or "*involves* an appeal to" the first-order chemical property of

---

[56] Godfrey-Smith elucidates as follows: "Saying that a trait did well because of its fitness is like saying a glass broke because of its fragility" (Godfrey-Smith, *Complexity and the Function of Mind in Nature*, 189-90).
[57] Block, "Can the Mind Change the World?" 162-3.
[58] Block, "Can the Mind Change the World?" 163.

the sleeping pill that is genuinely causally relevant to inducing sleep.[59] It is perhaps more accurate to say that our appeal to the dormitivity of the pill merely *locates* the first-order causally relevant property (in the chemical composition of the pill) without specifying it in any interesting level of detail, thereby `flagging' or outlining the causal path that culminated in my falling asleep. Switching metaphors, it helps to trace the contours of this causal chain without mentioning any of the causally relevant features from which it is forged. Let us call such explanations, in which we cite properties that have only this low grade of causal-explanatory relevance without also being causally relevant, "flagging" explanations, and the properties therein mentioned "flagging" properties.

The realisation that there are flagging properties alerts us to the danger of conflating mere causal-explanatory relevance (of even this weak variety) with genuine causal relevance, since flagging properties enter into the former, but not the latter, relation. Attention to this distinction should prevent any easy inference from the causal impotence of a property to its causal-explanatory irrelevance.[60] My focus, though, is on denying the

---

[59] Block, "Can the Mind Change the World?" 163. (Emphasis in the original)

[60] Thus we should reject the following inference that Louise Antony draws: "If anomalous monism is true, then the mental properties of reason are causally irrelevant to the production of action. And if that's the case, then the citation of a reason, even if it is the

contrapositive; causal-explanatory relevance (of even this low grade) does not imply causal relevance. From the fact that a flagging property can legitimately be cited in a causal explanation of so-and-so, and is thus in some sense causal-explanatorily relevant to it, it does not follow that that property is causally relevant to the production of so-and-so.

IV. Intensionality Preserved

It might be thought that the notion of a flagging explanation extensionalises explanatory contexts, and will therefore be found objectionable to Davidsonians and others[61] who assert the intensionality of explanatory contexts. For it appears that one can mark out (`flag') the causal route that culminated in the effect that is to be explained, and thereby rule out alternative causal explanations, simply by referring to the cause in some way or other. It does not matter how one picks out the cause; as long as one does so, one will have succeeded in providing a flagging explanation. This claim is to be rejected. To see

---

citation of a cause, is not in itself a causal explanation of the subsequent event" (Louise Antony, "The Causal Relevance of the Mental: More on the Mattering of Minds," *Mind and Language* 6 [1991]: 295-327 [at 323]). Two pages later, Antony elevates the grounds for this inference to the status of a principle, the *Causal Potency Condition*, according to which, "A citation of a cause counts as explanatory only if the properties cited in the description of the cause are causally potent properties" (Antony, "The Causal Relevance of the Mental," 325).

[61] Notably, P.F. Strawson in his "Causation and Explanation," in *Essays on Davidson*, eds. Bruce Vermazen and Merrill Hintikka (Oxford: Oxford University Press, 1985), 115-35.

why, contrast the distinction between flagging explanations and explanations that cite causally relevant features with a distinction that is drawn by David-Hillel Ruben.[62] Ruben distinguishes between *giving* an explanation of something and merely "implying that there is *some* explanation of that thing, without actually giving it."[63] For example, I `give' an explanation when I say that the hurricane caused the loss of life; while I merely imply that there is some explanation of the loss of life when I say that the event reported in *The Times* on Tuesday is responsible for the fatalities. By making the latter claim, I succeed in referring to the cause of the deaths and yet do not succeed in explaining them, for I have not referred to the cause in the right way. As a result, I have merely indicated that an explanation can be had by differently describing the event to which I have referred. Ruben uses this example to illustrate the intensional nature of explanation.

When I merely imply that there is an explanation by allusion to *The Times*, I provide no new information about the causal path that culminated in the loss of life: my claim, taken on its own (in isolation from any further claim about what was reported in the newspaper on Tuesday), does not rule out any of the alternative causal paths that could

---

[62] David-Hillel Ruben, *Explaining Explanation* (London: Routledge, 1990), 164.
[63] Ruben, *Explaining Explanation*, 164.

have resulted in that event. I do, of course, provide you with the information that the cause, whatever it is, was mentioned in the Tuesday edition of *The Times*, but this is not the same as providing you with information about what it was that led to the loss of life. For all you have yet been told, the fatalities could have been the result of an earthquake, or a tidal wave, or an outbreak of the Ebola virus, or . . . . All of these alternative candidate causes remain in the running in the light of my characterisation of the cause as the event that was reported in *The Times*. By contrast, I rule out some candidate causes when I say that Mort fell asleep because he took a sleeping pill: I rule out any causal chain that does not proceed through the pill. Even though I have not specified the cause in any interesting detail, I have at least located it. I thus succeed in providing information about the causal history that lies behind Mort's falling asleep.

Michael Patton has objected[64] that the claim

> (C) The event reported in *The Times* on Tuesday caused the deaths.

does rule out alternative causal paths, namely, those that were not reported in *The Times* on Tuesday; so that the property of being reported in *The Times* on Tuesday is, by my

---

[64] In Michael Patton, commentary on "Does Anything Break Because It Is Fragile?" by Paul Raymont (paper presented at the annual Mid-South Conference in Philosophy, Memphis, TN, February 1998).

standards, a causal-explanatorily relevant feature. Against this objection, let us note that being reported in *The Times* on Tuesday is an extrinsic, inessential feature of the causal sequence that led to the deaths, one that just about *any* causal trajectory could satisfy. In this sense, characterising the cause in this way does not rule out alternative causal paths. Since practically any causal path could satisfy this description, we are not brought any closer to discovering which causal path in fact lies behind the deaths when the cause is characterised in this fashion.

Patton might dismiss this as irrelevant, since the event that is in fact the cause of the deaths is the *designatum* of the description, "The event reported in *The Times* on Tuesday", and clearly the claim that the deaths were caused by *that* event does rule out all causal paths except for the actual one.[65]

In response, recall that in order to ascertain whether the property of being reported in *The Times* on Tuesday is causal-explanatorily relevant to the fatalities, we must gauge the informational content of (C) by determining whether it rules out any of the alternative causal paths

---

[65] This is suggested by his remark that we here have "a case of rigid designation by a definite description." Similar reasoning may lie behind the quasi-Lewisean view described in William Child, *Causality, Interpretation and the Mind* (Oxford: Clarendon Press, 1994), 104. According to that view, "`The *Titanic* sank because of an event mentioned on page three of the *Guardian*' is no less of an explanation than `The *Titanic* sank because it hit an iceberg'."

that might have given rise to that effect. Patton does this by first fixing the semantic value of (C) (including the reference of its definite description) and then asking whether that propositional content excludes any of the alternative causal paths. But surely this is not the appropriate concept of information to apply here. Surely the type of information at issue has a more fine-grained and intensional nature deriving from its close dependence on belief and knowledge contents. For example, if Lois Lane knows that Superman can fly but not that Clark Kent can fly, then these two claims express different information even though they have the same extensional content.[66] Similarly, since one can know (C) without knowing anything about the cause of the deaths (except that it was reported in the newspaper), we may conclude that (C) conveys no new information (in the relevant sense of "information") that rules out any of the alternative causal paths.

It would be wrong to operate with this intensional notion of information if we were trying thereby to reach metaphysical conclusions (e.g., about whether the cause

---

[66] The example is borrowed from Terence Horgan, "Jackson on Physical Information and Qualia," *Philosophical Quarterly* 34 (1984): 147-53 (at 150). My response to Patton owes a great deal to Horgan's discussion of the distinction between extensional and intensional notions of information. He says, "If one lacks an item of knowledge then one lacks the corresponding item of information . . . . This close link between knowledge and information means that information inherits the intensionality of knowledge," (Horgan, "Jackson on Physical Information and Qualia," 150).

really *is* the event described in *The Times*). However, when we are deciding whether a claim rules out alternative causal explanations (in the course of reaching a verdict about causal-explanatory relevance), we should assess its informational content in the intensional sense of information. We are, after all, talking about *explanatory* relevance, and explanation is essentially epistemic and intensional. Moreover, it is important to keep in mind that what is being assessed for causal-explanatory relevance is a *property of* the cause (or the cause *under an aspect*) and not the cause *simpliciter*.[67] But if we adopt Patton's extensional concept of information, then our focus will be on the reference of the definite description in (C), and thus on the cause itself rather than one of its properties. It is only when we deploy the intensional notion of information that the focus remains on the property cited in the definite description (i.e., on the property by means of which we secure reference to the cause) and not on the cause itself.

In conclusion, the notion of a flagging explanation is compatible with the view that explanation is intensional. For in giving a flagging explanation, the cause cannot be referred to in just any old way. It must be characterised in

[67] As will later be emphasised, the same can be said of causal relevance, a fact that is often overlooked in debates about mental causation.

a manner that at least locates the cause and thus rules out some of the alternative causal histories that might have produced the effect in question. We do this when we characterise the cause in dispositional terms. Dispositions, then, while not causally relevant, do at least possess causal-explanatory relevance.

V. Why Moliere's Doctor is a Quack

With the notion of a flagging explanation in mind, we can better understand the inadequacy of the explanation given by Moliere's doctor. It will help to distinguish between three kinds of explanatory project. I shall do so in terms of the three kinds of why-question that give rise to them. First, we may ask why *e* is *G*. It is (at least sometimes) appropriate to answer this kind of question without citing properties whose relevance to the effect is enlightening in any detailed or interesting way. In such cases, it is sufficient merely to mention a flagging property, such as a disposition; for we thereby point to some other event, *c*, and suggest that something about *it* is responsible for bringing about this effect. In doing so, we provide new information about why *e* has come to be *G*.

Next, we may already know that *c* caused *e*, but ask why *c* and *e* count as being *causally* related. Pressed far enough, this question requires giving a theory of causation. In

answering it, we may say, for instance, that *c* and *e* are causally connected just in case one of *c*'s properties is linked by a strict law of nature to one of *e*'s properties; and that it is in virtue of these two properties that the two events count as being causally related. This, evidently, is the sort of why-question that LePore and Loewer have in mind when they introduce their notion of causal relevance$_1$.

Finally, we may already know that *c* caused *e* to be *G*, but ask why it did. That is, we may ask what it is *about c* that is responsible for *e*'s being *G*. This is the sort of question that was put to Moliere's doctor. He was asked why opium puts people to sleep. In answering this kind of why-question, it is not appropriate to mention a flagging property that just locates the causally relevant factor in the opium. After all, the fact that we are asking why *the opium* puts people to sleep shows that we have already located the cause of the sleep in the opium, and are now asking what the causally relevant factor of the opium is. Since our question indicates that we already know that the causal path goes through the drug, we are clearly not asking for a mere location of the causally relevant factor, but rather for a specification of it. We want to know what it is *about* the opium that is responsible for putting people to sleep. Those who accept the terms of the debate about causal

relevance (to which LePore and Loewer are contributing) are likely to construe this as an attempt to discover a first-order (likely chemical) property that opium has and in virtue of which it counts as a dormitive agent. Regardless, we expect Moliere's doctor to tell us about properties of the opium that have this more interesting and enlightening relation to the effect. Instead of doing so, he only gives us a flagging explanation, one that locates the causally relevant feature in the opium. He thus merely repeats what we already know (indeed, what our why-question presupposes), namely, that something about the opium is responsible for causing sleep.

On this reading, Moliere's joke illustrates the causal irrelevance of dispositions, for it derives its humour from the fact that it is pointless to refer to dispositions when we are asked to name causally relevant properties.

It should be noted that in some contexts, it is appropriate to respond to the third kind of why-question without specifying properties that are relevant in some more interesting way to the effect. For instance, as Georges Rey has pointed out, it could be appropriate to reply to the sort of question that was put to Moliere's doctor by saying that the person who took the opium had an allergic reaction

to it.[68] Of course, the allergic reaction would be a reaction *to the opium*, so that to give this answer is to locate the causally relevant factors partly in the drug and partly in the opium-taker's body. We do likewise when we explain the effect by saying that the opium has a dormitive virtue. The difference between these two answers is that the latter one suggests that the causally relevant factors in the opium-taker's body are typical of other people's bodily constitutions, so that we can expect the drug to induce sleep in most other people too. By contrast, the answer that posits an allergic reaction implies that there is something atypical about the opium-taker's body and its reaction to the drug, so that we should not expect opium to cause sleep in most other people. This answer thus focuses our attention on the causally relevant factors that lie within the opium-taker's body; it marks them as being more pertinent to the explanatory task at hand. It should be clear from the context, however, that in Moliere's play the emphasis is instead on the causally relevant factors that lie within the opium. They are more salient for our explanatory purposes. This is evident from the fact that in the play, the question at issue is not why opium puts a particular person to sleep, but rather why it puts people (generally speaking) to sleep.

---

[68] Georges Rey, *Contemporary Philosophy of Mind* (Oxford: Blackwell Publishers, 1997), 207 n. 36.

This way of framing the question indicates that a person who falls asleep does not exhibit an atypical reaction (e.g., an allergic one) to the drug.

So it is clear from the context in Moliere's play that the focus of attention is on the causally relevant factors that lie within the opium (and on how they interact with a typical human body to induce sleep). That is what is being asked about. The humour in Moliere's joke derives from the fact that the doctor does not succeed in identifying any such causally relevant factors merely by positing a *virtus dormitiva* in the opium, for to posit such a disposition is just to tell us what we already know, namely, that the drug has something, *as yet unspecified*, that puts people (in general) to sleep.[69]

---

[69] Rey claims that in "La Maladie Imaginaire" (*sic*) there are other appropriate answers to the question that was put to the doctor that do not identify the causally relevant factors. He says, "It is not altogether clear precisely wherein the problem of dormative virtue explanations is supposed to lie. After all, something's *being a sleeping pill* can be explanatory of why it put someone to sleep (it wasn't an accident, an allergic reaction, something that had been combined with the pill . . .)" (Rey, *Contemporary Philosophy of Mind*, 207 n. 36). (Emphasis and ellipsis in the original) Note, however, that the first and third possibilities (adumbrated in his parenthetic remark) preclude its being *the pill* that caused the sleep. After all, to say that it was "an accident" is precisely to deny that the pill caused the person to fall asleep--something else did, and it was merely a coincidence that it did so right after the pill was taken. To say that something had been combined with the pill is to locate the causally relevant factor in that other thing (the "something" that was combined with the pill), and is thereby to deny that *the pill* put the person to sleep. In Moliere's story, these possibilities have already been ruled out, since the characters in the play already know that it is *the opium* that generally causes sleep; they know that it does so not just by fluke or accident in one particular case, and not just when it is combined with something else. Hence, *in the context*, something's being a sleeping pill (or having a dormitive virtue) is not explanatory, for it does not rule out any possibilities that have not already been ruled out by the people who framed the question. It does not meet their request for an

VI. Costly Intuitions

My criticism of LePore and Loewer challenges their use of counterfactuals as a test of the sort of causal relevance that we take mental features to have. I claim that dispositions are not causally relevant to the effects in terms of which they are defined, but that they nonetheless pass the counterfactual test for causal relevance to those effects (as given in LePore's and Loewer's condition iii). From this I conclude that the counterfactual test is inadequate as a test of causal relevance. It is tempting to respond to this argument by turning it on its head. Why not assume the second premise together with the denial of the conclusion, and infer from these two assumptions that the first premise is false (i.e., infer that dispositions really are causally relevant to the effects in terms of which they are defined)? Such an inversion seems especially appealing if we apply a cost-benefit analysis to the claims at issue. By denying the conclusion, we get an intuitively plausible test of causal relevance that our mental features pass. All we have to do in return is to curb the intuitions that supported the first premise: dormitivity, immunity, fitness

---

identification of the causally relevant properties of the opium. In short, the alternative explanations envisioned by Rey are not really answers to the why-question that was put to Moliere's doctor, for they do not explain the explanandum that is presented in that question. Instead, they are at best attempts to correct a mistaken presupposition of the question (e.g., the assumption that it is the opium that puts people to sleep, or that opium has this effect on most people --not just on those who have an allergy to it).

and other dispositions really do have some robust form of causal relevance which can be understood in terms of their accounting for their bearers' production of sleep, health and survival respectively. By reconciling ourselves to this mildly counterintuitive view, we spare ourselves the extremely counterintuitive consequences of epiphenomenalism.

This seems to be the approach favoured (tentatively) by Block.[70] In his view, we can avert the dire consequence of epiphenomenalism by denying the first premise and conclusion of the above argument, thereby affirming a counterfactual test of the sort of causal relevance that we regard mental features as having. Indeed, Block sees this as a reason for preferring a counterfactual test to a nomological one;[71] since it is, in his view, not clear that one can avert epiphenomenalism if causal relevance is nomologically construed.[72]

It is puzzling that Block regards counterfactual tests of causal relevance as being free of the sorts of difficulties that afflict nomological tests. For instance, he notes that any nomological standard of causal relevance must face the problem that, "There can be . . . nomological correlation of *F* with *G* without a causal relevance relation

---

[70] Block, "Can the Mind Change the World?" 157.
[71] Block, "Can the Mind Change the World?" 159, 166.
[72] Block, "Can the Mind Change the World?" 146-8. In fact, according to Block, nomological conceptions of causal relevance seem positively to support epiphenomenalism (Block, "Can the Mind Change the World?" 157-8).

between *F* and *G*."[73] As an example, he asks us to suppose that there is a metal rod connecting a fire to a bomb. If the rod's thermal conductivity is increased, the rod will transfer enough heat to the bomb to cause it to explode. Block adds that according to the Wiedemann-Franz law, we cannot (under normal conditions) increase a thing's thermal conductivity without also raising its electrical conductivity. (It is not clear, from Block's presentation, whether the implication goes the other way as well: we cannot increase the rod's electrical conductivity without also boosting its thermal conductivity. Presumably it does, for in order for Block's counterexample to work, it must be the case that whenever such a rod's electrical conductivity is increased, there is an explosion. But this would not follow if there were ways of increasing the rod's electrical conductivity without increasing its thermal conductivity, and thus without causing the bomb to explode.) Given this law-like correlation, it follows that an increase in the electrical conductivity of the rod is nomologically sufficient for the bomb to explode. Thus, on a nomological test of causal relevance, the cause's being an increase in electrical conductivity is causally relevant to the explosion. And yet, Block adds, clearly it is only the increase in thermal conductivity that is causally relevant.

---

[73] Block, "Can the Mind Change the World?" 146.

While this may be so, it is hard to see why these considerations should not also pose a problem for counterfactual tests of causal relevance. After all, nomological connections ground counterfactuals. If the increase in electrical conductivity really is nomologically sufficient for the explosion, then there should be a counterfactual to the effect that if the cause had not been an increase in the rod's electrical conductivity, then the effect would not have been an explosion. In which case, the counterfactual test is equally guilty of yielding the false claim that the rod's rising electrical conductivity is causally relevant to the explosion.[74] It seems, then, that the difficulties raised by Block's counterexample equally beset both the nomological and the counterfactual conceptions of causal relevance, even though Block only presses the case against the nomological account.[75]

A further difficulty for Block's approach is that it requires us simply to set aside any intuitions that militate

---

[74] Similar problems arise from Leibniz's Pre-established Harmony, according to which there are law-like correlations between the activities of the various monads without any causal relations (and thus without any relations of causal relevance) between them. In such a model, the correlations are laws of nature, and ground counterfactuals, and yet neither the laws nor the counterfactuals suffice for relations of causal relevance.

[75] David Robb maintains that such `fork' cases as the one that Block considers invalidate both the nomological and the counterfactual tests of causal relevance, although he does not make note of Block's apparently inconsistent application of the `fork' criticism only to the nomological criterion, while endorsing a counterfactual test. See David Robb, "The Properties of Mental Causation," *The Philosophical Quarterly* 47 (1997): 178-94 (at 181).

against the causal relevance of dispositions to their manifestations. This becomes less easy to do if we discover additional counterintuitive consequences that arise from the attribution of any robust kind of causal potency to dispositions. Block mentions one such additional consequence, according to which we must countenance "a bizarre systematic overdetermination."[76] For whenever a first-order property is causally relevant to an effect, there will be a causally relevant, second-order, dispositional property, namely, the one that consists in the having of some or other first-order property that is causally relevant to that effect. There will also be a causally relevant, third-order, dispositional property, namely, the one that consists in the possession of a second-order property that is causally relevant to that effect; and so on, *ad infinitum*.

