

## The Epistemological Approach to Mental Causation

Prof. Dr. Sven Walter  
Institute of Cognitive Science  
University of Osnabrueck  
Albrechtstraße 28  
D-49076 Osnabrueck  
Germany

phone: ++49 541 969 3360

fax: +49 541 969 3381

email: s.walter@philosophy-online.de

### Abstract

Epistemological approaches to mental causation argue that the notorious problem of mental causation as captured in the question “How can irreducible, physically realized, and potentially relational mental properties be causally efficacious in the production of physical effects?” has a very simple solution: One merely has to abandon any meta-physical considerations in favor of epistemological considerations and accept that our explanatory practice is a much better guide to causal relevance than the metaphysical reasoning carried out from the philosophical armchair. I argue that the epistemological approach to mental causation does not enjoy any genuine advantage over theories which treat the problem of mental causation as a genuinely metaphysical problem.

### 1. Mental Causation

Mental causation, our mind’s ability to make a causal difference to the course of the world, is part and parcel of our ‘manifest image’ of the world. We take ourselves to be freely deliberating agents that do what they do because they have the beliefs and desires they do. That there is mental causation seems obvious. How there can be such a thing as mental causation, however, is not at all obvious.

What is needed and what philosophers since Descartes have attempted to provide is an account of how the mind fits into the causal structure of an otherwise physical world such as to be able to exert a genuine causal influence. Providing such an account is one of the most daunting tasks in current philosophy of mind, given that there are a series of arguments to the conclusion that a fundamentally non-physical mind can not make a difference to the causal course of the world.<sup>1</sup> On the other hand, most of us find epiphenomenalism—the claim that the mind is merely a causally otiose by-product of physical processes—hideous and repellent. Without mental causation, it seems, we would not be the kinds of beings we take ourselves to be, and we would be unable to occupy the place in the world we take ourselves to occupy (see Walter, 2006b, ch. 6.3; 2007). We thus have to resolve the quandary between our yearning for the causal effectiveness of our mind on the one and our lack of understanding of how it could be causally effective, combined with a series of arguments to the extent that ours apparently is not a world in which there is mental causation on the other hand.

Some hold that this problem has a strikingly simple solution. The arguments that allegedly demonstrate the impossibility of mental causation (see note i), they say, are not indicative of a deep philosophical problem but result from a sort of metaphysical paranoia. To rid ourselves of the specter of epiphenomenalism, we only have to stop lending too much weight to metaphysical armchair arguments and instead base our confidence in our mind's causal effectiveness on our explanatory practice. According to this view, the metaphysical arguments that seem to threaten mental causation are “symptomatic of a mistaken set of priorities” (Burge, 1993, p. 97). Were we to appreciate the practically indispensable role the attribution of mental states plays in our everyday life and the sciences, the problem of mental causation would immediately ‘melt away.’

Such a deflationary perspective on mental causation is only rarely explicitly defended in print, but one encounters it frequently when one tries to convey the importance of the problem of mental causation to philosophers not primarily concerned with the philosophy of mind, to colleagues from outside the philosophy department, or to philosophical laymen, for they often hold that there cannot be a serious problem with mental causation simply because mentalistic explanations quite obviously are successful and indispensable.

Section 3 will argue that the move from a metaphysical to an epistemological level of analysis does not add anything to our understanding of mental causation. Before that, section 2 will say a bit more about the view at issue.

## 2. The Epistemological Approach to Mental Causation

The epistemological approach to mental causation assumes that the causal efficacy of our mind can be established independently of and prior to any of the metaphysical considerations that make mental causation look suspicious in the first place. That mental properties are relational properties, that they fail to supervene upon the local make-up of a causal system and that the physical is a self-sufficient causally closed system in which any appeal to non-physical causal properties appears to be unnecessary does not undermine the fact that many of our explanations treat mental properties as causally efficacious. The point is simple enough: Since the “probity of mentalistic causal explanation is deeper than the metaphysical considerations that call it into question” (Burge, 1993, p. 115-116), epistemology always trumps metaphysics and no armchair argument can ever establish epiphenomenalism as a serious option:

My suggestion is to take as our philosophical starting-point, not a metaphysical doctrine about the nature of causation or of reality, but a range of explanations that have been found worthy of acceptance. These include, pre-eminently but not exclusively, scientific explanations. They also include commonplace explanations that explain the phenomena that we encounter in everyday life ... If we put aside the metaphysical picture and begin with the explanations that work, causation becomes an explanatory concept. (Baker, 1993, p. 92-93; emphasis S.W.)

Since Lynne Rudder Baker has offered the most elaborate defense of the epistemological approach, I will concentrate on her view, making some passing remarks about other proponents along the way.