Block is prepared to accept this regress of overdetermining causally relevant factors because it does not, in his opinion, exhibit the features that make overdetermination worrisome. According to him, we are usually wary of positing overdetermining causes "because it is wrong, other things equal, to postulate coincidences."[77] For instance, if we know that Mort fell asleep because he

[76] Block, "Can the Mind Change the World?" 158.
[77] Block, "Can the Mind Change the World?" 159.

took a sleeping pill, we will be reluctant to endorse the claim that he, by coincidence, also fell asleep because he had had little sleep the night before. But since the above-mentioned regress of overdetermination involves no such coincidence, Block does not find it objectionable.[78]

Others have found overdetermination to be problematic for reasons that Block does not take into consideration. For example, in a paper in which they argue for the causal impotence of dispositions, Elizabeth Prior, Robert Pargetter and Frank Jackson rest their case centrally on the claim that if dispositions were causally potent, they would systematically overdetermine the effects in terms of which they are defined.[79] These authors have no objection to overdetermination if it is simply a coincidence of several sufficient conditions. However, in their view, the overdetermination in which dispositions are implicated is not so innocuous; it is not simply a case of there being more than one antecedent sufficient condition for the effect. It is instead a case of the effect's having more

---

[78] Tim Crane adopts a similar view in Crane, "The Mental Causation Debate," 232. Strangely, Crane misinterprets Block as refusing to countenance overdetermination (Crane, "The Mental Causation Debate," 233). But Block *says*, "We are normally reluctant to accept overdetermination because it is wrong, other things equal, to postulate coincidences. . . . But no . . . coincidence would be involved in the series of higher-and-higher-order causally efficacious properties I mentioned. If accepting such a series of causally efficacious properties is a price that must be paid for avoiding the problems to be mentioned, it can be paid" (Block, "Can the Mind Change the World?" 159).

[79] Elizabeth Prior, Robert Pargetter and Frank Jackson, "Three Theses About Dispositions," *American Philosophical Quarterly* 19 (1982): 251-7 (at 255-6).

than one *operative* sufficient condition. Prior, Pargetter and Jackson do not elaborate on their notion of an "operative" causally sufficient condition, but they seem to have in mind something like the following. An assassin slips a pill into the Generalissimo's dinner. The pill can kill in two ways: it contains chemical 1, which can stop the heart from pumping, and chemical 2, which can stop the lungs from functioning. These properties are each sufficient to kill the Generalissimo, but one of them will "get there first," so to speak; that is, one of them, say, chemical 1, precipitates a chain of effects that culminates in the Generalissimo's death, thus cutting short the chain of effects triggered by chemical 2, which would have culminated in his death had it not been cut short by the heart attack. It might be thought possible for neither causal chain to be cut short, and for both causal sequences to be simultaneously operative in generating the effect. However, it seems that this is precisely the type of overdetermination that Prior, Pargetter and Jackson aim to rule out.

It is unclear whether we can rule out the very possibility of this kind of overdetermination, as Prior, Pargetter and Jackson seem to want to do. After all, it seems *prima facie* possible for an effect to have more than

one operative cause. Nevertheless, on the basis of this understanding of overdetermination they are likely to find unacceptably problematic the regress of overdetermining causally relevant factors that Block seems willing to countenance. For now the worry that arises will be not simply that there are, in any given case, infinitely many antecedent conditions that are severally sufficient for the effect; it is instead that there are, in each case, infinitely many causally sufficient conditions that are actually operative on that occasion in bringing about the effect. It may be argued that this puts a greater strain on our intuitions than the sort of overdetermination that Block considers. Is it so easy to accept that whenever there is a causal relation, infinitely many, individually sufficient features of the cause actually come into play (or become "operative") in determining the effect?

V. How to Interpret Block's Examples

It might be thought that Block's attempt to dispatch epiphenomenalism by according causal powers to dispositions (and by accepting a counterfactual test of causal relevance) results in an overly hasty dismissal of his own examples, which seem to show that dispositions only sometimes are, and often are not, causally relevant to their manifestations.[80] In one example, he asks us to suppose that the redness of a

[80] Block, "Can the Mind Change the World?" 155-6.

bullfighter's cape provokes a bull. The redness of the cape is thus causally relevant to the bull's anger. Now consider a second-order, dispositional property of the cause: the property of having some or other first-order property that provokes the bull; that is, the property of being provocative. Presumably, it is only the first-order property of redness that is causally relevant to the bull's anger, not the second-order property of provocativeness. After all, provocativeness does not itself provoke the bull. As Block says, "The bull is too stupid for that."[81]

Now let us suppose that the bull is much smarter (and very sensitive). The bull can conceptualise second-order properties, and realises that the cape is provocative. It still sees the redness of the cape and is angered by it, but now additionally takes affront at the fact that it is a provocative cape, and is angered still further. Provocativeness seems to have gained something here, and what it has gained seems to be best described as causal potency.

The point can perhaps be made more clearly by contrasting (as Block does) the stupid bull with a member of the Society for the Prevention of Cruelty to Animals, who is

---

[81] Block, "Can the Mind Change the World?" 155. Of course, as Block notes, in real life the redness is irrelevant, since bulls are colour-blind. To avoid needless complications, I shall continue to focus on the fictional case in which the bull really is provoked by the cape's redness.

angered by the provocativeness of the cape *instead* of being angered by its redness. Here again, provocativeness seems to have something that it lacks in the case of the stupid bull. It is plausible to characterise this `something extra' as causal relevance.

Block is willing simply to set aside these intuitions.[82] He does not, however, offer any alternative proposals for how to assay them in a manner that does not require ascribing causal relevance to provocativeness only in the smart-bull (or SPCA) case and *not* in the stupid-bull case. In the absence of any such account, one might find the intuitions elicited by these cases less tractable than Block does. Indeed, it may even seem that Block loses sight of a distinction that marks a genuine difference, specifically, the difference between cases in which dispositions are efficacious and those in which they are not.

Unless we can dispel this impression, we will have to acknowledge that dispositions are usually causally impotent, like provocativeness in the stupid-bull case. Only in those rare cases in which they are recognised by intelligent agents do dispositions possess any robust kind of causal relevance to their manifestations. And yet counterfactual accounts of causal relevance do not acknowledge this difference between the stupid- and smart-bull cases, for on

---

[82] Block, "Can the Mind Change the World?" 157.

a counterfactual account the cape's disposition to produce anger comes out equally causally relevant (to the bull's anger) in both scenarios: in both cases it is true that if the cape had not been provocative, the bull would not have become angry.

These considerations are not conclusive, though, for there is another way of interpreting the smart-bull case that does not accord efficacy to the cape's provocativeness. This approach exploits a peculiarity that is shared by all the examples adduced by Block as putative instances of causally relevant dispositions. In each such case, the disposition is recognised by an intelligent agent, who then acts (or reacts) on the basis of this recognition. For instance, the smart bull and the SPCA are angered not by the provocativeness *per se* but rather by their belief that the cape is provocative. This belief might be thought to `screen off' the provocativeness from the ensuing anger; that is, it might be thought to render the disposition irrelevant to the effect, since the same effect would have resulted even if the cape had not been provocative but the bull (or SPCA) had nevertheless believed (mistakenly) that it was.[83]

On this reading, Block's examples are consistent with exactly the opposite of the view that he initially took them

[83] My thanks to William Seager and Ronald de Sousa, who independently raised this point.

to illustrate. He initially took them to show that dispositions are causally relevant only when recognised by an intelligent agent, and are otherwise inefficacious. In light of the `screening off' considerations, however, it seems to be precisely when they are thus recognised that dispositions are impotent, for that is when they are screened off; and it is open to one to maintain that they are otherwise efficacious.

It is not clear, though, that the `screening off' considerations have this result. To see why, note that in any causal chain that culminates in some effect (*E*), the immediate cause of *E* screens off all earlier members of that causal chain from being causally relevant to *E*. For example, in the chain {*C, D, E*} the properties of *D* screen off those of *C* from *E*: if *D* had occurred without having been produced by *C*--if, say, *D* had been brought about by *B* instead--then *E* would still have occurred (assuming that *D* is causally sufficient for *E*). As a result, if we take screening off to be a definitive test of causal relevance, then only the properties of *D* can be causally relevant to the production of *E*. But this reasoning is surely unsatisfactory, as it would render even the redness of the cape causally irrelevant to the bull's anger. After all, the cape's being red is not the immediate cause of the bull's anger, for it

only provokes the bull by causing light to be reflected in the manner in which red things typically do under normal conditions, which in turn causes a certain reaction in the bull's eyes, which in turn causes . . ., which in turn causes the bull's anger. The cape's being red is screened off by each successive event in this causal chain. And yet, intuitively, we do not take this to deprive the cape's being red of its causal relevance to the bull's anger. Similarly, in the smart-bull case, the cape's provocativeness is not necessarily screened off from causal relevance to the bull's anger by subsequent factors in the causal chain that culminated in that anger.

Still, the `screening off' criticism draws our attention to a weakness in any argument that relies on Block's examples; for it exploits the fact that in each of those examples, a disposition appears to be efficacious only because an intelligent agent recognises it and acts on the basis of that recognition. This at least points to the possibility that the causal relevance does not penetrate back beyond the intelligent agent's recognition, all the way to the disposition. After all, the intelligent recognition of all manner of features can prompt various sorts of responses, without this entailing that the features recognised are themselves *causally* relevant to the response.

For instance, the recognition of two's being an even number might lead me to act in a certain way, but it would be problematic (to say the least) to infer that this abstract, mathematical feature exerted any *causal* influence on my behaviour. Similarly, my recognition that a certain disjunctive predicate had been satisfied can cause me to respond in a certain way (e.g., to answer "yes" to the question whether it had been satisfied), and yet, as William Seager and David Owens have maintained, it is not clear that unduly *disjunctive* properties can ever be causally relevant to anything.[84] This is all by way of suggesting that acts of recognition by an intelligent agent serve as a sort of buffer against the backwards propagation of causal relevance to the features thereby recognised.[85] In light of this, we cannot simply infer from the smart-bull or SPCA examples that the disposition therein recognised (viz., provocativeness) is causally relevant to any of the effects of that recognition (viz., the ensuing anger). Hence, it is not clear that we should explicate the intuitive difference between the stupid- and smart-bull cases by saying that the

---

[84] William Seager, "Disjunctive Laws and Supervenience," *Analysis* 51 (1991): 93-8; and David Owens, "Disjunctive Laws?" *Analysis* 49 (1989): 197-202.

[85] Thus mental causation is puzzling `at both ends', so to speak. It is puzzling how the mental causes anything, and it is puzzling how the mental can be sensitive to features that are themselves utterly devoid of any causal clout, and cause certain things to happen on the basis of this sensitivity.

disposition has causal relevance in the latter case but not in the former one.

Even if provocativeness is causally relevant when it is recognised, this does not lend any credence to the claim that it is inefficacious when it is not thus discerned. There certainly is a difference between the smart- and stupid-bull cases, but there is no reason to assume that the difference lies in the fact that provocativeness is efficacious only in the former case, and not in the latter one. Indeed, it seems more plausible to interpret the difference as consisting simply in the fact that in the smart-bull case an act of recognition is interposed between the cape's provocativeness and the ensuing anger, whereas for the stupid bull there is no such mediating act of recognition. It is open to one to maintain that in both cases the cape's provocativeness is causally relevant to the anger, and that the only difference is that the smart bull is aware of this causally relevant factor while the stupid bull is not.

For these reasons, I will not rely on Block's examples, and his interpretation of them, in my case against according any robust kind of causal efficacy to dispositions. In fact, in what follows, I shall argue for a view that is inconsistent with the argument from Block's examples. For

while Block allows that dispositions do sometimes have real efficacy (viz., when they are recognised), I shall attempt to show that they never do.

VI. Unwieldy Disjunctions

Frank Jackson denies efficacy to dispositions on the grounds that since fragility (e.g.) is simply *defined* as `being such as to be disposed to break when struck (or dropped or . . .)', the liability to break when struck is essential to being fragile.[86] But, as Hume taught us, none of a state's genuinely causal powers is possessed by it essentially.[87] Hence (Jackson concludes), fragility itself has no causal power with respect to breaking. More generally, no type of state that is defined in terms of a causal power (i.e., no dispositional state) can itself be said to bestow that causal power upon its bearers.[88] The causal power with respect to breaking, for example, is

---

[86] Jackson, "Essentialism, Mental Properties and Causation," 257.

[87] Something like this premiss seems to be the motivation behind LePore's and Loewer's condition iv. However, unlike condition iv, Jackson's reasoning does deny causal relevance to dispositions. This is because according to Jackson's premiss, a property is causally irrelevant to an effect if it is defined as a general tendency to produce effects of that type. By contrast, in condition iv, a property is irrelevant to an effect's being thus-and-so if it is defined as that which makes that token effect to be thus-and-so.

[88] Eli Hirsch takes a similar line. He writes, "A property defined dispositionally in terms of certain causal powers cannot itself be said to have those causal powers. Consider the property *Q* of being able to melt things. I assume that *Q* has no . . . causal power to melt things. Something has *Q* in virtue of having another property *P* that does have the power to melt things. . . . It is a necessary fact that anything that has *Q* is able to melt things but a contingent fact that anything that has *P* is able to melt things" (Eli Hirsch, *Dividing Reality* [New York: Oxford University Press, 1993], 62-3).

conferred upon a vase by virtue of its having a given molecular structure, not by virtue of its being fragile. So we should not number fragility among the causally relevant factors that contributes to the breaking of the vase.

Jackson does not deny that there are necessary truths that postulate genuinely causal connections. For instance, it is necessarily true that the bridge over the River Kwai was destroyed by whatever destroyed it; or, to use Jackson's example, it is a necessary truth that fatal accidents cause death.[89] Note, however, that in these cases the necessity arises from the manner in which the cause is described. By contrast(continues Jackson), the necessity that attaches to the tendency of fragile things to break when struck is not simply a product of how we choose to refer to fragility. Regardless of *how* we characterise the state of being fragile, we end up talking about something that is, by definition, typically connected to breaking when struck. This sort of necessary connection, deriving as it does from the state that gets described rather than from the mode of describing it, cannot be causal if Hume was right about the contingency of the causal relation.

Jackson's argument may seem convincing as far as it goes, but it can be questioned whether it goes far enough. He assumes that dispositions are definable as second-order

---

[89] Jackson, "Essentialism, Mental Properties and Causation," 257.

properties that involve existential quantification over first-order properties. Thus Jackson interprets, `being such as to be disposed to break when struck' to mean `having the second-order property of having some or other first-order property that tends to cause breaking when struck'. However, those who advocate the causal efficacy of dispositions are likely to regard them as being more closely wedded to their causally relevant, first-order bases than this model allows. Specifically, they may favour defining each disposition purely extensionally, as the simple enumeration of its various causal bases. Thus, fragility is identified with the second-order property of having property *F1* or *F2* or . . . *Fn*, where *F1-Fn* are the empirically possible causal bases of fragility.[90] The causal connection to breaking when struck no longer figures essentially in the definition of `fragility', but has instead been downgraded to the status of a handy reference-fixer that is used for establishing reference to the properties, *F1-Fn*. `Being fragile' now amounts to no more than having one of these first-order properties. Jackson's argument has no force against the causal efficacy of fragility thus construed. After all,

---

[90] The grade of possibility here must be empirical, since if we included in our disjunction properties that serve as causal bases of fragility in non-actual worlds in which the laws of nature differ from the actual world's, then any state could serve as a causal basis of fragility. The same would hold for all other dispositions, so that every disposition would be defined by a disjunction that enumerates every possible state, and all dispositions would be identified with just this one disjunctive property; i.e., there would be only one disposition.

given that the extensional definition contains no mention of the causal connection to breaking when struck, and given that none of the properties that it lists (*F1-Fn)* is itself essentially a typical cause of breaking when struck, it is hard to see why fragility should be thought to be so.

This view of dispositions is modelled on Jaegwon Kim's erstwhile view that each mental property is reducible to the disjunction of its empirically (or `physically') possible physical realisers.[91] Unfortunately, once this parallel is noticed, it becomes evident that even though the strategy of extensional definition effects a tighter connection between dispositions and their causally relevant first-order bases, dispositions thus conceived do not inherit the causal status of their first-order realisers. This is because such `wildly disjunctive' properties as the ones to which Kim would reduce mental features (and to which we have considered reducing dispositions) are not fit to appear in genuine causal laws and the causally significant counterfactuals that they support, and because appearance in such claims

---

[91] Jaegwon Kim, "Concepts of Supervenience," in his *Supervenience and Mind,* 53-78. See also Jaegwon Kim, "Supervenience as a Philosophical Concept," in *Supervenience and Mind,* 131-160 (esp. 151-4). Kim seems to have forsaken this view in favour of narrower, species-specific, type-type reductions (Jaegwon Kim, "Multiple Realisation and the Metaphysics of Reduction," in *Supervenience and Mind,* 309-35; see also Jaegwon Kim, "Reductive and Nonreductive Physicalism," chap. 9 in his *Philosophy of Mind* [Boulder, Colorado: Westview Press, Inc., 1996] esp. 218-21 and 233-6).

would appear to be at least necessary (though--as I maintain--not sufficient) for being causally relevant.

In saying this, one need not dispute the claim that disjunctions of properties, even infinitely long disjunctions, are themselves properties. The point is rather that even if they are properties, the possession of them in no way augments the causal powers of their bearers. For something can have the disjunctive property, *F1* or *F2* or . . . *Fn* (to which fragility has supposedly been reduced), only by having one of the properties that is cited in its disjuncts, say, the molecular structure *F1*. But, insofar as we accept this debate's underlying premiss that properties can be efficacious, any such thing would appear to have whatever causal tendencies it has only because it is *F1*. The fact that it also has the disjunctive property of being *F1* or *F2* or . . . *Fn* adds nothing to its causal powers.[92]

It might seem odd to deny causal efficacy to a disjunctive property in spite of its appearance in generalisations that support counterfactuals and predictions. Nevertheless, none of the generalisations in which a wildly disjunctive property appears can be taken to

---

[92] I follow here the argument given by Terence Horgan in his criticism of Kim (Terence Horgan, "From Supervenience to Superdupervenience," *Mind* 102 [1993]: 555-86 [at 576-7]). Essentially the same reasoning was given by D.M. Armstrong in D.M. Armstrong, *A Theory of Universals* (Cambridge: Cambridge University Press, 1978), 20. Armstrong there wrote, "Suppose that *a* has *P* but lacks *Q*. The predicate `*P* V *Q*' applies to *a*. Nevertheless, when *a* acts, it will surely act only in virtue of its being *P*. Its being *P* V *Q* will add no power to its arm."

assert a law-like, causal connection between events having the disjunctive property and anything else, for none of these claims is confirmable by its instances. The importance of confirmability as a test of lawhood has been emphasised by David Owens.[93] He notes that while genuine laws are susceptible of confirmation by their instances, generalisations that invoke wildly disjunctive properties in their antecedents and non-disjunctive features in their consequents are not. For example, consider the generalisation,

> (*P*) For all *x*, if *x* has *F1* or *F2* or . . . *Fn*, then *x* breaks when struck.

Suppose we observe something that has *F1* and that breaks when struck, and that thus provides an instance of (*P*). This instance does not count as evidence in favour of (*P*), since it equally supports the following `rival' of (*P*) (i.e., a claim that yields predictions that are contrary to those yielded by [*P*]):

---

[93] David Owens, "Disjunctive Laws?" 197-202; see also William Seager, "Disjunctive Laws and Supervenience," 93-8. Owens is concerned to show that wildly disjunctive properties do not appear in causal laws. By contrast, Seager focuses on laws that are formulated with a biconditional and that support reductions. By showing that wildly disjunctive properties are not fit to appear in reductive laws, he thereby undermines Kim's attempt to reduce mental features to wildly disjunctive physical properties. I will not pursue this strategy with respect to the putative reduction of dispositions to the wildly disjunctive enumerations of their causal bases, but will instead argue that even if they could be so reduced, they would not be causally relevant to any effects. By thus conceding, for the sake of argument, that dispositions are susceptible of such a reduction I hope to arrive at a negative verdict concerning recently proposed tests of causal relevance.

(*Q*) For all *x*, if *x* has *F1* then it breaks when struck, and if *x* has *F2* then it does not break when struck.

But if an instance of (*P*) gives us no more reason for believing (*P*) than one of its rivals, then it does not confirm (*P*). In short, while the fact that something which is *F1* is seen to break when struck certainly lends credence to the general claim that whatever is *F1* breaks when struck, it gives no indication as to the behaviour of things that are *F2*, and thus equally supports any generalisation that conjoins the claim that *F1*-things break when struck with any claim whatever concerning the behaviour of *F2*-things. To give equal support to all such generalisations is to confirm none of them.

Note that since any given instance of (*P*) will count as such only by virtue of instantiating one of the disjuncts in (*P*)'s antecedent (as well as the property described in [*P*]'s consequent), we can always in like manner construct a rival of (*P*) that is equally supported by that instance, thereby showing that the instance at hand does not confirm (*P*). Since (*P*) is not confirmed by any of its instances it is not a causal law, and the counterfactuals that it supports are not indicative of causal relations.

Ned Block disputes this result. He has argued that disjunctive properties, such as jadehood and dormitivity (to

use his examples), are projectible.[94] According to Block, "Given that secobarbitol and diazepam resemble one another in one way, in causing sleep, that gives us some reason to expect that they resemble one another in another real property."[95] For example, if we find that secobarbitol is carcinogenic, then (says Block) we will have reason to expect that diazepam is too.

It is hard to see why we should expect there to be this further resemblance between the two sedatives. Granted, they already resemble each other in respect of causing sleep. But this is only a resemblance in respect of an extrinsic feature (i.e., in respect of what sort of effect they produce in a certain type of organism). Why should we take this to be indicative of any further similarities? After all, two sedatives may differ greatly in their intrinsic natures and in the mechanisms by means of which they induce sleep in human beings--a pharmaceutical company does not corner the market on all sedatives merely by taking out a patent on secobarbitol. With this sort of objection in mind, Block stresses that the claim that diazepam is carcinogenic receives only a small degree of confirmation from our discovery that secobarbitol causes cancer.[96] But does it

---

[94] Ned Block, "Anti-Reductionism Slaps Back," *Philosophical Perspectives* 11 (1997): 107-32 (at 126-7).
[95] Block, "Anti-Reductionism Slaps Back," 127.
[96] Block, "Anti-Reductionism Slaps Back," 127.

receive even a small degree of confirmation? Block rests his case on the premiss that, "Any *real* resemblance makes *another* real resemblance a bit more likely."[97] But surely this is too strong. Secobarbitol resembles every other substance in *some* way or other (more so if we are to classify a resemblance in respect of extrinsic features as a "real resemblance"); and yet it would be extremely counterintuitive to conclude from this that the discovery of its carcinogeneity confirms (even to a small degree) the claim that all these other substances are carcinogenic too. Moreover, even if we were to follow Block in saying that it confirms this claim to some infinitesimally small degree, the degree of confirmation at issue is clearly insufficient to establish the causal relevance of the disjunctive property involved. To revert to an earlier example, the red brick that broke the window resembles the red curtain in respect of colour; if Block is right, this lends some small degree of confirmation to the claim that the curtain will break the window; but even if redness is thus projectible, this goes no way towards establishing the causal relevance of the brick's redness to its breaking the window.[98]

---

[97] Block, "Anti-Reductionism Slaps Back," 127.

[98] Block himself would be unperturbed by this result, since he wishes ultimately to sunder projectibility from causal relevance and takes Kim to task for not doing likewise (Block, "Anti-Reductionism Slaps Back," 129).

Whichever route we take, then, we must acknowledge that the wildly disjunctive properties which we have been considering are not causally relevant. It is interesting to note that in spite of their lack of any interesting sort of causal relevance, wildly disjunctive properties do nonetheless appear in counterfactuals. For instance, (*P*) supports the claim that if the paperweight had been *F1* or *F2* or . . . *Fn*, then it would have broken when struck. (*P*) is also closely bound up with the counterfactual claim that if the glass had not been *F1* or *F2* or . . . *Fn*, then it would not have broken when struck. Wildly disjunctive properties can appear in both these sorts of counterfactuals *without* thereby having their causal credentials authenticated, that is, without it being the case that they possess the sort of causal relevance that we take mental properties to have.

This result is interesting even if we do not attempt to reduce dispositions to the disjunctions of their causal bases, because regardless of what we think about dispositions, the fact remains that some wildly disjunctive properties, in spite of their causal irrelevance, satisfy LePore's and Loewer's conditions i-iv. They thus afford another counterexample (in addition to dispositions) to the counterfactual criterion of causal relevance.

VII. Too Remote To Matter

There seems then to be no advantage in identifying dispositions with the disjunctions of their first-order realisers. Thus conceived, they become too disjunctive to have any bearing on what sorts of events cause what sorts of effects. Indeed, David Lewis regards dispositions as being too disjunctive to matter even when we think of them as Jackson does, namely, as higher-order existential properties (e.g., as the property of having some or other feature that typically causes breaking when struck). According to Lewis, "The existential property, unlike the various bases, is too disjunctive and too extrinsic to occupy any causal role."[99] It is not clear whether Lewis is right in claiming that dispositions, even when conceived as existential properties, are too disjunctive to be causally relevant. However, he is right to regard them as being too extrinsic to matter.

What Lewis has in mind here can be better understood by considering what he has to say about a putatively higher-order property that is not a disposition: the property of losing heat. Lewis begins by asking us to suppose that

> Heat is whatever property it is that occupies a certain causal role. . . . So there are many different ways that the poker might lose heat, depending on what sort of world it is in. . . . Its molecular motion might decrease, in a world where molecular motion is what occupies the role;

[99] David Lewis, "Causal Explanation," in his *Philosophical Papers* (Oxford: Oxford University Press, 1986), 2:214-40 (at 224).

> or it might lose caloric fluid, in a world where caloric fluid is what occupies the role. . . . So it seems that losing heat is quite a disjunctive affair; and what's worse, extrinsic, since whether one property or another occupies the heat-role depends on what goes on throughout the world in question, not just on the region of it where the poker is.[100]

According to Lewis, this militates against the view that the poker's losing heat is a genuine event which causes the poker's contraction.

His reasoning is based on the assumption that causation is local. That is, what happens here and now depends on what happened here recently; other, more remote factors (e.g., factors that involve "what goes on throughout the world in question") cannot be the proximate cause of what happens here and now. Contemporary physics may offer counterexamples to this view,[101] but none of these examples suggests that something as sweeping as what happens throughout the entire actual world is causally relevant to determining what happens in a given place at a given time.

Lewis is willing to extend the same sort of reasoning to cover dispositions, such as fragility.[102] Roughly, a thing's being fragile is not sufficiently local to be

---

[100] David Lewis, "Events," in his *Philosophical Papers*, 2:241-69 (at 267).
[101] I have in mind here the Einstein-Podolsky-Rosen prediction, which was confirmed by John Stewart Bell.
[102] Lewis, "Events," 268.

causally relevant to anything happening here and now (e.g., to the vase's breaking).

Lewis's points are relevant to this discussion of LePore's and Loewer's proposed test of causal relevance; for LePore and Loewer offer their test as a criterion of the *efficacy* of *properties*, so that dispositions, since they pass this test, would in some sense have to be efficacious or causally potent. But how could they be, given that they are (as Lewis notes) extrinsic? In what follows, these points will be developed, although no claim will be made that Lewis's reasons for holding this view are the same as the reasons that are offered here.

Let us suppose that the vase has molecular configuration *C*, and that because of this it breaks when I drop it. Is it also true that it breaks because it is fragile (assuming that *C* is a causal basis of fragility)? No, not if it is really the case that what happens here and now directly depends only on local causally relevant factors, and not on things that are far away or in the remote past. For the vase's being fragile is intimately bound up with such factors. In fact, its being fragile consists partly in how things are in the future and in nearby non-actual worlds. This is because its being fragile is a matter of its state (in this case, *C*) being linked by

the laws of nature to breaking-when-struck.[103] That is, the laws of nature are part of what makes it true that the vase is fragile; they help to constitute this fact.[104] But insofar as the having of a disposition supervenes not only on local factors but also on causal laws, it supervenes on whatever it is that makes the causal laws true; thus, on the most plausible account of what makes causal laws true, it supervenes on how things are throughout the actual world and in nearby non-actual worlds. But if this is so, then the having of a disposition is much too broad to be causally relevant. Surely what happens here and now does not *causally* depend upon what happens throughout the actual world and in nearby *non*-actual worlds.

To clarify, it is not being suggested that relational features cannot be causally relevant. (Surely the fact that I am standing *in front of* the firing squad is causally relevant to what happens next.) Instead, it is being argued that the vase's being fragile is an extremely broad state of

---

[103] Again, while this may not be the sort of reasoning that Lewis has in mind, he would agree at least with this premiss of the argument. In a more recent paper he says, "Dispositions are an intrinsic matter. (Except perhaps in so far as they depend on the laws of nature. I myself would wish to insist on that exception. . . .)" (David Lewis, "Finkish Dispositions," *The Philosophical Quarterly* 47 [1997]: 143-58 [at 147-8]).

[104] D.M. Armstrong calls this a "Soft doctrine of powers." He says, "What we need is that the particular should have the property *F*, together with the totality of the relevant laws of nature. . . . These states of affairs . . . are sufficient as truthmakers for truths of unmanifested powers and dispositions. For the unmanifested power would appear to *supervene* upon these truthmakers" (D.M. Armstrong, *A World of States of Affairs* [Cambridge: Cambridge University Press, 1997], 82).

affairs, as broad as the scope of the laws that connect *C* to breaking when struck; and that it is thus too broad to exert any causal influence on the here-and-now.

It might be objected that the mere fact that the vase's being fragile depends upon causal laws should not weigh against its being causally potent; if it did, then nothing would be efficacious with respect to anything else, since the answer to any question about causal efficacy will of course depend (at least in part) upon causal laws. Note, however, that this dependence is not itself causal but rather conceptual. More specifically, the laws of nature help us to see which local states are efficacious with respect to which sorts of effect; so, to be sure, the laws of nature help us to discern the relations of efficacy; but what *gets* (or bears) the efficacy are the local states, not the laws of nature themselves. In short, the causal laws are relevant to the question of what is causally potent, but it is the local states that are said to *be* causally potent. But this would not be the case if we allowed the having of a disposition to be causally potent; for, since the having of a disposition consists (at least in part) in the causal laws being such as they are, this would render the laws themselves (along with the remote factors on which their

truth depends) causally efficacious with respect to the effect.

It may be questioned whether the having of a disposition really is extrinsic. To illustrate that it is, consider the same vase (or one of its counterparts) with the same intrinsic makeup, including *C*, in a world where the laws of nature are different. According to these other-worldly laws, the vase will not break if struck. In fact, the laws in this other world are such that almost nothing can break the vase. Is the other-worldly vase fragile? It is hard to see how it could be accurately characterised as such. And yet it has precisely the same intrinsic nature as the vase in the actual world. Therefore, being fragile is not part of the actual vase's intrinsic nature. It is instead an extremely broad feature of the vase, one that involves the causal laws that associate the vase's intrinsic nature with breaking-when-struck. Any state that is this broad is too extrinsic to matter.[105]

---

[105] The views considered in this section are clearly at odds with the position of Sydney Shoemaker and others, who identify every genuine property with whatever causal powers it confers upon its bearers. On this view, every property is dispositional, and instead of each disposition having a categorical basis, it's dispositions 'all the way down', so to speak. For Shoemaker's statement of this view, see Sydney Shoemaker, "Causality and Properties," in *Time and* Cause, ed. Peter van Inwagen (Dordrecht: D. Reidel Publishing Company, 1980), 109-35. See also Sydney Shoemaker, "Properties, Causation, and Projectibility," in *Applications of Inductive* Logic, ed. L. Jonathan Cohen and Mary Hesse (Oxford: Clarendon Press, 1980), 291-312. Chris Swoyer favours Shoemaker's view in Chris Swoyer, "The Nature of Natural Laws," *The Australasian Journal of Philosophy* 60 (1982): 203-23. If Shoemaker is right, then the reasoning presented in this section would imply that no properties are causally relevant. But Shoemaker is not right. I agree

A critic may doubt that the vase really would retain *C* in worlds in which the laws of nature are different. For, given that *C* is the vase's intrinsic molecular configuration, it is hard to see how *C* could be held constant whilst the laws of nature are varied to such an extent that the vase, with that intrinsic nature, remains virtually indestructible. Surely this derangement of the laws of nature would ramify through the laws of molecular bonding, so that the vase's internal molecular configuration would itself differ in these non-actual worlds.[106]

There are two ways in which this objection might be met. First, instead of considering possible worlds in which universal laws of nature differ from those of the actual world, consider worlds in which none of the laws have universal scope. In some such world that contains merely a patchwork of 'gappy' regularities, it may just be a brute fact that a vase with molecular configuration *C* (the inner structure of which is described by one set of generalisations) does not break when dropped or struck. Admittedly, we are here considering bizarre, remote worlds, but we are not, after all, doing so in order to settle questions about natural or nomological possibilities.

---

with the criticisms of his view that have been offered by Richard Swinburne in Richard Swinburne, "Properties, Causation, and Projectibility: Reply to Shoemaker," in *Applications of Inductive Logic*, ed. Cohen and Hesse, 313-20; and by Stephen Yablo in Yablo, "Mental Causation," 263-4 n. 39.

[106] My thanks to William Seager for raising this objection.

Instead, we are merely trying to decide a conceptual question (viz., whether dispositions are extrinsic). It seems permissible to venture further afield into bizarre and remote worlds when only conceptual issues are at stake.

Secondly, even if the objection succeeded when applied to fragility, there are other dispositions that can still be used to show that at least some dispositions are too broad to be causally efficacious with respect to their manifestations. For example, the opium's intrinsic nature can be held constant while varying other factors sufficiently so that the opium is no longer dormitive. This time, however, the variation in other factors need not involve changes to any causal laws. Instead, we need only change the human body's constitution just enough so that opium no longer triggers in it a series of reactions that culminates in sleep. These changes leave the opium untouched; it retains the same old intrinsic nature, but due to changes wrought elsewhere (viz., in the human body) it no longer counts as a dormitive agent.[107] Since we can thus take away opium's dormitivity merely by changing things that are external to opium, the dormitivity must itself be a very

[107] This is an interesting, but seldom noted, difference between dispositions like fragility and those like dormitivity. The difference seems to arise from the fact that in the case of dormitivity, but not in the case of fragility, the manifestation of the disposition lies outside the bearer of the disposition, so that it is possible to take the disposition away from its bearer by making changes in the locus of the disposition's manifestation rather than in its bearer.

broad feature of opium, one that is too broad to be efficacious.

The foregoing considerations provide an additional reason for denying that (at least some) dispositions are causally potent, even though they pass the counterfactual test. They also suggest more directly that mental states, as conceived by the functionalist, are too broad to be causally efficacious with respect to anything. The point is not limited to content; it is not the old worry about externalism--that is to say, the old worry that if content is broad (in the ways that Hilary Putnam and Tyler Burge have suggested), then it is hard to see how it can contribute to my arm's raising. It is instead a worry that pertains to all mental states insofar as we adopt the functionalist's account of them. For if mental properties are higher-order functional states that consist in the having of lower-order features that realise a certain causal role, then they too (like dispositions) would seem to embrace not only the lower-order intrinsic features that realise them, but also the laws that connect those lower-order intrinsic features to their causes and effects.

VIII. What We Really Want

I have thus far criticised those attempts to overcome the threat posed by epiphenomenalism that rely on an appeal

to a counterfactual test of causal relevance (taking LePore's and Loewer's account as the clearest example of such attempts). My criticisms have taken two forms. First, I have claimed that even if mental properties pass the counterfactual test, they may be (and--according to functionalism--are), like dispositions, higher-order properties that have at best only a weak form of causal-explanatory relevance, and that thus come out looking causally deficient in some crucial respect, since it seems that only the first-order realising properties that serve as their causal bases enjoy the full-fledged efficacy that the participants in this debate have in mind. Secondly, I have argued that dispositions and certain wildly disjunctive properties are not causally relevant in any rich or interesting sense, but do nonetheless pass the counterfactual test, and are therefore counterexamples to LePore's and Loewer's claim that this test is a test of the causal relevance that we take mental properties to have.

I would like now to ask why it is that counterfactual criteria fail as genuine tests of the sort of causal relevance that we regard mental features as having. In a nutshell, the answer is that passing a counterfactual test at best only shows that the property in question is necessary for the prediction of a given type of effect to be

warranted; it does not show that having that property is alone sufficient to yield that prediction.[108] Indeed, it appears that property instances that clearly are sufficient in this sense, and that therefore clearly are causally relevant to it, actually fail the counterfactual test.

These points can be illustrated by means of an example that was proposed by Douglas Ehring. Ehring asks us to consider "a colour detector that activates at the presence of a certain precise shade of red, scarlet."[109] According to Ehring, the detector does not activate in the presence of any other shade of red. In this case, the counterfactual criterion yields the result that the property "being scarlet" is causally relevant to the detector's activation; for if the item that is being scanned by the detector (say, a chair) had not been scarlet, then the machine would not have been activated. But the counterfactual criterion also yields the result that the property of being red is causally relevant to the activation, for it is also true that if the chair had not been red, the detector would not have been

---

[108] For simplicity, I shall hereafter call a property's sufficiency to support the prediction of an effect (strongly enough for the prediction to be justified) its *causal sufficiency* for that effect; similarly, a property's mere necessity for the justification of a prediction will be its *causal necessity* for the given effect. Also, it is assumed that in each of the following examples, we have already observed whether each of the properties to be considered is in fact followed by the effect whose prediction is in question; so that some of the properties to be considered do not (given our background knowledge) justify the relevant prediction.

[109] Douglas Ehring, "Mental Causation, Determinables and Property Instances," *Nous* 30 (1996): 461-80 (at 466).

activated. The same can be said as we move further up the chain of determinables away from the determinate shade, scarlet, so that even the property of being coloured turns out to be causally relevant to the machine's activation: if the chair had not been coloured, the device would not have been activated. Now while these increasingly determinable features (being red, etc.) are *in some sense* causally relevant to the machine's activation, it is only in the very weak sense of being required or causally necessary for that effect. Ehring apparently takes this to be the mark of the sort of causal relevance that is at issue in the mental causation debate, but it is important to see that it is not what we have in mind when we seek to vindicate the causal relevance of the mental. After all, even though "being red" is causally required in order to activate the detector, it is not sufficient to do so; if the chair had been crimson, the detector would not have been activated. When we set out to vindicate the causal relevance of the mental, we have in mind something stronger than the weak form of causal relevance that is shared by the properties of being red or being coloured in this example. We want to defend the claim that my desire for a Coke, coupled with my belief that there is Coke in the fridge, is in itself sufficient to support the claim that I will open the fridge.

One problem, then, with a counterfactual test of causal relevance (such as the one proposed by LePore and Loewer) is that it tracks a form of relevance that is weaker than what we seek for mental properties, for it only tracks properties that are required for a certain outcome rather than properties that are sufficient with respect to it. Worse still, the counterfactual test takes us *away* from causal sufficiency and towards *mere* causal necessity in the added sense that many causally sufficient properties *fail* to satisfy it. To see that this is so, consider the case in which I tip the scales because I weigh 150 pounds, and in which anything that weighs more than 120 pounds also tips the scales. Does my property of weighing 150 pounds pass the counterfactual test? Apparently not, for it seems to be false that if I had not weighed 150 pounds, I would not have tipped the scales. This is because it seems likely that in the nearest non-actual worlds in which I do not weigh 150 pounds I instead weigh 151 pounds, or some similar weight, which is equally sufficient to tip the scales. Or consider the claim that the pattern of neuronal firing in my brain caused my arm to rise. Is it the case that if just *that* actual, fully determinate pattern had not been realised in my brain, then my arm would not have risen? Again, it seems that the answer must be "no", for in the nearest non-actual

worlds in which my neurons do not realise precisely that pattern, they instead realise an almost exactly similar pattern (with the difference affecting only one or two brain cells) that is also sufficient to cause my arm to rise. In short, if we were to adopt the counterfactual test of causal relevance, we would arrive at the conclusion that few (if any) of the fully determinate features that are actually on hand to herald the arrival of an effect are in fact relevant to it, since they are not actually required for the production of that effect, given that other, similar properties would equally have sufficed for it. And yet, surely, the fully determinate properties that are actually sufficient with respect to an effect are in some sense causally *relevant* to it. This is the sort of causal relevance that we regard mental properties as having. To the extent that counterfactual criteria fail to track this sort of causal relevance, they fail to address the question that we have in mind when we inquire about the causal relevance of mental properties.

It might be objected that the counterfactual test proffered by LePore and Loewer can readily be supplemented in a way that captures the stronger kind of causal relevance that we seek for mental features. One could simply say that a property is causally relevant to a given effect if it

passes the counterfactual test, or if some determinable of it does. This won't work, though, because it would render burgundy causally relevant to the detector's activation, since one of this shade's determinables (viz., red) passes the counterfactual test: if the chair (which is in fact scarlet) had not been red, the detector would not have been activated. But burgundy is clearly not causally relevant to the activation of the machine. This difficulty cannot be remedied by stipulating that the property that is being assessed for causal relevance must at least be one that the chair actually has, for the same problem arises if we suppose that the chair is part burgundy and part scarlet. We want a test of causal relevance according to which the detector is activated because the two-tone chair is scarlet, not because it is burgundy. The amended counterfactual criterion that we are considering does not yield this result.