Baker criticizes what she calls the ‘Standard View,’ a generic position comprising functionalism, eliminativism, and various type- and token-identity theories, characterized by the assumption that mental states or properties are identical to or constituted

by brain states or properties of the brain. In order for an account of mental causation to be acceptable to defenders of the Standard View, it must conform to a “comprehensive metaphysical theory to which they have prior commitment” (Baker, 1995, p. 19), and if their metaphysical theory seems to require it, they even take a position like epiphenomenalism seriously. For Baker, in contrast, there can be no question that it is the metaphysical theory that has to go in such a case. Whenever “successful explanatory practice is pitted against a priori conditions of adequacy,” Baker maintains, “it seems to me prudent to stick with what works—the successful explanatory practices—and to forsake the a priori conditions of adequacy” (Baker, 1995, p. 119).

According to Baker, the Standard View comprises three claims: (1.) human behavior can be understood as the occurrence of behavioral events (i.e. actions) with causes (i.e. beliefs and desires) which are internal to the agent (Baker, 1993, p. 76); (2.) “causation is ... ‘objective,’ a ‘real relation’ in nature, instances of which are independent of anyone’s explanatory interests” (Baker, 1993, p. 75); and (3.) “all reality depends on physical reality, where physical reality consists of a network of events” (Baker, 1993, p. 75). (3.) includes the claims that (3a.) the physical realm is causally closed and that (3b.) every instantiation of every property supervenies upon the instantiation of a micro-physical property or a set of micro-physical properties (Baker, 1993, p. 79). These assumptions play a crucial role in the arguments that allegedly threaten the causal efficacy of the mental. And since (at least some of) these arguments can be generalized to show that the chemical, biological, geological etc. is equally otiose, “the metaphysical assumptions with which we began inevitably lead to skepticism ... about macro-level causation generally” (Baker, 1993, p. 77).

Since it would be ridiculous to give up nearly all the explanations we have ever offered for anything, Baker argues, this amounts to a reductio of the Standard View (Baker, 1995, p. 28). Since it is untenable that the metaphysical assumptions of the Standard View undermine our explanatory practice beyond recognition, Modus Tollens requires that the claims expressed in (1.), (2.) and (3.) cannot (all) be correct. In particular, Baker suggests replacing (2.) with a conception of causation more in tune with Practical Realism.<sup>ii</sup> Her suggestion is to reverse the priorities of epistemology and metaphysics and understand causation in terms of explanation: “it seems methodologically mis-

guided to begin with a Standard View metaphysics that precludes the causal explanations we want, need, and are willing to pay millions to find. It is better to start with explanations that work” (Baker, 1995, p. 121):

If we put aside the metaphysical picture and begin with the explanations that work, causation becomes an explanatory concept. ... If we reverse the priority of explanation and causation that is favored by the metaphysician, the problem of mental causation just melts away. (Baker, 1993, p. 93)

Whether a property is causally efficacious or not is hence determined by whether it figures in the explanations that make up our explanatory practice. Simply put: Causal efficacy reduces to explanatory relevance. And there can be no doubt that most macrolevel properties, including mental properties, do have the required explanatory potential: We write bills, fill in revenue forms, lie, hold electoral addresses, read instruction manuals, author books, buy theft protection devices and gossip because by doing so we can, via a manipulation of the attitudes of our fellow human beings, systematically manipulate their behavior. According to the epistemological approach, it is this explanatory connection that makes the properties in question causally efficacious, not any abstract philosophical reasoning.

Apart from Baker, the most prominent proponents of the epistemological approach are Tyler Burge (see Burge, 1993) and Frank Jackson and Philip Pettit (see Jackson & Pettit, 1988, 1990a, 1990b, 1992). According to the latter’s program explanation account, for instance, mental and other macrolevel properties are not causally efficacious in the same sense as basic physical properties are causally efficacious, but they nevertheless do make a causal difference in virtue of figuring in so-called ‘program explanations’ which convey information not conveyed by explanations in terms of the causally efficacious physical properties (see Walter, 2005).<sup>iii</sup>

The key idea common to all these views is that the problem of mental causation can be solved by moving the discussion from a causal to an explanatory level at which the problem allegedly does not arise. The next section argues that this idea is misguided.

### 3. The Epistemological Approach to Mental Causation: A Problem

The epistemological approach makes two important claims:

- (1) Causal efficacy can be reduced to explanatory relevance.
- (2) Mental (and other macrolevel) properties are explanatorily relevant.