It might also be thought that the two cases described above (involving the fully determinate pattern of neuronal firing in my brain and my weighing 150 pounds respectively) are just examples of the sort of `screening off' worry that LePore and Loewer address.[110] To say that one feature (*P*) `screens off' another property (*M*) is to say that the cause in question has both *P* and *M*, and that even if it had lacked

[110] LePore and Loewer, "Mind Matters," 638-40.

M but retained *P*, it still would have produced the same effect that it actually produced. LePore and Loewer entertain the proposal that we add as a fifth condition on causal relevance$_2$ the requirement that no property that is screened off from a given type of effect is relevant$_2$ to it. They reject this proposal on the grounds that it would render causally irrelevant certain properties that obviously are relevant.

It should be evident that my above argument does not simply recapitulate the `screening off' worries that LePore and Loewer have already addressed. For notice that in order for there to be a genuine case of `screening off', the cause must *actually* possess both the screening off and the screened off properties. However, in the case in which I tip the scales by weighing 150 pounds, the property that `takes over' (so to speak) in the nearest non-actual world (in which I do not weigh 150 pounds) and guarantees the same effect is *not* one that I actually possess. For it is none other than the property of weighing 151 pounds, which I do not possess in the actual world.

Perhaps it will be objected that my argument is, nonetheless, just the same old screening off worry (that LePore and Loewer have already considered) in disguise. For while I may not actually weigh 151 pounds, I do actually

possess the determinable feature of weighing more than 120 pounds; and (the critic may add) it is this determinable property that, in the above example, screens off my property of weighing 150 pounds from being causally relevant to my tipping the scales. It might thus appear that my argument really does only amount to the screening off worry that LePore and Loewer have already taken into account, with the added twist that in my example we have a determinable feature screening off one of its own determinates from the effect.

To see why this is not so, recall that in the sort of case that LePore and Loewer discuss, we are to consider the claim that the neurological properties that realise a given mental feature would have produced the *same* effect (e.g., would have caused my arm to raise) even if they had not been accompanied by that mental property, and thus screen off the mental property. It is important to note that the counterfactual at issue here is *not* simply of the form of LePore's and Loewer's crucial third condition on relevance$_2$. The counterfactual in their third condition had the following form:

(*a*) $-Mc > -Be$.

In the case that LePore and Loewer discuss, the counterfactual that has this form is, "If I had not been in

mental state *M*, my arm would not have risen." In their view, this counterfactual is true (and this is what--in their opinion--establishes the causal relevance of the mental), for in the nearest non-actual worlds in which I am not in state *M* my arm does not raise. Note, however, that in those nearby non-actual worlds in which I am not in *M* I also do not have the neurological properties that realised *M* in the actual world; for *M* strongly supervenes on those neurological features, from which it follows that in any nearby world in which I lack *M*, I also lack those neurological properties (and any other features on which *M* strongly supervenes).[111] The point of all this is simply that in the case that LePore and Loewer consider, in which we ask what would happen if I had those same neurological features *without* having the mental property *M*, we cannot simply be asking what happens in the *nearest* non-actual worlds in which I lack *M*; for, given the strong supervenience of *M* on those neurological features, the nearest non-actual worlds in which I lack *M* are precisely worlds in which I also lack

---

[111] A very similar point was made by A.C. Ewing in response to the claim that epiphenomenalism implies that a zombie (who is devoid of consciousness) could have written *Hamlet*. Ewing counters that, "Even on [the epiphenomenalist's] view, since mental processes inevitably result when the physiological brain-processes have attained a sufficient degree of complexity, and very complex brain-processes are undoubtedly needed for the production of such works, they could not be produced without complex mental processes also" (A.C. Ewing, *Value and Reality* [London: George Allen & Unwin Ltd., 1973], 77-8). Unlike LePore and Loewer, Ewing does not take this to establish the efficacy of mental processes, and hence regards it as being part of *the epiphenomenalist's* position.

those neurological features; so we must be going further afield, to more distant non-actual worlds, until we reach one that is (so to speak) beyond the orbit of worlds that define the actual world's strong supervenience relations, a world in which those same neurological properties *do* appear without being accompanied by *M*.[112] In this sort of case, then, we are not simply evaluating a counterfactual of the form of (*a*). Instead, the counterfactual at issue has the form

(*b*) $(-Mc \ \& \ Pc) > Be$.[113]

As has been noted, LePore and Loewer reject the proposal that any property that can be substituted for *M* in (*b*) (i.e., any property that is similarly screened off from *e*'s being *B*) should be barred from being causally relevant$_2$ to *e*'s being *B*.

It should be clear though that in the sort of case that was described in my earlier argument, I was not illicitly (contrary to LePore's and Loewer's wishes) treating the

---

[112] This is what David Lewis has in mind when he speaks of "a logical peculiarity of counterfactuals: their `variable strictness'. It can happen that two counterfactuals

> If it were that *p*, it would be that not-*q*
> If it were that *p* and *q*, it would be that *r*

are true together, and that the truth of the second is not merely vacuous truth. Because the first counterfactual is true, the supposition that *p* and *q* is more far-fetched, more `remote from actuality', than the supposition just that *p*. But we are not forbidden to entertain a supposition merely because it is comparatively far-fetched. Variable strictness means that some entertainable suppositions are more far-fetched than others" (Lewis, "Finkish Dispositions," 150).

[113] This is (S) in LePore and Loewer, "Mind Matters," 639.

truth of (*b*) as a bar to *M*'s causal relevance. For in that argument we really are confining our attention to the *nearest* non-actual worlds in which I lack the property of weighing 150 pounds. Thus, we really *are* evaluating a counterfactual that has the form of (*a*) (i.e., the form of the counterfactual in LePore's and Loewer's crucial third condition on causal relevance$_2$), for we are simply considering what would have happened if I had lacked the property of weighing 150 pounds; and it seems quite plausible that in the nearest worlds in which I lack this property, I have some very similar property (e.g., weighing 151 pounds) that `takes over' and guarantees the same effect that was produced in the actual world by my weighing 150 pounds. As a result, the counterfactual, "If I had not weighed 150 pounds, I would not have tipped the scales," is false. So if we accept LePore's and Loewer's counterfactual test of causal relevance, then my weighing 150 pounds is not causally relevant to my tipping the scales.

Note that this reasoning can be generalised to apply to many other fully determinate features, for it derives from the manner in which determinables often screen off their determinates from a given effect. More specifically, it seems that if we substitute a determinate predicate for "*M*" in (*a*) and (*b*) and a determinable predicate for "*P*" in (*b*),

then (usually) (*a*) and (*b*) are equivalent, for their antecedents come out true in the same worlds; that is to say, the nearest worlds in which the antecedent of (*b*) is true just *are* the nearest worlds in which (*a*)'s antecedent is true; so that, unlike the screening off case that LePore and Loewer consider, we need not venture further afield to more distant worlds in evaluating (*b*) than we do in evaluating (*a*). But then, since the counterfactual that we are considering is equivalent to one that has the form of (*a*), and since LePore and Loewer accept that sort of counterfactual as their crucial third condition on causal relevance$_2$, they cannot avoid the conclusion that many fully determinate properties (e.g., the determinate pattern of neuronal firing in my brain) are not causally relevant. This is a compelling reason for abandoning their test, if we accept (as we surely must) that such determinate properties as the actual pattern of neuronal firing in my brain are relevant to the production of certain effects, and are causally relevant in the strong sense of being causally sufficient for them.

3. Still Looking . . .

> For example, a particular point of grammar is present in the soul, which is the subject, but is not said of any subject, and a particular whiteness is present in a body (for every colour is in a body), which is the subject, but is not said of any subject. . . . And without qualification, that which is an individual and numerically one is not said of any subject, but nothing prevents some of them from being present in a subject; for a particular point of grammar is present in a subject but is not said of any subject.
>
> Aristotle[114]

I. Relevance Reclaimed?

In the second chapter David Lewis's writings were cited in support of the view that dispositions are causally irrelevant. Lewis has never seemed entirely pleased with this result. Indeed, he has called it a "disagreeable oddity"[115] that must be dispatched if the identification of dispositions with second-order properties is to win our unequivocal support. In a recent paper, he takes himself to have done just that.[116] He begins by saying that, "Sometimes, an event . . . is a having of a certain property by a certain thing"; and sometimes, he continues, "Two different properties are had in the same single event."[117] Consider, for instance, the event that consists in the "having of the

---

[114] Aristotle, *Categories, chap. 2*, in *Aristotle: Selected Works*, trans. Hippocrates G. Apostle and Lloyd P. Gerson (Grinnell, Iowa: The Peripatetic Press, 1982), 29-30. My thanks to Ronald de Sousa for calling this passage to my attention.
[115] Lewis, "Causal Explanation," 224. Lewis seems more willing to affirm the inefficacy of dispositions in his later paper (Lewis, "Events," 268).
[116] Lewis, "Finkish Dispositions," 151-2.
[117] Lewis, "Finkish Dispositions," 152.

[first-order] causal basis"[118] of the glass's disposition to break when struck. According to Lewis, "This same event is a having of the second-order property," (viz., fragility).[119] That is, the glass's possession of the molecular structure which serves as the causal basis of its fragility is all there is to its being fragile. The `havings' of these two properties are one and the same entity. Thus, since the glass's possession of that molecular structure is a cause of its breaking, so too is its fragility.

Cynthia and Graham Macdonald have adopted a similar strategy to vindicate the causal relevance of the mental. The Macdonalds couch their discussion in terms of "property instances", but their property instances are the same sorts of entities as Lewis's events. According to the Macdonalds, "Different properties may be instantiated in the same `individual property'",[120] where an `individual property' is an instance of an attribute by a particular thing. Consider, for example, a scarlet sweater. Its being red is not something over and above its being scarlet. Rather, its possession of the property of being scarlet is all there is to its being red.[121] In this particular case, the `havings'

---

[118] Lewis, "Finkish Dispositions," 152.
[119] Lewis, "Finkish Dispositions," 152.
[120] Cynthia Macdonald and Graham Macdonald, "Mental Causes and Explanation of Action," in *Mind, Causation and Action*, ed. L. Stevenson, R. Squires and J. Haldane (Oxford: Basil Blackwell, 1986), 35-48 (at 38).
[121] In the Macdonalds' words, "To be an exemplification of the former just is, in this case, to be an exemplification of the latter, despite

of these two properties are one and the same entity. Thus, the two different properties, *being red* and *being scarlet*, are had in the same single property instance.

Suppose now that the sweater's being scarlet is causally relevant to the activation of a colour detector. According to the Macdonalds, once we see that the sweater's being scarlet is causally relevant to the detector's activation we must acknowledge that so too is its being red. After all, since the sweater's being red and the sweater's being scarlet are one and the same property instance, its being red is causally relevant whenever its being scarlet is.[122]

It is the Macdonalds' view that the connection between mental and physical phenomena is in some ways parallel to the relation that determinables (such as red) bear to their determinates. More specifically, they maintain that even though mental properties cannot be reduced to physical properties, each mental property instance is identical with a physical property instance, and that mental property instances are thus efficacious with respect to behavioural effects whenever the physical property instances with which

---

the distinctness of the properties themselves" (Macdonald and Macdonald, "Mental Causes and Explanation of Action," 39).

[122] As the Macdonalds put it, "Any causally efficacious case in which a more determinate form of that property [viz., colour] is exemplified is a case in which the exemplification of colour *itself* is efficacious, by the extensionality of the causal relation" (Macdonald and Macdonald, "Mental Causes and Explanation of Action," 39). (Emphasis in the original)

they are identical are efficacious. As the Macdonalds put it,

> An instance of the property, being a brain event *B*, can be an instance of the mental property, being a pain. Moreover, if an instance of the former is causally efficacious, then so is the latter.[123]

This approach indicates an initially promising strategy for establishing the causal relevance of properties whose causal significance has been impugned. Lewis applies it to dispositions. The Macdonalds' work illustrates that it can equally well be applied to any suspect property (e.g., determinables and mental features), so long as that property is realised by or supervenes upon basic physical features whose causal relevance is not in question.

---

[123] Macdonald and Macdonald, "Mental Causes and Explanation of Action," 39. It is unclear whether, at the time at which their paper was written, the Macdonalds regarded the relation between mental and physical properties as a species of the determinable-determinate relation. In the passage from which the above quote was taken, they write, "Mental properties correlate in a one-many way with physical properties (though in no systematic way), with the consequence that any instance of the former is an instance of one or another of *some more determinate physical property. Just as* to be red is to be coloured, one might say, to be an instance of the property, being a brain event *B*, is to be an instance of the mental property, being a pain." (Emphasis added) The parenthetic denial of a "systematic" correlation between mental and physical properties militates against the claim that physical properties are determinates of mental features. However, the text that I have emphasised suggests the contrary view. This discrepancy is remedied in a later paper, in which the Macdonalds repudiate the claim that mental features have physical properties as their determinates (Cynthia Macdonald and Graham Macdonald, "How to be Psychologically Relevant," in *Philosophy of Psychology*, vol. 1 of *Debates on Psychological Explanation*, ed. Cynthia Macdonald and Graham Macdonald [Oxford: Basil Blackwell, 1995], 60-77 [at 74 n. 10]). They nonetheless continue to believe that mental and physical properties share the same property instances.

II. The Merits of the Macdonalds' View

Before reaching a verdict on the Macdonalds' approach it will help to clarify the issues if we pause briefly to compare their position with the counterfactual strategy that was examined in the second chapter. Unlike LePore and Loewer, the Macdonalds do not begin by formulating a criterion of causal relevance and by then attempting to show that mental features meet that standard. Instead, they try to demonstrate that the physical properties that intuitively are causally relevant are closely enough wedded to mental states for the latter properties (more accurately, their instances) to partake in the efficacy of the physical features.

This approach is similar to a strategy that was considered in the second chapter, namely, the attempt to confer causal relevance upon dispositions by identifying them with the disjunctions of their causal bases. Both strategies aim to bring the properties whose causal status is in doubt (e.g., dispositions and mental properties) closer to their obviously causally relevant physical realisers, in the hope that they can then inherit the efficacy of their physical bases. However, the two strategies are importantly different. For while the view that was considered in the second chapter proceeded at the

level of *properties* and the identities that may or may not hold among them, the Macdonalds' view imposes no restrictions on the relations that may or may not obtain among the properties involved. Instead, they limit their claims to identity relations that obtain among *instances* of properties. It is thus open to a Davidsonian to take this route, for it is compatible with the anomalous monist's rejection of systematic connections among mental and physical types.[124]

An additional novelty of the Macdonalds' approach derives from its broader scope: it is more ambitious in the range of worries that it aims to combat. To see why this is so, recall that the contemporary debate concerning the problem of mental causation arose within the context of criticisms of Davidson's anomalous monism. The worry was not simply that in the world depicted by anomalous monism, mental features turn out, as a matter of fact, to be inefficacious. Instead, the concern was that if anomalous monism is true, then it is not clear how mental properties even *could* be causally relevant to any effect. Given their absence from strict laws of nature, it seems that they are just not the *sorts* of properties that could (even in

[124] Indeed, in their 1986 paper the Macdonalds were specifically concerned to arrive at a demonstration of the causal relevance of mental features consistently with their own endorsement of anomalous monism (Cynthia Macdonald and Graham Macdonald, "Mental Causes and Explanation of Action,"). It is doubtful, though, that Davidson himself would be sympathetic to their approach.

principle) enjoy any causal significance. Accordingly, those who grapple with the problem as it arises in the context of anomalous monism begin by proposing a general test of causal relevance, which they then claim is met by mental features. It may be objected, however, that while they may thus show that mental properties are at least fit for causal relevance, that is, that there is nothing in the very nature of such properties that prevents them from being efficacious, it remains to be seen whether they are in fact causally relevant to any actual event. One might, for example, criticise the counterfactual test of efficacy on the grounds that it does not suffice for causal relevance, since the putative efficacy of the mental properties that satisfy this test may yet be precluded by certain general features of the actual world. For instance, it might be precluded by the (alleged) causal closure of the physical realm, or by the causal completeness of the physical world together with the absence of pervasive overdetermination.

Jaegwon Kim has directed this sort of criticism at LePore and Loewer (among others).[125] According to Kim, if mental properties are distinct from physical properties (as non-reductive physicalists maintain), and if the physical realm is causally closed, then it seems that mental

---

[125] Jaegwon Kim, "Explanatory Exclusion and the Problem of Mental Causation," in *Information, Semantics and Epistemology*, ed. Enrique Villanueva (Oxford: Basil Blackwell, 1990), 36-56 (at 43-5).

properties are not causally relevant to any physical effects; for, according to the principle of closure, only physical events and properties contribute causally to the production of physical events, and, according to the principle of distinctness, mental properties are not physical properties. It might be felt that the principle of closure is too strong, and should be replaced by the more modest principle of completeness, according to which for every physical event, there are physical events and properties that were sufficient to produce it, or at least to fix the probability of its occurrence. This more modest principle allows that there might indeed be other, non-physical factors in a physical event's causal history, as long as they were merely overdetermining causes that did not bring about any result (or yield any probability of an outcome) that was not already fixed by the purely physical elements in the causal chain. Unfortunately, this possibility is not a promising basis for an account of mental causation. For even if overdetermination is possible, it is surely not as pervasive as it would need to be in order for every human action to be an effect both of physical and mental antecedents. Thus we face the problem of according causal potency to the mental in a world in which

all the causal work has apparently already been done by physical states.

When confronting this worry it is no help to appeal to a putative test of causal relevance (e.g., the counterfactual or the nomological tests) and claim that mental properties pass that test, for, whatever test we adopt, it will (at least) equally well be satisfied by physical features. We will thus be thrown back into the dilemma of either according causal relevance exclusively to one set of properties (either to the mental or to the physical features) or countenancing the rampant overdetermination of human actions by both mental and physical states. If we affirm the causal closure of the physical realm, or at least its causal completeness together with the denial of rampant overdetermination, then the ineluctable consequence is the impotence of the mental.

Unlike the counterfactual strategy, the Macdonalds' approach does at least purport to allay these fears. For once we make property *instances* (rather than properties themselves) the units of causal relevance, we render unproblematic the ascription of causal relevance to mental states, so long as those states are held to be identical with physical tokens. This is not tantamount to the postulation of outside influences operating on the physical

system, for the mental instances that are said to be efficacious just are physical entities. Moreover, it does not require overdetermination, since in any given case there is only *one thing* (one property instance) that is said to be operative in bringing about the effect. Granted, this single property instance is describable in a plurality of ways (e.g., as an instance of a physical property and as an instance of a mental feature), but that should not be taken to imply that there is a plurality *of entities* (more specifically, property instances) acting on this occasion.

It is in its response to the problem of overdetermination that the chief virtue of the Macdonalds' and Lewis's approach is thought to lie.[126] The worry was that the ascription of causal relevance to suspect properties (such as mental and dispositional features) implicates them in an implausibly ubiquitous overdetermination. The response is that by making the units of causal relevance property instances (or Lewis's "events") and by identifying all such instances with instances of physical features, we can assign causal relevance to any property instance under any

[126] It is interesting to note that in his most recent discussion of the issue, Lewis summarizes the positive case for the impotence of dispositions exclusively in terms of an argument from overdetermination (Lewis, "Finkish Dispositions," 152); and yet, in his earlier papers (Lewis, "Causal Explanation," 224 and Lewis, "Events," 268), he does not even mention overdetermination when presenting the reasons for taking dispositions to be everywhere idle.

description while retaining an ontology and causal story that is as austere as we like.

III. The Identity of *Havings*

Douglas Ehring has objected that the Macdonalds overlook one of the conditions for the identity of property instances construed as exemplifications of universals.[127] To wit, they have failed to see that exemplifications of universals are identical only if they are exemplifications of the same property. After all, a property exemplification is simply a having of a property by a given object at a given time. All three of these items (the property, the object, and the time) are essential "components" of the property instance and should figure in its identity conditions.[128] So, contrary to what the Macdonalds claim, the sweater's being red is *not* the same property instance as its being scarlet. Ehring concludes that if we wish to regard an instance of a determinable as being identical with an instance of one of its determinates, and as inheriting the latter's efficacy, then we shall have to forsake the ontology of particulars exemplifying universals in favour of the nominalist's tropes (and resemblance classes thereof) in

---

[127] Ehring, "Mental Causation, Determinables and Property Instances," 463.

[128] "Exemplifications will have various `components' including universals and it is hard to see how exemplifications with different universal `components' could be identical" (Ehring, "Mental Causation, Determinables and Property Instances," 463).

order to find a suitable metaphysical grounding for our claims.

While we can agree that the properties that are instanced should figure in the identity conditions of exemplifications, it is not clear why Ehring believes that a property exemplification must be an instance of only *one* property. Why can't we just opt for more coarse-grained property instances, each of which is an instance of more than one feature? We should of course acknowledge that if *a* and *b* have different components, then *a* is not identical with *b*. But this does not preclude there being *one* property instance that is at once an exemplification of both redness and scarlet, so that *it* (that one exemplification) has both of these "component" properties among its identity conditions. Such an exemplification is not simply a redness-instance or a scarlet-instance. It is rather a redness-and-scarlet-instance; nothing could be it without being an instance of *both* of those properties.[129]

There is some intuitive support for these coarser-grained property instances. There is clearly a sense in

---

[129] Tim Crane takes this to be what distinguishes the Macdonalds' property instances from facts (on at least some conceptions of facts); thus, while the fact that I am in pain at *t* is different from the fact that I am in brain state *B* at *t* (assuming that *being in pain* is not the same property as *B*), "What the Macdonalds mean is that a single property instance has as `components' a mental property and a physical property" (Crane, "The Mental Causation Debate," 222). As will soon be explained, this is a misinterpretation of the Macdonalds if by calling both properties "components" Crane means that they are both constitutive of the event in question.

which what makes the sweater red is *not* the very same thing as what makes it size eight or a turtleneck, but *is* the same thing as what makes it scarlet. Its being scarlet is all there is to its being red, but is not all there is to its being size eight. Moreover, this manner of speaking is not restricted to standard examples of the determinate-determinable relation, for we are inclined to say similar things about dispositions and their realisations.

Ehring's objection can be met by adopting this strategy. As it happens, though, the Macdonalds are unlikely to take this approach, for it is inconsistent with what Cynthia Macdonald has said in anticipation of Ehring's criticism.[130] According to her, we must distinguish "between constitutive and characterising properties . . . of events."[131] While it is true that every event is a property instancing, and thus has as one of its constitutive (i.e., essential) components a property of which it is an instancing, it does not follow that every property figures as an essential component in the events which serve as its instantiations. That is to say, not all of the properties that are exemplified in a given property instance figure in the identity conditions of that instance; some of them are inessential to that event. According to Macdonald, any given

---

[130] Cynthia Macdonald, *Mind-Body Identity Theories* (London: Routledge, 1989), 143-55.
[131] C. Macdonald, *Mind-Body Identity Theories*, 147.

mental property will be inessential to the events that are its instantiations.[132] Thus, such events are merely characterised, but not constituted, by mental features.

It is interesting to note that Lewis has similarly downgraded dispositional and mental properties to the status of inessential aspects of their instantiations.[133] His reason for doing so is that all such properties are definable in terms of causal roles. If we assume that one and the same event could have occupied any number of different causal roles, then the occupation of any given causal role is inessential to that event. Hence, if mental and dispositional features are to be conceived of in terms of causal roles, then any event that is a having of a mental or dispositional property could have occurred without being a having of that property.[134]

It must be concluded that both Lewis and the Macdonalds have ready answers to Ehring's criticism, and that even if one finds their answers implausible (because it does not

---

[132] As she says, "The view that mental properties of persons are constitutive of the events that are exemplifyings of them (hence that mental properties of events are essences of them) is at best dubious and arguably false on the view of essences favoured by many" (C. Macdonald, *Mind-Body Identity Theories*, 152).

[133] Lewis, "Events," 268.

[134] In Lewis's words, "There is a genuine event which is accidentally classifiable in terms of fragility; essentially, however, it is a possession of such-and-such molecular structure. . . . And if I am right to think that mental states are definable as occupants of causal roles, then no genuine event is essentially classifiable as my being in pain. There are pain events, no doubt of it; but they are pain events only accidentally. . . . Essentially, the events are firings of neurons" (Lewis, "Events," 268). It should be noted that these are not Macdonald's reasons for holding this view.

seem that pain, e.g., could be anything less than essential to its instantiations), one can reply to Ehring in the manner described at the beginning of this section. It seems then that the identification of mental and physical property instances is a viable metaphysical option. At least, it cannot be ruled out on the basis of Ehring's criticism.

IV. An Equivocation

The crux of the Macdonalds' strategy is their claim that causal relevance is a relation between particulars (viz., property instances) rather than types. Setting aside for the moment the question whether this claim is true, it seems that the Macdonalds' defence of it involves an equivocation by means of which they draw conclusions about causal relevance from a consideration of the nature of causality. The equivocation seems clearest in the following passage:

> If we do insist that *causality is a relation between token events*, and that it is instances of properties associated with event types which are causally efficacious, then the "Principle of the Nomological Character of Causally-Relevant Properties" should be amended so as to finish ". . . Causally-Relevant Instances of Properties".[135]

Their point is that instead of taking causal relevance to be a relation between properties, we ought to regard it as a

---

[135] Macdonald and Macdonald, "Mental Causes and Explanation of Action," 37. (Emphasis added)

relation between the tokens of those properties. Clearly, though, the (emphasised) premiss that is supposed to motivate this view is a claim only about *causality*. No conclusions about *causal relevance* can be drawn without further argument--unless, of course, the Macdonalds take the relations of causality and causal relevance to be the same relation; but in that case we should simply rest content with Davidson's anomalous monism, for the problem of the causal relevance of the mental would simply not arise.

John Heil sketches a view similar to that of the Macdonalds. Unfortunately, his discussion, like theirs', involves a continual running together of causal relevance and causation. This is perhaps most evident when Heil tells us that,

> Discussions of mental causation are especially prone to type-token confusions. *In considering events as participants in causal transactions*, for instance, we are concerned, not with types of event, but with token events, dated, non-repeatable, particular occurrences.[136]

Again, it is clear from the emphasised text that Heil is speaking about events standing in the relation of causation. However, given the context,[137] he seems to believe that he is

---

[136] Heil, *The Nature of True Minds*, 136. (Emphasis added) There is also an aside in which Heil says, "For simplicity, I shall follow Searle and speak here of properties or characteristics causing and being caused, though, strictly, it is *instances* of properties or characteristics that have aetiological significance" (Heil, *The Nature of True Minds*, 127 n. 22). (Emphasis in the original)

[137] In light of what Heil says earlier in his chapter on the problem of mental causation, it is clear that he takes the problem to be the

not simply making an uncontroversial point about that relation, but is instead arriving at a more ambitious and interesting result about the nature of the *causal relevance* relation, to wit, that it too must be a relation between particulars rather than types. Heil characterises these particulars as "property instances or exemplifications--Plato's `moving forms', D.C. Williams's `tropes'."[138] While not identifying mental and physical tropes with each other, Heil does regard each mental trope as being "realised" by a physical property instance, where this means that the latter instance constitutes the former one and that the mental property (considered as a type) supervenes on the physical property.[139]

Whatever the merits of this account might be, it must be said that Heil's defence of it involves the same sort of legerdemain that we saw in the Macdonalds' reasoning, by means of which quite legitimate points about causation are subtly transferred to the different (and so far mysterious) relation of causal relevance. Moreover, it cannot be the case that Heil simply takes the relations of causation and causal relevance to be one and the same, for he agrees that

---

question whether mental properties are causally relevant, and not merely the question whether mental tokens cause anything (see esp. Heil, *The Nature of True Minds*, 104-7 and 121-2).

[138] Heil, *The Nature of True Minds*, 138.

[139] "The liquidity of Clara's soup is realised by its molecular structure only if liquidity supervenes on molecular structure and the former `trope' is constituted by the latter" (Heil, *The Nature of True Minds*, 138).

Davidson's account is not enough. That is to say, he agrees that even if mental tokens are causes, there remains a *further* question as to whether they are causally relevant;[140] and he regards his trope account as being crucial to answering this further question.[141] But this further question would not even arise if causation and causal relevance were the same relation. Heil's argument, then, involves the conflation of what he himself takes to be two distinct relations. He, like the Macdonalds, relies on an equivocation.

V. Counterexamples to the Trope Account

Heil does not believe that his remarks about tropes are in themselves sufficient to allay the fear that epiphenomenalism might be true.[142] He denies that the mere realisation of a mental property instance by a causally relevant physical property instance is enough to guarantee the causal relevance of the mental trope. To illustrate this

---

[140] This is especially clear from Heil, *The Nature of True Minds*, 122.
[141] This is evident from Heil, *The Nature of True Minds*, 123, where it is suggested that only a trope account can make sense of the fact that a supervening property (in this case, being liquid) "matters causally," i.e., is causally relevant. Pierre Jacob adopts a similar approach in Pierre Jacob, *What Minds Can Do* (Cambridge: Cambridge University Press, 1997), 218-9.
[142] It is thus strange to find Heil repeatedly cited as a proponent of the trope solution. E.g., he is characterised as such in Crane, "The Mental Causation Debate," 222 n. 23; in Robb, "The Properties of Mental Causation," 188 n. 21; and in Paul Noordhof, "Do Tropes Resolve the Problem of Mental Causation?" *The Philosophical Quarterly* 48 (1998): 221-26 (at 222 n. 5). More recently, Heil has offered a response to the problem of epiphenomenalism that (despite his protestations to the contrary) appears to be eliminativist (John Heil, *Philosophy of Mind* [London: Routledge, 1998], 200-1).

point, he appeals to the well-worn example in which Ella shatters a glass by singing, "Break not my heart."[143] Clearly, her singing causes the glass to break, but it does not produce this effect in virtue of being a singing of "Break not my heart"; this content is causally irrelevant. This is so in spite of the fact that its token in this instance is realised (according to Heil) by Ella's singing at just that pitch and amplitude, and in spite of the fact that this latter physical token is causally relevant to the effect.

A similar example has been suggested by James Robert Brown.[144] Having a given structure may be causally relevant to the vase's breaking. Moreover, its delicate structure may be what makes the vase beautiful. Its beauty, then, is realised in its structure. It does not follow, however, that the vase's beauty is causally relevant whenever its structure is. Regardless of whether we have here two tropes related by the realisation relation (as in Heil), or one property instance or event that falls under two different descriptions (as in the Macdonalds and Lewis), it is simply

---

[143] Heil, *The Nature of True Minds*, 139-40. The example was first given in Fred Dretske, *Explaining Behavior: Reasons in a World of Causes* (Cambridge, MA: MIT Press, 1988), 79.

[144] The example was given in James Robert Brown, commentary on "Does Anything Break Because It Is Fragile?" by Paul Raymont (paper presented at the annual meeting of the Ontario Philosophical Society, Toronto, Ont., October 1997).

not the case that these connections underwrite the transfer of causal relevance from the vase's structure to its beauty.

Stephen Yablo advances a similar counterexample against the Macdonalds' account to show that two properties can share the same property instance without both being causally relevant to that instance's effects.[145] Yablo's criticism has the added virtue of being equally effective even if the properties involved are related to each other not just by supervenience but by the more intimate relation of determination. So, for example, suppose that I, weighing 165 pounds, tip the scales, and that anyone weighing more than 120 pounds would also tip them. Clearly, my having the property of weighing 165 pounds is causally relevant to my tipping the scales, as is my determinable property of weighing more than 120 pounds. By contrast, my weighing less than 180 pounds is surely not causally responsible for that effect, despite the fact that it too is a determinable of weighing 165 pounds. Thus, even though in my case the properties of weighing 165 pounds and of weighing less than 180 pounds are had in the same single property instance, and even though these features are related to each other as determinate to determinable, this does not guarantee that they will both be causally relevant whenever one of them is.

---

[145] Yablo, "Mental Causation," 259 n. 32.

To clarify, the criticism is *not* directed at the claim that the property instance of my weighing less than 180 pounds is a cause of my tipping the scales. It would be wrong to deny this claim, since its truth follows from the identity of the aforesaid property instance with the event of my weighing 165 pounds (which really did cause the scales to tip), together with the extensionality of the causal relation. Instead, the criticism is intended to show that it is not in virtue of its being an instantiation of the property of weighing less than 180 pounds that the event causes the scales to tip; and that its being a having of this property is therefore causally irrelevant to that effect. In short, the Macdonalds' (and Lewis's) willingness to countenance property instances that incorporate more than one property leaves them open to a similar objection to the one that confronts Davidson, namely, that not all of those properties need be relevant to the property instance's effects.

VI. Responses to Yablo

The Macdonalds consider Yablo's counterexample in a recent paper.[146] Their response seems to be that while the counterexample illustrates the futility of their strategy as a means for establishing the causal *relevance* of mental *properties*, it does not undermine their claim that mental

[146] Macdonald and Macdonald, "How to be Psychologically Relevant," 68.

property *instances* are *causes*. Moreover, they only had this latter claim in mind when they set out to establish the "efficacy" of the mental. Thus, establishing the causal relevance of mental properties is a quite different task from the one that concerned them in the paper that Yablo criticises. Fair enough, one wants to reply, but if this is how we are to have interpreted their earlier paper, then it seems once again that their story about property instances makes no progress beyond Davidson's account and is, at best, merely a transposition of his view into a metaphysics in which the causal relata are more fine-grained than his events. We cannot by this manoeuvre escape the worry that confronts Davidson's theory, namely, the difficulty that not all of a cause's features are relevant to its production of a given effect, so that the causal relevance of mental properties is not established merely by showing that they characterise events that are causes.[147]

David Robb has a bolder response to Yablo.[148] According to Robb, Yablo's putative counterexample in fact illustrates one of the chief virtues of the property instance strategy, namely, the insight that the bearers of causal relevance are

---

[147] It is also hard to see how the Macdonalds' 1986 paper could have met its stated aim of responding to Ted Honderich's criticisms of Davidson, since Honderich seems to have been concerned with the causal *relevance* of mental properties in Davidson's account, and not simply with the question whether mental events are causes (Honderich, "The Argument for Anomalous Monism,").

[148] Robb, "The Properties of Mental Causation," 191-4.

particulars (viz., property instances or *tropes*, to use his preferred term)[149] rather than types. Robb happily affirms the truth of both

> (i) I tipped the scales in virtue of weighing 165 pounds,

and

> (ii) I tipped the scales in virtue of weighing less than 180 pounds.

Both of these claims are true, in Robb's view, because, despite appearances, what follows ". . . in virtue of . . ." is a singular term that refers to a particular property instance. Only these particulars, and not the types to which they belong, can be said to be causally relevant. Accordingly, statements (i) and (ii) are both true because they each pick out the same trope (viz., my weighing 165 pounds) and say of it that *it* is causally relevant with respect to my tipping the scales. The manner of its description does not determine whether it, that property instance, is causally relevant. On Robb's approach, then, ". . . in virtue of . . ." is an extensional context: it does not matter how we refer to the trope, for as long as we do succeed in referring to it, our claim will be true (if in fact that trope was causally relevant).

---

[149] Robb uses the term `trope' where the Macdonalds use `property instance'. He gives `trope' a much broader usage than Ehring and the Macdonalds allow it. Whereas they regard tropes as the exclusive preserve of nominalists, Robb treats the usage of `trope' as being neutral between nominalism and realism with respect to universals (see esp. Robb, "The Properties of Mental Causation," 186).

VII. The Trouble With Tropes

Robb's account of (i) and (ii) does not accurately represent the force of ". . . in virtue of . . ." claims. Clearly the point of such claims is not to express some fetishistic attachment to a particular trope, according to which the scales tipped because of the presence of *that* token. The point is rather to assert that the scales tipped because the cause was of a certain type, so that *other* tokens of the same type can be expected to produce the same sort of effect. In short, what follows the ". . . in virtue of . . ." is not a singular term but is instead a general term specifying a type.

Robb's denial of this severs the link between causal relevance and explanation, thereby rendering the former notion devoid of content. To see why, consider how meaning accrued to the notion of causal relevance in the first place. We tested for causal relevance in terms of what is explanatory, and it was because not every way of picking out the cause was explanatory that we concluded that not all of its properties were causally relevant. For example, picking out the hurricane as "The event reported on page four of the Times" goes no way towards explaining the fatalities, so the hurricane's having been reported in the Times is not causally relevant to this effect. Similarly, it is the

explanatory impotence of our characterising the brick's impact on the window as "The window's touching a red object" that leads us to deny the relevance of the brick's redness to its breaking of the window. In view of this, the contexts that express relations of causal relevance appear to be intensional, for causal relevance at least requires being explanatory and can thus be born only by particulars *under descriptions* (or under types) and not simply by particulars in and of themselves.

A further difficulty arises from Robb's attempt to account for our inclination to regard (ii) as being false. According to him, we are so inclined not because (ii) really is false but instead because it pragmatically implies a falsehood. Whereas (i) implies the truth that weighing 165 pounds is sufficient to tip the scales, (ii) implies the falsehood that weighing less than 180 pounds is also sufficient to yield this effect. More accurately, it implies that having a *weighing-less-than-180-pounds* trope is sufficient to tip the scales.[150] It is unclear, though, how (ii) can imply any such perfectly general claim, a claim about the behaviour of *other* tropes of that type, if the scope of its ". . . in virtue of . . ." clause really is limited to one particular token (the one that is identical with my weighing 165 pounds). This is because an explanation

[150] Robb, "The Properties of Mental Causation," 193.

that just appeals to some particular thing in itself (without the manner of the thing's description figuring crucially in the explanation) implies no claims about the behaviour of anything *other* than that very particular. If I say, "He did it because of Carol," I say nothing about what people other than Carol might have caused him to do. If there is any implication at all (even a pragmatic one) about what others might cause him to do, it must be by way of some other explanation of the same effect that does *not* simply appeal to a particular (in this case, Carol). For example, it must be by way of the claim that he did it because of Carol's intelligence, or because of her sadness, or whatever. And Carol's sadness in turn implies nothing about what other states of mind might have led him to do if it is considered solely as a particular trope, as a token belonging to disparate physical and mental types, and not in terms of its being a token *of sadness*, of that type.

I take this to be true of explanations generally: without the requisite generality, they carry no implication about the causal propensities of other things. If this is true, then Robb's account of our inclination to regard (ii) as being false is inconsistent with (ii)'s alleged ascription of causal relevance to a particular trope instead of to a type. For if the only point of (ii) is to ascribe

causal relevance to a trope (and not to a trope under a description), then it is hard to see how (ii) could even pragmatically imply anything about the behaviour of other tropes (including other *weighing-less-than-180-pounds* tropes), *unless* at some point we make a claim about the behaviour of the trope *qua* its being a trope of the property of weighing less than 180 pounds. But this is precisely the sort of claim that Robb will not allow.

The foregoing arguments against the ascription of causal relevance to tropes or other particulars share this guiding insight: relevance is a denizen of logical space. Just as *things* cannot entail or be inconsistent with other *things*, likewise particulars *in themselves* cannot be relevant or irrelevant to other particulars. Their relevance is entirely a matter of how they are described. By contrast, Robb regards tropes as being causally relevant independently of how they are characterised. For him, a trope is not relevant *qua* being a trope of the type "weighing more than 120 pounds" or *qua* being a trope of the type "weighing less than 180 pounds". As he says, "Tropes are not relevant *qua* this or that, they are causally relevant (or not), period."[151] But no *thing* is just "relevant, period," for relevance is inherently *qua*sal: only things under descriptions can properly be said to be relevant (or not).

[151] Robb, "The Properties of Mental Causation," 191.

VIII. The Moral of the Story

One lesson that can be drawn from all of this is that the concept of causal relevance is an unstable compound that places competing and irreconcilable demands on its bearers. On the one hand, the Macdonalds and Heil are onto something: causal relevance is supposed to be (at least in part) *efficacy*; to be causally relevant is to be efficacious. However, as they rightly point out, abstracta and intensional items (e.g., manners of description) simply cannot fulfil this requirement for being causally relevant. These items are just not the right sorts of thing to *cause* anything to happen; they are not part of the causal flux, and hence are not genuinely efficacious with respect to anything. Concrete particulars, such as mental or physical tokens, can make things happen. Abstract objects and modes of presentation cannot.

On the other hand, examples like the one involving the red brick (where it seems that the brick's redness is causally irrelevant to the window's breaking) seem to illustrate a different requirement on causal relevance, namely, that causal relevance is a species of explanatory relevance, so that to be causally relevant is to be explanatory. Unfortunately, the only things that can meet this demand are precisely the ones that cannot fulfil the

first requirement, to wit, intensional items--things *under descriptions;* for concrete particulars are never just "relevant period" but can instead only be relevant under an aspect. (Yablo's counterexample illustrates this same point with respect to fine-grained particulars such as tropes and property instances.)