Jaegwon Kim denies (1). According to him, causal efficacy cannot be reduced to explanatory relevance because even if it is granted that macrolevel properties do play a vital explanatory role, their causal status remains problematic for the reasons stated in the metaphysical arguments of the Standard View, in particular his Argument from Causal Exclusion:

[T]he presence of two causal stories, each claiming to offer a full causal account of a given event, creates an unstable situation requiring us to find an account of how the two purported causes are related to each other. ... That the explanations arise in different areas of inquiry ... or that they are responses to different epistemic or pragmatic concerns, makes no difference. ... [The problem] arises from the very notion of causal explanation and what strikes me as a perfectly intuitive and ordinary understanding of the causal relation. If this is right, turning away from metaphysics to embrace epistemology, or away from causation to embrace explanation, will not dissipate the need for an account of mental causation. (Kim, 1998, p. 64-67; emphasis S.W.)

However, I cannot see how Kim can insist that there is a 'need for an account of mental causation' over and above the fact that mental properties are explanatorily relevant without thereby begging the question against the epistemological approach, given that the latter precisely claims that causation just is an explanatory concept, which leaves no room for any problem of mental causation over and above the explanatory relevance of mental properties (see Walter, 2006a).

Prima vista, (2) seems unquestionable. As proponents of the epistemological approach point out correctly, it would be absurd to deny that mental and other macrolevel properties are explanatorily relevant, given that a lot, if not most, of our most successful explanations appeal to mental and other macrolevel properties.

I do not deny that (1) and (2) can both be true. What I deny is that one can explicate, in purely epistemological terms, a notion of explanatory relevance strong enough to vindicate our confidence in the causal efficacy of mental and other macrolevel properties. In purely epistemological terms, one can only explicate a notion of explanatory relevance according to which, although (2) is unquestionable, (1) is false. If one wants a more substantial notion of explanatory relevance, (1) and (2) may be true, but one then has to invoke notions which aren't in a clear and straightforward sense 'merely epistemological' since they can be employed equally well by someone who would in Baker's sense endorse the metaphysical assumptions of the Standard View. Either way, the central tenet of the epistemological approach—'Causal efficacy reduces to explanatory relevance'—does not, by itself, contribute to our understanding of mental causation.

What does the claim that a property is causally efficacious in virtue of its playing an indispensable role in our 'successful explanatory practice' mean? If we take this to mean that every property that figures in a statement de facto offered and accepted as an explanation possesses the right kind of explanatory relevance (and thus is thereby ipso facto causally efficacious), then (2) can hardly be denied. But that is certainly not what the proponent of the epistemological approach has in mind. Apart from other reasons why such a notion of explanatory relevance would be strange, (1) would be false under this reading of 'explanatory relevance.' Otherwise containing phlogiston would have been a property causally responsible for the combustibility of, say, wood when combustion processes were explained by appeal to phlogiston, but ceased to be so when the oxygen theory of combustion was discovered and 'explanations' in terms of phlogiston were discarded. Patently, this is nonsense: containing phlogiston was never causally responsible for anything since nothing ever contained phlogiston, and therefore a property's being causally efficacious cannot be reduced simply to its figuring in a statement de facto offered and accepted as an explanation.<sup>iv</sup>

Apart from the existence of such pseudo-explanations, another reason why causal efficacy cannot be reduced to explanatory relevance simpliciter is the existence of non-causal explanations. Two paradigmatic cases that come to mind here are ‘compositional explanations’ involving identities or the realization relation: the water’s boiling can be explained by the fact that some molecules have such-and-such kinetic energy sufficient to overcome the intermolecular binding forces, but that does not render the latter causally responsible for the former—the water’s boiling just is the molecules’ having such-and-such kinetic energy; one cigarette may have the macrolevel property of being carcinogenic in virtue of containing ingredient a<sub>1</sub>, another cigarette may have the same property in virtue of containing ingredient a<sub>2</sub>, and we may explain why these cigarettes are carcinogenic by saying that they contain a<sub>1</sub> and a<sub>2</sub>, respectively, although their containing a<sub>1</sub> and a<sub>2</sub> does not cause their being carcinogenic (realization being a synchronic relation).

Hence, if we try to vindicate the epistemological approach by drawing attention to the explanations that we actually offer in our everyday life and in our scientific practice, (2) is unquestionable, but (1) is false—causal efficacy cannot be reduced to explanatory relevance simpliciter. It is thus important for the proponent of the epistemological approach to clarify the notion of explanatory relevance used in her argument. What is needed is a substantial notion of explanatory relevance according to which the properties cited in pseudo-explanations or non-causal explanations do not qualify as explanatorily relevant. If the epistemological approach is to get off the ground, there thus must be some additional factor that renders some statements, like those in terms of mental properties, capable of grounding the causal efficacy of the properties in question, while pseudo-explanations or non-causal explanations can be part of our explanatory practice without ipso facto rendering the properties in question causally efficacious.