Has the first requirement (the requirement of efficacy) been misinterpreted? After all, it is not clear that when we say that a property is causally relevant we are thereby committed to the claim that it actually caused anything; we may instead only be claiming that the property in question is *that in virtue of which* the cause caused whatever it caused. Regrettably, though, this use of the phrase, "that in virtue of which", merely recapitulates the dilemma. For either this phrase means something like *the active ingredient, the vital component that gave rise to the effect,* or it means *the crucial aspect that accounts for why the effect happened.* On the first reading, "that in virtue of which" picks out a fine-grained cause, while, on the second reading, to be *that in virtue of which* the cause produced its effect *just is* to be explanatory of that effect. Yablo's counterexample shows that the things that are picked out by "that in virtue of which" on its first interpretation (viz., tropes or property instances) can be

cited in unexplanatory ways, and thus are not the items that the phrase picks out when it is read in the second way. This is just what one would expect, since the former entities are particulars (causes) while the latter ones are intensional items (ways of characterising causes that explain why their effects ensued). Hence, even when interpreted by means of the phrase "that in virtue of which", causal relevance turns out to be an unsatisfiable concept: nothing could possibly meet the irreconcilable demands that it encapsulates, for nothing could be *both* a particular throwing its weight around in the causal flux *and* an essentially explanatory way of characterising the cause.

It must be concluded that the very notion of causal relevance is an unstable compound that runs together elements of the metaphysical relation of causation with the pragmatic or epistemological notion of explanation. Under closer scrutiny, the elements of this mixture inevitably separate out in such a way that those who wish to retain this concept are tugged in opposite directions, speaking now of particulars causing things to happen and now of propositional items explaining one another, under the pretence that they are in fact talking about just one relation ("causal relevance") whose relata are at once concrete particulars and propositional entities.

It may be thought that this is too hasty. Surely, one may object, the only lesson to be drawn at this point is that causal relevance, whatever it is, is not a relation between particulars; it may yet turn out to be a metaphysical relation between the realist's properties. This seems unlikely, though, for much the same reason as was given early in the preceding section: to conceive of causal relevance as being *any* type of metaphysical relation is to sever the link between causal relevance and explanation, thereby rendering the former notion devoid of content. After all, each of the realist's properties is itself susceptible of more than one description, and not all of these descriptions characterise the property in such a way as to provide information about why the effect in question occurred. Like Robb's tropes, then, any such property can be picked out in *un*explanatory ways. This is sufficient to show that the realist's properties are not the bearers of causal relevance, for something is causally relevant only if it is explanatory.[152] The point is not simply that a bearer of causal relevance must be explanatory *under some description or other*. Rather, the point is the stronger claim that a bearer of causal relevance must be explanatory *full-stop*,

---

[152] In Chapter 2, sec. III (the "flagging" section), I said that something can be explanatory without having the sort of relevance that we regard mental properties as having. Here I wish only to deny the converse: it is not the case that something can have that sort of relevance without being explanatory.

such that merely to cite it is thereby to explain. Of course, this assimilates it to the manner in which the cause is cited, which, again, is to render it an intensional item. This puts causal relevance clearly into the field of epistemology and pragmatics. It is in no sense a metaphysical relation.

One might remain unconvinced. After all, a critic may ask, why not simply allow that when I say, "He opened the fridge in virtue of having the property mentioned on page five," I *am* in fact citing a causally relevant property, but in a way that is not explanatory? Why not just allow that something can be causally relevant without this guaranteeing that the mere citation of it will always be explanatory? The short answer is that if this is so, then causal relevance gets us no further than causation, for these are just the sorts of claims that Davidson makes about causation (viz., that something can be *a cause* without the mere citation of it being explanatory). The whole point of introducing the notion of causal relevance was to do justice to the feeling that only *some* ways of referring to the cause are explanatory. So if it turns out that only *some* ways of referring to a causally relevant property are explanatory, have we made any progress by introducing this notion? If one replies that the mere possibility of referring to causally

relevant properties in unexplanatory ways is not a genuine problem that should compel us to seek a remedy, then why not simply agree with Davidson when he says much the same thing about the possibility of referring to causes in unexplanatory ways? Perhaps we should simply have heeded his advice and refrained from the quest for a remedy which led us to posit relations of causal relevance in the first place. Or, at least, if we do see a legitimate role for talk of causal relevance, then it must be as talk about an epistemological or pragmatic relation rather than a metaphysical relation between things in the world (whether they be tropes or properties).

In short, then, *whatever* causal relevance turns out to be, the manner of description is decisive for its obtaining between any two relata. This was just the lesson of the previous section, namely, that causal relevance can only obtain between things under descriptions, not between things in themselves.

4. Explanatory Relevance

> If we knew thoroughly the nervous system of Shakespeare . . . we should be able to show why . . . his hand came to trace on certain sheets of paper those crabbed little black marks which we . . . call the manuscript of Hamlet. We should understand the rationale of every erasure and alteration therein . . . without in the slightest degree acknowledging the existence of the thoughts in Shakespeare's mind. The words and sentences would be taken, not as signs of anything beyond themselves, but as little outward facts, pure and simple.
>
> William James[153]

> What is left entirely unexplained is just the play of Hamlet, as such. The play, as such, is not merely the material thing which we describe as constituted by certain black marks on certain sheets of paper. It is rather the *meaning* of these marks and of their arrangement. . . . The only possible explanation is that the thought and will of Shakespeare expressed themselves in and through the written characters.
>
> G.F. Stout[154]

> What I called jottings would not be a *rendering* of the text, not so to speak a translation with another symbolism. The text would not be *stored up* in the jottings. And why should it be stored up in our nervous system?
>
> Wittgenstein[155]

I. Beyond The World of Little Outward Facts

The above quotations of James and Stout represent one of the most popular, and venerable, ways in which to account for the explanatory role of appeals to mental states. This strategy had its inception in the parallelism of Leibniz and various Cartesian philosophers, who were grappling with the apparent idleness of mind in a world in which all motions were to be explained in mechanical terms only. Their

---

[153] James, *The Principles of Psychology*, 1:132.
[154] G.F. Stout, *A Manual of Psychology* (London: W.B. Clive, 1899), 99-100.
[155] Ludwig Wittgenstein, *Zettel*, ed. G.E.M. Anscombe and G.H. von Wright, trans. G.E.M. Anscombe (Berkeley: University of California, 1967) 612.

solution was to posit two synchronised, but independent, causal chains, one mental, the other mechanical: never the twain shall meet, but always shall agree. This view, though it may seem a desperate expedient, prescinds from the perhaps equally desperate postulation of physical gaps that are plugged by mental forces.

James and Stout offer their own brand of parallelism, according to which what stands outside of, and parallel to, the physical world are meanings. Moreover, meaning-bearers, insofar as they are meaningful, are not amenable to a merely physical explanation. Thus, *Hamlet*, the play as such, cannot be explained simply by reference to the workings of Shakespeare's physiology. All that we can thereby explain are "those crabbed little black marks" on the paper, taken as purely physical markings ("little outward facts") rather than as meaningful expressions. Moreover, all human behaviour has this double aspect, whereby it can be regarded either as mere bodily motion or as meaning-saturated action, susceptible of being interpreted in a variety of ways. Of course, this bifurcation between meaningful and meaningless aspects pertains to the explanans as well as the explanandum. If the meanings of *Hamlet* are not "stored up" (to use Wittgenstein's phrase) in the little black marks on the paper, neither are they to be found in the little grey

marks in the brain. The parallelism is complete: on one side we have physiology and the mere bodily motions that it can be called upon to explain, while on the other side we have Shakespeare's meaningful "thought and will" (as Stout says), which is not stored up in his physiology, and which alone can explain his actions.

The central claim in this view is that our actions are individuated essentially in terms of their meanings, and that, as such, they are wholly absent from the physical perspective. They do not appear in the world as described by the physical sciences, for the language of those sciences is simply not equipped to register the presence of intentional phenomena (*qua* intentional), including meaningful thoughts and actions.

This has been a prevalent theme throughout the twentieth century. One of its more well-known expressions in analytic philosophy can be found in Roderick Chisholm's and Peter Geach's criticisms of behaviourism.[156] Chisholm and Geach argued that the attempt by the behaviourist to analyse mental states exclusively in terms of behavioural dispositions founders on the realisation that the conditionals that express the dispositions must, if they are to be even plausible, make reference in their antecedents to

---

[156] Roderick Chisholm, *Perceiving* (Ithaca, NY: Cornell University Press, 1957); and Peter Geach, *Mental Acts* (London: Routledge and Kegan Paul, 1957), 8.

mental background conditions. Since they thus always presuppose intentional locutions, the conditionals in question cannot provide an exhaustive analysis of all such language. The upshot is that if we were to confine ourselves to the behaviourist's language of physical objects and mere bodily motions (e.g., sound waves produced by the vibration of the vocal cords), then there could be no hope of capturing the sense of intentional language. In short, within the behaviourist's purely physical perspective, mental phenomena do not appear.

A similar critique of John Watson's behaviourism was offered thirty years earlier by the social psychologist, William MacDougall.[157] MacDougall maintained that behaviourism is self-defeating, since the very data that it purports to account for (viz., intentional behaviour) cannot even be "described intelligibly and profitably" if we restrict ourselves to the categories of the physical sciences. According to MacDougall, psychology, as one of the *Geisteswissenschaften*, is an autonomous science that deploys a conceptual apparatus that is fundamentally alien to the physical sciences.

---

[157] John B. Watson and William MacDougall, *The Battle of Behaviorism* (New York: W.W. Norton & Company, Inc., 1929), 91-2. MacDougall says that he earlier presented his criticism of behaviourism in William MacDougall, Presidential Address to the Psychological Section of the British Association, Toronto, 1924.

In framing this critique, MacDougall was influenced by German social theorists, such as Wilhelm Dilthey and Max Weber. For them, the distinctive mark of the human sciences is their concern with meaning.[158] This fixation on meaning derives from the central role in the human sciences of action. "Action," according to Weber, "is rationally evident chiefly when we attain a completely clear intellectual grasp of the action-elements in their intended context of meaning."[159] Actions, that is, can only be identified by grasping their meanings. For example, when we identify the man's supporting the rifle at shoulder length as his *aiming* the rifle, we exploit the same sort of capacity that is at work when we identify a facial pattern as an expression of anger.[160] In both cases, we exhibit a "direct observational understanding of the subjective *meaning* of a given act as such."[161] Beyond this, we have a capacity not only to identify but also to explain actions in terms of their

---

[158] I rely here on Max Weber's methodological essay, "Basic Sociological Terms," in *Economy and Society*, ed. Guenther Roth and Claus Wittich, original translation by Talcott Parsons (1947) revised by Roth and Wittich (New York: Bedminster Press, 1968), 3-62; originally published in 1922; as well as on Martin Hollis's summary, "Philosophy of Social Science," in *The Blackwell Companion to Philosophy*, ed. Nicholas Bunnin and E.P. Tsui-James (Oxford: Blackwell, 1996), 358-87 (at 368-70). Weber refers to neo-Kantians (Heinrich Rickert, Ferdinand Toennies and Georg Simmel) as precursors of his own view (Weber, "Basic Sociological Terms," 3-4). However, as Hollis notes (Hollis, "Philosophy of Social Science," 368), Dilthey himself claimed to be following Hegel in expounding his views on the *Geisteswissenschaften*.

[159] Weber, "Basic Sociological Terms," 5.

[160] These examples are borrowed from Weber, "Basic Sociological Terms," 8.

[161] Weber, "Basic Sociological Terms," 8. (Emphasis added) Weber adds that we make use of this same sort of understanding whenever we interpret verbal utterances.

meanings. "This," says Weber, "is rational understanding of motivation, which consists in placing the act in an intelligible and *more inclusive context of meaning.*"[162]

Weber briefly considers the possiblity that, "Future research may be able to discover non-interpretable [i.e., non-intentional] uniformities underlying what has appeared to be specifically meaningful action."[163] However, this would not, in his view, undermine the social sciences. This is because, "The recognition of the causal significance of such factors would not in the least alter the specific task of . . . sociological analysis or . . . the other sciences of action, which is the interpretation of action in terms of its subjective meaning."[164] Thus, the social sciences cannot be displaced by the physical sciences, since the former sciences invoke a unique form of understanding that enables us to identify the intention and meaning behind the agent's behaviour. Without this kind of understanding, the agent's behaviour can only be understood as mere behaviour, mere bodily motion. This is how the agent's behaviour appears through the lens of the physical sciences. In the human sciences, by contrast, we apply a new set of categories,

---

[162] Weber, "Basic Sociological Terms," 8. (Emphasis added)
[163] Weber, "Basic Sociological Terms," 7-8.
[164] Weber, "Basic Sociological Terms," 8. It is not clear why Weber continually includes the modifier "subjective", since he denies the existence of "an objectively `correct' meaning or one which is `true' in some metaphysical sense" (Weber, "Basic Sociological Terms," 4).

including the concepts of purpose and value, which enable us to see bodily motions as *meaningful* actions.

Once again, then, our actions are said to be identifiable as actions only in terms of their meanings, and are thus held to be absent from the physical perspective. Moreover, actions, as explananda standing outside of the physical sciences, may plausibly be thought to have explanations that are themselves alien to those sciences. We thus have some real explanatory work that can be done by (in this case) the human sciences.

This theme in the social sciences is, in some respects, very similar to the strategy adopted by James and Stout. In both cases, we start with a mental or social explanans that is individuated in terms of its meaning, and that thus becomes problematic. How can meanings explain, given that they have no legitimate place in the physical sciences? The solution is to distinguish between two kinds of explananda that are often run together in our talk about behaviour: there is mere bodily motion, which can be explained exclusively in physical terms; and there is action, which must be identified in terms of its meaning, and which thus also stands outside of physics, chemistry and biology. As such, actions become the proprietary concern of folk psychology and the human sciences.

II. Recent Applications of This Strategy

Acceptance of this approach does not require some vague metaphysical notion of causal relevance, since all this talk about giving psychology some explanatory work to do can be cashed out exclusively in terms of explanatory relevance. This strategy also involves no commitment to dualism. James and Stout were in fact responding to old-fashioned epiphenomenalism, and may well have assumed the dualism implicit in that doctrine. It should be clear, though, that a proponent of this sort of approach need not deny the identity of mental (i.e., meaning-bearing) and physical events, but may instead insist only on the irreducibility of properties concerning meaning to physical types.

With these clarifications in mind, it is interesting to note that variants of James's strategy have recently been proffered in an attempt to delimit the manner in which mental facts explain. Ausonio Marras, for example, believes that mental features owe their explanatory power to the fact that, "Explanatory contexts . . . are *nonextensional* and *context-dependent.*"[165] Whether something is a good explanation depends on how we type-identify both the cause and the event to be explained, and this in turn depends upon the context. To use Marras's example, my getting a drink of

---

[165] Ausonio Marras, "The Causal Relevance of Mental Properties," *Philosophia* 25 (1997): 389-400 (at 397). (Emphasis in the original)

water may be identical with a sequence of bodily movements, but these aspects of my behaviour, "Though tokenable by the same event on a given occasion, are obviously distinct types of event, and call for distinct explanations."[166] For Marras, intentional explanations of behaviour simply do not compete with physical accounts of the same events, for although the same event figures in the explananda of these two explanations, it appears under different descriptions in each one. Thus, since "an event is an *explanandum* only as *described*,"[167] the mental and physical explanations are explanations of different explananda. In this way, Marras believes, we can overcome worries about explanatory exclusion: as explananda that are couched in the language of folk psychology, actions are the exclusive preserve of intentional explanations. These explanations are the only ones available to us when the explanandum at issue is an action. They have no explanatory rivals.

Lynne Rudder Baker suggests a similar approach.[168] She puts her case in terms of an example involving the collapse of a savings and loan institution. Suppose we identify the bad investment that prompted its collapse. Call the microphysical states that constitute the bad investment a *U*-

---

[166] Marras, "The Causal Relevance of Mental Properties," 398.
[167] Marras, "The Causal Relevance of Mental Properties," 398. (Emphasis in the original)
[168] Lynne Rudder Baker, *Explaining Attitudes* (Cambridge: Cambridge University Press, 1995), 134-5, and 148-50.

state. Baker rejects the claim that the explanations of the bank failure respectively in terms of the bad investment and in terms of the U-state are in competition with one another, and that the latter explanation displaces or invalidates the former one in the way that explanations that appeal to oxygen came to replace explanations that appeal to phlogiston. According to Baker, the explanations appealing to phlogiston and oxygen are in competition with each other because they "share a single explanandum."[169] By contrast, "The U-state explanation and the investment explanation explain different things";[170] for the U-state explanation at best only explains the microphysical phenomena that constitute the bank failure, whereas "the investment explanation explains the bank failure *as* a bank failure."[171] Like Marras, then, Baker maintains that when the event to be explained is characterised in intentional terms, only an intentional explanans can meet our explanatory needs.[172] Also like Marras, she stresses considerations having to do with how we count explananda. Given the nonextensionality of explanation, it is just not true that explanations of the same event will always share the same explanandum and thus

---

[169] Baker, *Explaining Attitudes*, 134.
[170] Baker, *Explaining Attitudes*, 134-5.
[171] Baker, *Explaining Attitudes*, 135. (Emphasis in the original)
[172] Baker is clear about the intentional nature of the explanation of the bankruptcy. She says, "Nothing is a bankruptcy or an investment in a world without complex economic practices, practices that could not exist in a world without attitudes" (Baker, *Explaining Attitudes*, 128).

be in competition with each other. Hence, we can affirm the identity of each action with a physical token and still insist that actions as such remain the exclusive proprietary concern of intentional explanations, and are not explained by the physical explanations of the physically described tokens with which they are identical.

Jennifer Hornsby is a third recent proponent of James's strategy.[173] Hornsby denies that actions are accessible from the impersonal point of view, and takes this to show that they cannot be explained by being located in the law-governed world of the physical sciences. For her, an action can only be explained by citing the agent's reasons for so acting.

III. Monism Preserved

Unlike Marras and Baker, Hornsby denies that actions, and the reasons that explain them, are identical with physical tokens.[174] She takes Davidson to task for his monism, implying that to be a monist at all (even an anomalous monist) is thereby "to view the mental impersonally"[175] and to render actions "accessible from the impersonal point of view."[176] This, however, is not a fair characterisation of anomalous monism, and involves an

---

[173] Jennifer Hornsby, "Agency and Causal Explanation," in *Mental Causation*, ed. Heil and Mele, 161-88.
[174] Hornsby, "Agency and Causal Explanation," 169-74.
[175] Hornsby, "Agency and Causal Explanation," 171.
[176] Hornsby, "Agency and Causal Explanation," 169.

exploitation of an ambiguity in the phrase, "the presence of actions to an impersonal point of view."[177] Hornsby reads this expression *de re*, such that actions are said to be *present to* the impersonal point of view as long as they are identical with items that can be characterised in any terms whatever (e.g., in purely physical terms) within the impersonal perspective. On this reading, Davidson affirms, and Hornsby denies, that actions are present to the impersonal point of view. However, a Davidsonian can still deny that actions are recognisable *as actions* within the impersonal perspective. This approach allows for the identity of actions with physical tokens, but emphasises that those tokens only count as actions within the holistic and normative discourse of folk psychology. They do not register as actions from the impersonal point of view, within the discourse of the physical sciences. In this sense, Davidson can retain his monism while denying that he thereby views the mental impersonally, or makes actions accessible from the impersonal point of view.

As an illustration of this point, consider what would be involved in maintaining that actions are accessible from the impersonal perspective, in the strong sense of being recognisable as actions from that vantage point. This very claim is made by those who are exercised by a putative

[177] Hornsby, "Agency and Causal Explanation," 169.

problem in our conception of agency. The problem was clearly articulated by Thomas Nagel.[178] It involves an apparent dissolution of agency when the agent and her actions are viewed "from an objective or external standpoint."[179] According to Nagel, when action is thus viewed,

> Some of its most important features seem to vanish under the objective gaze. Actions seem no longer assignable to individual agents as sources, but become instead components of the flux of events in the world of which the agent is a part.[180]

On this model, actions are thought to be discernible as actions within the objective viewpoint. The agent as a source of action fades from view, but her actions remain. They are just *there*, *happening* rather than being done by someone. They are thus held to be identifiable as actions without there being any identification of an agent whose actions they are.

---

[178] Thomas Nagel, *The View From Nowhere* (Oxford: Oxford University Press, 1986), 110-20.

[179] Nagel, *The View From Nowhere*, 110.

[180] Nagel, *The View From Nowhere*, 110. Similar thoughts may be at work in Ludwig Wittgenstein, *Philosophical Investigations* (Oxford: Basil Blackwell, 1953), sections 611-632 (esp. at 611 and 620). Interestingly, section 611 contains an allusion to Schopenhauer. According to Christopher Janaway, Schopenhauer, like Nagel, was perplexed by this juxtaposition of my action as something that I do with my action as an event that merely happens or occurs (Christopher Janaway, *Self and World in Schopenhauer's Philosophy* [Oxford: Clarendon Press, 1989], 246-7). Schopenhauer expounded a dual-aspect position, in which the subjective view of the mind (for which "consciousness" is fundamental) is contrasted with the objective view that is offered by the empirical sciences (chiefly zoology and physiology). In his words, "On the purely *objective* path, we never attain to the inner nature of things, but if we attempt to find their inner nature from outside and empirically, *this inner always becomes an outer in our hands*" (Arthur Schopenhauer, *The World as Will and Representation*, trans. E.F.J. Payne [New York: Dover, 1969], 2:273-4). (Second emphasis added)

It is this assumption (that there can be recognition of an action antecedently to the identification of an agent) that generates the problem considered by Nagel, and that Hornsby wishes to reject. According to her, "Seeing something as an action requires the identification of a person."[181] Moreover, a person is a "causally complex whole"[182] that exhibits a rational pattern by conforming to certain norms of consistency and coherence. It is only by virtue of this approximation to an ideal of rationality that the behaviour of the causally complex system can be interpreted in terms of such concepts as belief, desire, intention *and action*. Thus, without the backdrop of a complex system conforming (at least roughly) to the ideal of rationality, that is, without a person, the concept of *action* simply has no application.

This is, of course, just what a Davidsonian would say, and in saying it, he would side with Hornsby, against Nagel, in claiming that actions do not appear in the sub-personal framework. In saying this, he means that nothing counts as an action, or a belief, or a desire, antecedently to our attainment of the personal level, at which the norms of rationality apply. A thing only takes on any of these designations insofar as it coheres with the normative and

---

[181] Hornsby, "Agency and Causal Explanation," 174.
[182] Hornsby, "Agency and Causal Explanation," 172.

holistic network of intentional states that the agent is interpreted as having. And if it fits into that web, then it is identifiable as *that agent's* action, and not as some state that can somehow (like the smile of the Cheshire cat) float free of the agent whose action it is while remaining recognisable as that action.

Davidsonians and Hornsby agree on this much. Their disagreement arises from Davidson's monistic claim that actions, like beliefs and desires, are events or processes that fall under physical as well as psychological descriptions. Moreover, of all the considerations that Hornsby adduces to establish the absence of actions from the impersonal viewpoint, only one militates against this monistic claim.[183] She asks us to consider all of the events that appear in the impersonal view and that are likely to be thought of as antecedents, parts or consequences of the action. These will include "a whole collection of events leading from some happening in the depth of Peter's brain all the way to an event beyond his body in which his desire's being satisfied consists."[184] According to Hornsby, it is impossible to delimit precisely the action's boundaries in terms of these events. There is simply no way in which to specify exhaustively which of these events the

[183] Hornsby, "Agency and Causal Explanation," 174-5.
[184] Hornsby, "Agency and Causal Explanation," 174.

action consists of. From this she concludes that the action itself is absent from the impersonal viewpoint, in the strong sense of not being identifiable in any terms whatever (not necessarily as an action) from that perspective. The series of events initiated within Peter's brain and emanating outwards from his bodily movements inhabit "an impersonal point of view, from which it is impossible to locate actions."[185]

This strong conclusion does not follow from Hornsby's argument. Granted, actions have fuzzy boundaries, but then so too do riots and storms. These latter have their place in the impersonal view, even though they too lack precise spatial and temporal boundaries. More generally, the fuzziness of macro-physical entities does not preclude their being physical and thus present to the impersonal view. Moreover, Hornsby herself directs our attention to the impersonally identifiable events in question (with which actions allegedly cannot be identified) by asking us to consider both the events "which the action caused"[186] and the "causal antecedents"[187] of the action. This suggests, though, that we have already located the action in the impersonal matrix in which these events are precipitated. After all, it is hard to see what our talk of "locating" something in the

[185] Hornsby, "Agency and Causal Explanation," 175.
[186] Hornsby, "Agency and Causal Explanation," 174.
[187] Hornsby, "Agency and Causal Explanation," 174.

impersonal view can amount to unless it just means finding a place for it on one stretch of the impersonal causal chain, such that it can be picked out by referring to causal antecedents and outcomes that are identifiable on that chain.

IV. Anomalous Monism and the Jamesian Strategy

To reiterate, apart from this difference between Davidson and Hornsby, Davidsonians should find Hornsby's outlook to be quite congenial to their own. They can agree that actions are not identifiable as actions until we reach the personal level. They can also agree that the physical sciences, addressing as they do only the sub-personal levels, simply do not concern themselves with the explananda that occupy us when we set out to explain actions. These sciences at best only explain the events and processes, which are the actions, under sub-personal (viz., physical) descriptions. To do so is not to explain them *as actions*, and is thus to leave aside the explananda which are the focus of folk psychology.

In this way, anomalous monism might seem to be the natural heir to the strategy articulated by James and Stout. Read in this way, the theory assigns an explanatory role to intentional features by demarcating a range of intentional explananda, *actions*, that only reasons can explain. However,

there are two reasons why Davidsonians should eschew this strategy.

First, the Jamesian strategy concedes too much to the epiphenomenalist. There is a whole range of explananda that we typically explain in mental terms but that Hornsby, Baker and Marras represent as being beyond the pale, explainable in the terms of the physical sciences only. For example, if I open the refrigerator to get a Coke, the refrigerator door will be open, the contents of the pop bottle will have been diminished, and the bottle will occupy a different position on the refrigerator shelf as a result of my action. While the action may not itself be describable as an action within the terms of the physical sciences, it seems that each of these explananda can be captured in those terms. Moreover, these physically characterised explananda are outcomes of my action, and can be explained as such. Why does the bottle now occupy this position instead of the one it had ten minutes ago? It would seem to be a perfectly satisfactory answer to say that after I poured myself a drink, I left the bottle there. In short, action leaves its mark on the world, and this remains true when the world of which we are speaking is the world of "little outward facts". Our conception of agency derives in large measure from this capacity of our actions to reverberate through the sub-

personal levels of description and to reshape the world as described in those terms. This much of our agency is not protected by the Jamesian strategy.

The second reason for rejecting the Jamesian strategy is that it belies an acceptance of the picture that gives rise to worries about "explanatory exclusion". In seeking to demarcate some domain of explananda which can then be held out as the exclusive proprietary concern of intentional explanations, we are already subscribing to the view that in order for properties to be explanatorily relevant, they must stake a claim to some range of explananda which they alone can explain. Here, explanations are conceived as rivals, contending with each other for the rights to any given explanandum. They must so contend, it is thought, because any explanation of explanandum *E* excludes all others; if an explanans couched in the language of the physical sciences has already explained *E*, then intentional explanations must "shove off" (so to speak) and find some other fact to explain. If we accept this picture, then it makes sense to start looking for some group of explananda that elude the physical sciences, and for the explanation of which intentional states thus face no plausible rivals.

We can start to undermine the view that explanations compete against each other by noting first that it can gain

no support from worries about overdetermination. Overdetermination is only objectionable when it requires ontological profligacy, and it is only implicated in such excess when it is overdetermination *by causes*, rather than merely by *the descriptions* under which the causes fall. Only this latter kind of overdetermination is at issue in the present context, for here we are only considering explanational (as opposed to causal) overdetermination, that is, the possibility that multiple properties of the cause are explanatorily relevant to the effect, and not that a plurality of *things* (e.g., tropes or some other kind of cause) are present and acting to bring about the effect. In short, there is no violation of Ockham's razor, for we are not multiplying entities but only the properties of them that might, in a given context, stand out as being particularly salient to the explanatory task at hand.

There are additional, more positive reasons that can be mobilized against the view that explanatorily relevant properties are rivals that must compete against each other for their explanatory relevance. These considerations can be brought into clearer focus by investigating more thoroughly the implications of the conclusion of the previous chapter, namely, the denial that causal-explanatory relevance is any kind of extensional, metaphysical relation at all.

V. Why There is No Competition for Explanatory Relevance

Brian McLaughlin does not take this result to have much bearing on the debate about mental causation.[188] He takes Davidson to task for having claimed that the extensional nature of causation undermines the view that an event causes its effects "in virtue of" its properties.[189] According to McLaughlin, extensional accounts of causation do not have this result. That is, it is perfectly consistent for us to affirm that the causal relation obtains only between non-abstract, particular events, and that if event *c* causes event *e*, then *d* caused *e* if *d* is identical with *c*, while at the same time maintaining that *c* caused *e* in virtue of certain of *c*'s properties.[190] McLaughlin sets out to motivate this claim by means of an analogy. The analogy involves the relation, *weighing less than*. This relation is extensional: it too obtains between non-abstract particulars, and *weighing-less-than* contexts meet the standard set by the intersubstitutivity of co-referential terms. Nonetheless, McLaughlin adds, if *a* weighs less than *b*, it will do so *in virtue of* one of its properties, namely, its weight. Similarly, causes only cause their effects *in virtue of* some

---

[188] McLaughlin, "On Davidson's Response to the Charge of Epiphenomenalism," 27-40.
[189] Donald Davidson, "Thinking Causes," in *Mental Causation*, ed. Heil and Mele, 3-17.
[190] McLaughlin, "On Davidson's Response to the Charge of Epiphenomenalism," 30-1.

of their properties. Contrary to what Davidson says, this in no way conflicts with the view that causation is extensional. If we follow McLaughlin in saying this, then we shall once again have to face the worry that the features in virtue of which causes cause their effects are defined in such a way as to exclude mental properties from causal-explanatory relevance.

To clarify, McLaughlin (unlike Sosa and others) is not claiming that it is really *c*'s-being-*F* (rather than simply *c*) that causes the effect. He accepts the Davidsonian view that just as it is only concrete particulars that can weigh less than other things, so too is it only concrete particulars (events in this case) that can cause or be caused. McLaughlin merely wishes to point out that these extensional relations obtain only because the particulars involved have the properties that they have. To return to his analogy, "If *a* weighs less than *b*, then *a* has some weight, w_1, and *b* has some weight, w_2, w_1 is less than w_2, and *a* weighs less than *b* in virtue of this."[191]

This still contravenes the spirit, at least, of Davidson's response to his critics, for it involves a hypostatisation of weights, and takes their interrelation to be primary: *first* we have the *abstracta*, w_1 and w_2, and it

[191] McLaughlin, "On Davidson's Response to the Charge of Epiphenomenalism," 31.

is only *because* of their relation to each other that *a* stands in the *weighing-less-than* relation to *b*. It is at least as plausible to conjecture that the order of explanation runs in the opposite direction: *a* and *b* take on their respective weight designations only in virtue of the distribution of the *weighing-less-than* relation. That is to say, the relation between these two concrete particulars is primary, and the properties (weights in this case) in virtue of which it allegedly holds in fact derive from it. This more parsimonious approach avoids reifying weights. It avoids treating them as real things that stand in certain relations to each other. More pertinently, it avoids treating them as things the interrelation of which is the foundation of relations between real concrete particulars. This route certainly seems more congenial to Davidson's outlook, especially in its refusal to countenance a network of abstract objects which somehow undergirds and determines the way the world is (e.g., *a*'s weighing less than *b*).[192] Moreover, it allows Davidson to reject McLaughlin's claim that *a* weighs less than *b in virtue of* some relation between w_1 and w_2.

---

[192] Indeed, the inspiration for this paragraph is Davidson's remark that, "It is *events* that have the power to change things, not our various ways of describing them" (Davidson, "Thinking Causes," 12 [Emphasis in the original]). Also, "For me, it is events that have causes and effects. Given this extensionalist view of causal relations, it makes no literal sense . . . to speak of an event causing something as mental, or by virtue of its mental properties, or as described in one way or another" (Davidson, "Thinking Causes," 13).

Abstracta are inefficacious. They do not *make* one thing weigh more than another. Nor do they *make* one event cause another. That is to say, the cause does not cause its effect in virtue of any of its properties, not even its strictly nomic ones; rather, properties figure in strict or hedged generalisations only because the causal relations between concrete particulars are such as they are. William Child has put this view with admirable clarity.[193] While he does not frame his discussion as a response to McLaughlin, he does take himself to be developing the Davidsonian position in such a way as to undermine the putative implications of the "in virtue of" talk that McLaughlin and others try to exploit. According to Child, "Causation is a basic, natural [extensional] relation between events."[194] It is *basic* in the sense that it "does not obtain, or hold, *in virtue of* anything else."[195] If I interpret him correctly, part of Child's motivation for holding this view is his refusal to assign priority to properties, and the relations between them, as somehow shaping the aggregate of concrete particulars into a causally ordered series.[196]

---

[193] Child, *Causality, Interpretation and the Mind*, 189.
[194] Child, *Causality, Interpretation and the Mind*, 189.
[195] Child, *Causality, Interpretation and the Mind*, 189.
[196] Child draws an analogy to the relation of temporal precedence (instead of *weighing less than*) to illustrate this point about natural relations generally: "When *a* precedes *b*, that temporal relation does not hold in virtue of anything else more basic; its holding is itself a basic fact" (Child, *Causality, Interpretation and the Mind*, 189).

It is interesting to note that Frederick Stoutland, one of the first to argue that Davidson is committed to epiphenomenalism, now shares this reluctance. In his review of Child's book, he concurs with Child's verdict, saying, "This seems to me (now) the right response: if events are causally related, there is a (physical) law which the events instantiate, but they are not causally related *in virtue of* instantiating that law."[197]

This approach still allows nomic patterns to be significant as indicators of a causal connection between two events. After all, even though it is false that two events are causally related to each other in virtue of some nomic correlation between their respective properties, it remains the case that whenever two events are causally interrelated, some of the cause's features will in fact be nomically linked to some of the effect's features. We can, then, still use nomic correlations as a guide to locating causal connections by taking these correlations as signs that indicate the presence of a causal relationship between two events. That is, we can rely on them as symptoms, but not sources, of causal relations between particulars. But if

---

[197] Frederick Stoutland, critical notice of *Causality, Interpretation and Mind*, by William Child, *Philosophy and Phenomenological Research* 58 (1998): 711-15 (at 713). (Emphasis in the original) Stoutland retains his old view, similar to Melden's, that reasons are not causes. His position has only changed in respect of Davidson's alleged commitment to epiphenomenalism.

nomic connections between properties are one kind of upshot of causation, there surely are others as well. For instance, a counterfactual dependency can indicate the presence of a causal connection, and there seems to be no reason to deny that the rationally linked features of beliefs, desires and actions are not in their own way indicators of a causal connection between these events and states. Of course, these latter "sense-making" features are high-level properties that only show up (and signal the presence of causal connections) in those highly complex systems towards which it is fruitful to adopt the intentional stance. But there seems to be no good reason for denying that some of the symptoms of causal connections are only to be found in some, but not all, of the cases in which concrete particulars are causally interrelated, that is, that some of the relations between properties that indicate the presence of a causal link need not be on hand every time there is such a connection.

Here, then, is room for a robust pluralism in the spirit of Davidson. For this kind of causal significance (by way of indicating the presence of causal connections) gives no hint of being a scarce commodity over which properties must compete; that is, there is no reason to suppose that it lends itself to any of the "exclusion" worries. After all,

why cannot it be the case that *several* of the cause's properties are indicative or symptomatic of the causal relation that obtains between their bearer and some other event? Why should one indicator of a causal connection exclude any other? Clearly there is room for a causal relation to be simultaneously made evident by a plethora of relations between the properties of the cause and effect, relations which can reliably be taken to be signs of a causal connection.

In conclusion, the explanatory pluralism envisaged by James, Stout, Weber, Hornsby, Baker and Marras is attainable without having to demarcate a special realm of explananda which are held to be the exclusive preserve of intentional explanations. If we take the above perspective, from which physical properties no longer appear to be privileged explainers that alone can lay claim to causal significance, we thereby remove the motivation for trying to provide mental features with some explanatory work to do by privileging *them* with respect to some narrow range of explananda. Mental properties can partake of causal-explanatory relevance without our having to hit upon some non-physical explananda as their proprietary concern.

VI. Context Is Everything

It will be noticed that the weak form of causal significance described in the previous section is enjoyed by dispositions. True enough, metaphysically speaking, dispositional features really are on a par with mental and physical properties. When we get beyond dispositional features to consider their lower-level, realising properties we do not thereby get closer to *the cause*, for the cause is the event, not any of its properties. Instead, we approach lower-level properties that carry their own new (to us) informational load, where this new information affords us a greater opportunity for explanation, prediction and control. So, quite often, lower-level physical and mental properties do offer us something more than merely dispositional features offer, but this "something more" is not metaphysical. Instead, it is only to be understood in terms of pragmatics, for it can only be cashed out in terms of explanatory relevance, a relation that entails no exclusion principle.

Let us examine a little more closely the informational impoverishment of appeals to dispositions. We tend to feel dissatisfied with appeals to dispositional properties because they seldom provide us with a basis for abstracting sufficiently far from the present context to allow for the

formulation of new and interesting predictions and counterfactuals. For example, when Moliere's charlatan doctor tries to account for the tendency of opium to induce sleep by saying that it has a dormitive virtue, he at least succeeds (as has been noted) in locating part of the cause. He also gives us an indication of where to look in order to discern some of the properties that we are likely to find informative (in the light of the explanatory and predictive tasks at hand). He does not, though, give us an explanation that enables us to depart very far from the actual circumstances when we set out to infer new predictions and counterfactuals. On the basis of his claim that this powder has a disposition to put people to sleep, we may only conclude that people who take *this* powder will fall asleep, and that if I had ingested *this* powder, I too would have fallen asleep. Thus, his explanation only tells us what *this particular* (or something exactly similar to it) is capable of doing. It does not allow us to infer that *anything else*, anything that differs from the opium in some ways but that resembles it in a crucial respect, can also be expected to put normal human beings to sleep. That is to say, it does not give us the power to abstract away from this particular lump of powder (the actual cause) in our subsequent predictions and explanations. The charlatan, then, is a

quack not because he has failed to adduce some previously unknown metaphysical connection between the powder and its effect, something called "causal relevance". Rather, he is a quack because his explanation carries no novel predictive and explanatory power.

As was earlier noted, the charlatan's explanation merely repeats information that was already in (or presupposed by) the question he was trying to answer. Thus, since the why-question that sets the standard for evaluating his answer is a feature of the context, context is everything. So (as was suggested in connection with Rey's discussion of the Moliere case) if the question put to the charlatan had not been, "Why does opium put people to sleep?" but rather, "Why have these people fallen asleep?", then his answer would not have been so clearly vacuous, for it would have imparted some new information (viz., that the opium had something to do with their falling asleep). As Rey suggests, it would have been a perfectly good answer if our background knowledge had not included the claim that people generally do get sleepy after taking opium, so that (for all we know) these people might have fallen asleep due to an allergic reaction to the powder. Or consider the claim that the cord stretched because it was elastic. In some contexts this will be a relatively uninformative explanation, but in

others it will carry a more salient informational load. If we do not already know that the cord is elastic, then we may find it a perfectly good explanation to say that the cord did not hold the broom upright because it was elastic.[198] Here, the answer does broaden our explanatory and predictive horizons, for we now know something we did not know beforehand, namely, that any significant force exerted on the cord is not likely to meet with any resistance.

In short, the "something more" that lower-level physical and mental features usually have, and that dispositional properties so often lack, is a creature of the epistemology and pragmatics of explanation, not metaphysics. This is suggested by the observation that whether a property has this "something more" is a matter of contextual considerations having especially to do with the mind of the inquirer, both with the nature of her why-question and the extent of her background knowledge. This explains why LePore and Loewer failed to capture the "something more" by means of their four-pronged counterfactual test: they mistakenly believed the "something more" to be a metaphysical relation (such that a certain effect can be said to have been caused *in virtue of* a given property of the cause), and accordingly

---

[198] My thanks to John King-Farlow for this example, and for pressing this point generally in John King-Farlow, commentary on "Does Anything Break Because It Is Fragile?" by Paul Raymont (paper presented at the annual meeting of the Canadian Philosophical Association, Ottawa, Ont., June 1998).

formulated a test that would track that connection; but dispositional properties are metaphysically on the same footing as mental (and basic physical) features; so, naturally enough, dispositional properties satisfied the four-pronged test. The problem remains even if we try to recast the counterfactual test as a test only of an intensional relation of explanatory relevance, for dispositions would still pass it and, consequently, the "something more" would still elude it. Thus, we would still lack a test that can be applied in all contexts as a gauge of the sort of causal relevance that we take mental properties to have.[199]

VII. Conclusion

In a recent paper Ned Block objects to Jaegwon Kim's assimilation of efficacious properties with projectible kinds.[200] According to Block, projectibility is relative, for we must always ask, "Projectible with respect to what type of property?"[201] Moreover, "Kindhood comes in degrees,"[202] for it is based on similarity, which admits of degrees. Thus, since whether something is a kind depends on whether it is

---

[199] Thus, William Child, though he appeals to counterfactuals and supervenience as indicators of causal-explanatory relevance relations, ultimately gives up trying to formulate a set of sufficient conditions for causal-explanatory relevance (Child, *Causality, Interpretation and the Mind*, 213).
[200] Block, "Anti-Reductionism Slaps Back," 129.
[201] Block, "Anti-Reductionism Slaps Back," 128.
[202] Block, "Anti-Reductionism Slaps Back," 128.

projectible, kindhood is "relative and graded".[203] Block poses this as a problem for Kim. After all, asks Block, "How could causal efficacy be relative and graded?"[204] He continues, "Once one agrees that the notion of kind is relative and graded, unless one is prepared to see causation as relative and graded, kinds will be poor candidates for the key to causation."[205] True, kinds will be poor candidates for the key to causation, but only because all properties are, if by "key to causation" we mean that in virtue of which causal relations obtain. Causal efficacy is a different matter, though. In this passage Block, like Cynthia Macdonald and John Heil, all too easily slides from premisses about causation to conclusions about causal efficacy. While we can agree that causation is not relative and graded, it is not at all clear that the same can be said of causal efficacy. For if we are speaking of the efficacy of properties (and it is clear from the context that this is what Block has in mind), then we can only be speaking of an intensional relation of explanatory relevance, and this seems a clear case of something that is context-sensitive, and which it is perfectly natural to describe as relative and graded.

---

[203] Block, "Anti-Reductionism Slaps Back," 128.
[204] Block, "Anti-Reductionism Slaps Back," 129.
[205] Block, "Anti-Reductionism Slaps Back," 129.

This passage from Block's paper illustrates the current tendency to glide back and forth between causation and causal relevance (or "efficacy") as though they were one and the same thing. While it may be tempting to do this, we can resist the temptation by keeping clear about the vast difference between the relata that stand in these two relations. Causal relevance is just causal-explanatory relevance, and is thus intensional and context-sensitive. Only properties may enter into this relation. By contrast, only concrete particulars may enter into the natural, extensional relation of causation, and thus enjoy real efficacy. If we grasp the full implications of this difference, we shall see that properties (even basic physical features) are not really efficacious at all. They do not *make* anything happen. In view of this, the properties of basic physics can be seen not to occupy some privileged position, by alone possessing a strange sort of causal efficacy, by comparison with which mental (and other) properties appear to be causally deficient. Once we dispense with the belief that properties can enjoy any real efficacy, by somehow making causes cause whatever they cause, we will find in anomalous monism no commitment to any form of epiphenomenalism. More generally, we shall limit the application of such principles as closure and exclusion to

concrete particulars only. With this restriction comes the realisation of a more accommodating pluralism with respect to properties: a property may explain in one context without thereby excluding other properties from taking on an explanatory relevance as robust as any property is capable of enjoying.

Bibliography

Antony, Louise. "The Causal Relevance of the Mental: More on the Mattering of Minds." *Mind and Language* 6 (1991): 295-327.

Aristotle. *Categories*. In *Aristotle: Selected Works*, translated by Hippocrates G. Apostle and Lloyd P. Gerson. Grinnell, Iowa: The Peripatetic Press, 1982.

Armstrong, D.M. *A Theory of Universals*. Cambridge: Cambridge University Press, 1978.

Armstrong, D.M. *A World of States of Affairs*. Cambridge: Cambridge University Press, 1997.

Baker, Lynne Rudder. *Explaining Attitudes*. Cambridge: Cambridge University Press, 1995.

Block, Ned. "Can the Mind Change the World?" In *Meaning and Method: Essays in Honor of Hilary Putnam*, edited by George Boolos. Cambridge: Cambridge University Press, 1990.

Block, Ned. "Anti-Reductionism Slaps Back." *Philosophical Perspectives* 11 (1997): 107-32.

Braddon-Mitchell, David, with Frank Jackson. *The Philosophy of Mind and Cognition*. Oxford: Blackwell Publishers, 1996.

Broad, C.D. *The Mind and Its Place in Nature*. London: Routledge & Kegan Paul Ltd., 1925.

Broad, C.D. "Replies." In *The Philosophy of C.D. Broad,* edited by Paul Arthur Schilpp. New York: Tudor Publishing Company, 1959.

Brown, James Robert. Commentary on "Does Anything Break Because It Is Fragile?" by Paul Raymont, paper presented at the annual meeting of the Ontario Philosophical Society, Toronto, Ont., October, 1997.

Burge, Tyler. "Individualism and the Mental." *Midwest Studies in Philosophy* 4 (1979): 73-121.

Burge, Tyler. "Individuation and Causation in Psychology." *Pacific Philosophical Quarterly* 70 (1989): 303-22.

Child, William. *Causality, Interpretation and the Mind.* Oxford: Clarendon Press, 1994.

Chisholm, Roderick. *Perceiving*. Ithaca, NY: Cornell University Press, 1957.

Cohen, L. Jonathan, with Mary Hesse, eds. *Applications of Inductive Logic*. Oxford: Clarendon Press, 1980.

Crane, Tim. "The Mental Causation Debate." *Proceedings of the Aristotelian Society* (Suppl.) 69 (1995): 211-36.

Davidson, Donald. "Mental Events." In *Essays on Actions and Events*. Oxford: Clarendon Press, 1980. Originally published in Lawrence Foster and J.W. Swanson, eds., *Experience and Theory* (Amherst, MA: The University of Massachusetts Press, 1970).

Davidson, Donald. "Thinking Causes." In *Mental Causation*, edited by John Heil and Alfred Mele. Oxford: Clarendon Press, 1993.

Dretske, Fred. *Explaining Behavior: Reasons in a World of Causes*. Cambridge, MA: MIT Press, 1988.

Dretske, Fred. "Reasons and Causes." *Philosophical Perspectives* 3 (1989): 1-15.

Edwards, Paul, ed. *Immortality*. Amherst, NY: Prometheus Books, 1997.

Ehring, Douglas. "Mental Causation, Determinables and Property Instances." *Nous* 30 (1996): 461-80.

Elliot, Hugh S.R. *Modern Science and the Illusions of Professor Bergson*. London: Longmans, Green, and Co., 1912.

Ewing, A.C. *Value and Reality*. London: George Allen & Unwin Ltd., 1973.

Geach, Peter. *Mental Acts*. London: Routledge and Kegan Paul, 1957.

Godfrey-Smith, Peter. *Complexity and the Function of Mind in Nature*. Cambridge: Cambridge University Press, 1996.

Hebb, D.O. *Essay on Mind*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers, 1980.

Heil, John, with Alfred Mele. "Mental Causes." *American Philosophical Quarterly* 28 (1991): 61-71.

Heil, John. *The Nature of True Minds*. Cambridge: Cambridge University Press, 1992.

Heil, John, with Alfred Mele, eds. *Mental Causation*. Oxford: Clarendon Press, 1993.

Heil, John. *Philosophy of Mind*. London: Routledge, 1998.

Hess, Peter. "Actions, Reasons, and Humean Causes." *Analysis* XLI (1981): 77-81.

Hesse, Mary, with L. Jonathan Cohen, eds. *Applications of Inductive Logic*. Oxford: Clarendon Press, 1980.

Hirsch, Eli. *Dividing Reality*. New York: Oxford University Press, 1993.

Hollis, Martin. "Philosophy of Social Science." In *The Blackwell Companion to Philosophy*, edited by Nicholas Bunnin and E.P. Tsui-James. Oxford: Blackwell, 1996.

Honderich, Ted. "The Argument for Anomalous Monism." *Analysis* XLII (1982): 59-64.

Horgan, Terence. "Jackson on Physical Information and Qualia." *Philosophical Quarterly* 34 (1984): 147-53.

Horgan, Terence. "Mental Quausation." *Philosophical Perspectives* 3 (1989): 47-76.

Horgan, Terence. "From Supervenience to Superdupervenience." *Mind* 102 (1993): 555-86.

Hornsby, Jennifer. "Agency and Causal Explanation." In *Mental Causation*, edited by John Heil and Alfred Mele. Oxford: Clarendon Press, 1993.

Huxley, T.H. "On the Hypothesis That Animals Are Automata." In *The Philosophy of Mind*, edited by Brian Beakley and Peter Ludlow. Cambridge, MA: the MIT Press, 1992. Originally published in 1874.

Jackson, Frank, with Elizabeth Prior and Robert Pargetter. "Three Theses About Dispositions." *American Philosophical Quarterly* 19 (1982): 251-7.

Jackson, Frank, with Philip Pettit. "Functionalism and Broad Content." *Mind* 97 (1988): 381-400.

Jackson, Frank, with Philip Pettit. "Program Explanation: a General Perspective." *Analysis* 50 (1990): 107-17.

Jackson, Frank. "Essentialism, Mental Properties and Causation." *Proceedings of the Aristotelian Society* 95 (1995): 253-68.

Jackson, Frank. "Mental Causation." *Mind* 105 (1996): 377-413.

Jackson, Frank. "The Primary Quality View of Color." *Philosophical Perspectives* 10 (1996): 199-219.

Jackson, Frank, with David Braddon-Mitchell. *The Philosophy of Mind and Cognition*. Oxford: Blackwell Publishers, 1996.

Jacob, Pierre. *What Minds Can Do*. Cambridge: Cambridge University Press, 1997.

James, William. *The Principles of Psychology*. Vol. 1. New York: Henry Holt and Company, 1890.

Janaway, Christopher. *Self and World in Schopenhauer's Philosophy*. Oxford: Clarendon Press, 1989.

Kim, Jaegwon. "Concepts of Supervenience." In *Supervenience and Mind*. Cambridge: Cambridge University Press, 1993. Originally published in *Philosophy and Phenomenological Research* 45 (1984): 153-76.

Kim, Jaegwon. "The Myth of Nonreductive Physicalism." In *Supervenience and Mind*. Cambridge: Cambridge University Press, 1993. Originally published in *Proceedings and Addresses of the American Philosophical Association* 63 (1989): 31-47.

Kim, Jaegwon. "Explanatory Exclusion and the Problem of Mental Causation." In *Information, Semantics and Epistemology*, edited by Enrique Villanueva. Oxford: Basil Blackwell, 1990.

Kim, Jaegwon. "Supervenience as a Philosophical Concept." In *Supervenience and Mind*. Cambridge: Cambridge University Press, 1993. Originally published in *Metaphilosophy* 21 (1990): 1-27.

Kim, Jaegwon. "Multiple Realisation and the Metaphysics of Reduction." in *Supervenience and Mind*. Cambridge: Cambridge University Press, 1993. Originally published in *Philosophy and Phenomenological Research* 52 (1992): 1-26.

Kim, Jaegwon. *Supervenience and Mind*. Cambridge: Cambridge University Press, 1993.

Kim, Jaegwon. *Philcsophy of Mind*. Boulder, Colorado: Westview Press, Inc., 1996.

King-Farlow, John. Commentary on "Does Anything Break Because It Is Fragile?" by Paul Raymont, paper presented at the annual meeting of the Canadian Philosophical Association, Ottawa, Ont., June, 1998.

Kneale, William C. "Broad on Mental Events and Epiphenomenalism." In *The Philosophy of C.D. Broad*, edited by Paul Arthur Schilpp. New York: Tudor Publishing Company, 1959.

LePore, Ernest, with Barry Loewer. "Mind Matters." *The Journal of Philosophy* 84 [1987]: 630-42.

Lewis, David. "Causal Explanation." In *Philosophical Papers*, Vol. 2. Oxford: Oxford University Press, 1986.

Lewis, David. "Events." In *Philosophical Papers*, Vol. 2. Oxford: Oxford University Press, 1986.

Lewis, David. "Finkish Dispositions." *The Philosophical Quarterly* 47 (1997): 143-58.

Loewer, Barry, with Ernest LePore. "Mind Matters." *The Journal of Philosophy* 84 [1987]: 630-42.

Macdonald, Cynthia, with Graham Macdonald. "Mental Causes and Explanation of Action." In *Mind, Causation and Action*, edited by L. Stevenson, R. Squires and J. Haldane. Oxford: Basil Blackwell, 1986.

Macdonald, Cynthia. *Mind-Body Identity Theories*. London: Routledge, 1989.

Macdonald, Cynthia, with Graham Macdonald. "How to be Psychologically Relevant." In *Philosophy of Psychology*, Vol. 1 of *Debates on Psychological Explanation*, edited by Cynthia Macdonald and Graham Macdonald. Oxford: Basil Blackwell, 1995.

Macdonald, Graham, with Cynthia Macdonald. "Mental Causes and Explanation of Action." In *Mind, Causation and Action*, edited by L. Stevenson, R. Squires and J. Haldane. Oxford: Basil Blackwell, 1986.

Macdonald, Graham, with Cynthia Macdonald. "How to be Psychologically Relevant." In *Philosophy of Psychology*, Vol. 1 of *Debates on Psychological Explanation*, edited by Cynthia Macdonald and Graham Macdonald. Oxford: Basil Blackwell, 1995.

MacDougall, William, with John B. Watson. *The Battle of Behaviorism*. New York: W.W. Norton & Company, Inc., 1929.

McLaughlin, Brian P. "Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical." *Philosophical Perspectives* 3 (1989): 109-35.

McLaughlin, Brian P. "On Davidson's Response to the Charge of Epiphenomenalism." In *Mental Causation*, edited by John Heil and Alfred Mele. Oxford: Clarendon Press, 1993.

Marras, Ausonio. "The Causal Relevance of Mental Properties." *Philosophia* 25 (1997): 389-400.

Melden, A.I. *Free Action*. London: Routledge and Kegan Paul, 1961.

Mele, Alfred, with John Heil. "Mental Causes." *American Philosophical Quarterly* 28 (1991): 61-71.

Mele, Alfred, with John Heil, eds. *Mental Causation*. Oxford: Clarendon Press, 1993.

Moliere. *The Imaginary Invalid (Le malade imaginaire)*. In *The Dramatic Works of Moliere*, Vol. 3, translated by Charles Heron Wall. London: George Bell and Sons, 1908.

Nagel, Thomas. *The View From Nowhere*. Oxford: Oxford University Press, 1986.

Nietzsche, Friedrich. *Daybreak*, translated by R.J. Hollingdale, edited by Maudemarie Clark and Brian Leiter. Cambridge: Cambridge University Press, 1997. Originally published in 1881.

Noordhof, Paul. "Do Tropes Resolve the Problem of Mental Causation?" *The Philosophical Quarterly* 48 (1998): 221-26.

Owens, David. "Disjunctive Laws?" *Analysis* 49 (1989): 197-202.

Pargetter, Robert, with Elizabeth Prior and Frank Jackson. "Three Theses About Dispositions." *American Philosophical Quarterly* 19 (1982): 251-7.

Patton, Michael. Commentary on "Does Anything Break Because It Is Fragile?" by Paul Raymont, paper presented at the annual Mid-South Conference in Philosophy, Memphis, TN, February, 1998.

Pettit, Philip, with Frank Jackson. "Functionalism and Broad Content." *Mind* 97 (1988): 381-400.

Pettit, Philip, with Frank Jackson. "Program Explanation: a General Perspective." *Analysis* 50 (1990): 107-17.

Popper, Karl. "Autobiography." In *The Philosophy of Karl Popper*, edited by Paul Arthur Schillp. La Salle: Open Court, 1974.

Prior, Elizabeth, with Robert Pargetter and Frank Jackson. "Three Theses About Dispositions." *American Philosophical Quarterly* 19 (1982): 251-7.

Putnam, Hilary. "The Nature of Mental States." In *Mind, Language and Reality*. Cambridge: Cambridge University Press, 1975. Originally published as "Psychological Predicates," in W.H. Capitan and D.D. Merrill, eds., *Art, Mind, and Religion* (Pittsburgh: University of Pittsburgh Press, 1967).

Putnam, Hilary. "The Meaning of `Meaning'." In *Mind, Language and Reality*. Cambridge: Cambridge University Press, 1975. Originally published in Keith Gunderson, ed., *Language, Mind, and Knowledge* (Minneapolis: University of Minnesota Press, 1975).

Rey, Georges. *Contemporary Philosophy of Mind*. Oxford: Blackwell Publishers, 1997.

Robb, David. "The Properties of Mental Causation." *The Philosophical Quarterly* 47 (1997): 178-94.

Ruben, David-Hillel. *Explaining Explanation*. London: Routledge, 1990.

Santayana, George. "Reason in Common Sense." In *The Life of Reason*. New York: Charles Scribner's Sons, 1905.

Santayana, George. "The Efficacy of Thought." *The Journal of Philosophy* 3 (1906): 410-12.

Santayana, George. *Realms of Being*. New York: Charles Scribner's Sons, 1940.

Santayana, George. "*Apologia Pro Mente Sua*." In *The Philosophy of George Santayana*, edited by Paul Arthur Schilpp. New York: Tudor Publishing Company, 1940.

Schilpp, Paul Arthur, ed. *The Philosophy of George Santayana*. New York: Tudor Publishing Company, 1940.

Schilpp, Paul Arthur, ed. *The Philosophy of C.D. Broad*. New York: Tudor Publishing Company, 1959.

Schillp, Paul Arthur, ed. *The Philosophy of Karl Popper*. La Salle: Open Court, 1974.

Schopenhauer, Arthur. *The World as Will and Representation*, translated by E.F.J. Payne. Vol. 2. 1958. Reprint, New York: Dover, 1969. Originally published in 1818.

Seager, William. "Disjunctive Laws and Supervenience." *Analysis* 51 (1991): 93-8.

Searle, John. *Intentionality*. Cambridge: Cambridge University Press, 1983.

Shoemaker, Sydney. "Causality and Properties." In *Time and Cause*, edited by Peter van Inwagen. Dordrecht: D. Reidel Publishing Company, 1980.

Shoemaker, Sydney. "Properties, Causation, and Projectibility." In *Applications of Inductive* Logic,

edited by L. Jonathan Cohen and Mary Hesse. Oxford: Clarendon Press, 1980.

Sosa, Ernest. "Mind-Body Interaction and Supervenient Causation." *Midwest Studies in Philosophy* 9 (1984): 271-81.

Spencer, Herbert. *First Principles*. 6th ed. New York: D. Appleton and Company, 1900.

Stout, G.F. *A Manual of Psychology*. London: W.B. Clive, 1899.

Stoutland, Frederick. "The Causation of Behavior." In *Essays on Wittgenstein in Honor of G.H. von Wright, Acta Philosophica Fennica*, XXVIII. Amsterdam: North-Holland, 1976.

Stoutland, Frederick. Critical notice of *Causality, Interpretation and Mind*, by William Child. *Philosophy and Phenomenological Research* 58 (1998): 711-15.

Strawson, P.F. "Causation and Explanation." In *Essays on Davidson*, edited by Bruce Vermazen and Merrill Hintikka. Oxford: Oxford University Press, 1985.

Swinburne, Richard. "Properties, Causation, and Projectibility: Reply to Shoemaker." In *Applications of Inductive Logic*, edited L. Jonathan Cohen and Mary Hesse. Oxford: Clarendon Press, 1980.

Swoyer, Chris. "The Nature of Natural Laws." *The Australasian Journal of Philosophy* 60 (1982): 203-23.

Vivas, Eliseo. "From *The Life of Reason* to *The Last Puritan*." In *The Philosophy of George Santayana*, edited by Paul Arthur Schilpp. New York: Tudor Publishing Company, 1940.

Watson, John B., with William MacDougall. *The Battle of Behaviorism*. New York: W.W. Norton & Company, Inc., 1929.

Weber, Max. "Basic Sociological Terms." Original translation by Talcott Parsons (1947) (revised by Guenther Roth and Claus Wittich). In *Economy and Society*, edited by Guenther Roth and Claus Wittich. New York: Bedminster Press, 1968. Originally published in 1922.

Wittgenstein, Ludwig. *Philosophical Investigations*. Oxford: Basil Blackwell, 1953.

Wittgenstein, Ludwig. *Zettel*, translated by G.E.M. Anscombe, edited by G.E.M. Anscombe and G.H. von Wright. Berkeley: University of California Press, 1967.

Wittgenstein, Ludwig. "Lectures on Freedom of the Will." From notes taken by Yorick Smythies. In *Ludwig Wittgenstein: Philosophical Occasions 1912-1951*, edited by James C. Klagge and Alfred Nordmann. Indianapolis: Hackett Publishing Co., 1993.

Yablo, Stephen. "Mental Causation." *The Philosophical Review* 101 (1992): 245-80.