The problem with this is that if such a substantial notion of explanatory relevance, i.e. a notion of explanatory relevance which renders (1) and (2) true, can be spelled out at all, this can be done only by appeal to notions which can equally well be employed by someone who offers a ‘metaphysical’ account of mental causation in line with what Baker calls the Standard View. If correct, this shows that if the epistemological approach



works, it does not work because it is a characteristically epistemological, as opposed to a characteristically metaphysical, account of mental causation.

Since the obvious difference between pseudo-explanations and non-causal explanations on the one and those explanations that the proponent of the epistemological approach has in mind on the other hand is that only the latter are causal explanations, the proponent of the epistemological approach must answer the question 'What renders an explanation a causal explanation?' What the claim that a property is causally efficacious in virtue of its playing an indispensable role in our 'successful explanatory practice' must mean is that a property is causally efficacious for a given effect if it figures in (one of) its causal explanation(s). If that is what is meant by 'explanatorily relevant,' then (1) is presumably true, for it would sound strange to say that a property F figures in a causal explanation of an effect e and yet F is not causally responsible for e. The question, however, is why one should accept (2) on that a reading of 'explanatorily relevant.' Simply to say that our mentalistic explanations obviously are causal explanations would be tantamount to saying that mental properties obviously are causally efficacious, and that is an adequate description of our manifest image of the world, but not a substantial solution to the problem of mental causation.

The proponent of the epistemological approach must therefore supplement her account by a criterion which distinguishes among the statements that make up our explanatory practice those which are causal explanations from those which are pseudo-explanations or non-causal explanations.<sup>v</sup>

Once such a criterion is provided, however, the crucial step in the argument is not the reduction of causal efficacy to explanatory relevance, but the explication of the relevant kind of explanatory relevance by means of the criterion which delineates the notion of a causal explanation. Strictly speaking, the attempt to avoid the pertinent problems with respect to causation by reducing causation to explanation and holding that the pertinent problems do not arise with respect to explanation therefore does not bring us an inch closer to a solution to the problem of mental causation.<sup>vi</sup> Proponents of the Standard View who treat the problem of mental causation as a metaphysical problem must formulate a criterion which distinguishes causally efficacious from causally inefficacious properties and which can be fulfilled by mental properties. Proponents of

the epistemological approach who treat the problem of mental causation as an epistemological problem must formulate a criterion which distinguishes statements that are causal explanations from those which are not and which can be fulfilled by mentalistic explanations. In both cases, the important work is done by the criterion in question. The detour via the epistemological level, by itself, adds nothing. Saying that a property which fulfills a given criterion C possesses the right kind of explanatory relevance and that every property which possesses the right kind of explanatory relevance is ipso facto causally efficacious:

fulfilling criterion C = explanatory relevance of the right kind = causal efficacy,

is only a lengthy and cumbersome way to say that a property which fulfills criterion C is causally efficacious:

fulfilling criterion C = causal efficacy.

The appeal to epistemology does no crucial work in the argument. It ultimately 'cancels out.' The problem is not just that the epistemological approach is incomplete because a criterion which is needed is not provided—the problem is that even if such a criterion is provided, the metaphysical/epistemological-distinction which figures so prominently in the epistemological approach does no work in the proposed solution to the problem of mental causation.<sup>vii</sup>

To assume that the epistemological approach succeeds where a metaphysical approach founders is to assume that there can be a criterion able to delineate the class of causal explanations but unable to delineate the class of causally efficacious properties in the theories of those who endorse the metaphysics of the Standard View. First of all, it is unclear what such a criterion might be. Second, and more importantly, the criteria that people like Baker or Jackson and Pettit have in fact proposed do not seem to make use of any characteristically epistemological features.

Having argued that the problem of mental causation just melts away "[i]f we put aside the metaphysical picture and begin with the explanations that work" (Baker, 1993,

p. 98), Baker goes on to say what exactly it is that renders a property explanatorily relevant, i.e. what exactly it is that makes an explanation an explanation that works. Explanatory relevance, she says, is just a matter of figuring in appropriate counterfactuals that relate causes and their properties to effects (and their properties):

When Jill returns to the bookstore to retrieve her keys, what she thinks is that she left her keys on the counter and that she wants them back. What she thinks affects what she does in virtue of the following explanatory fact: if she hadn't thought that she had left her keys, then, other things being equal, she wouldn't have returned to the bookstore; and given that she did think that she had left her keys, then, other things being equal, her returning was inevitable. (Baker, 1993, p. 93)

Counterfactual dependencies of this kind are Baker's criterion for what it is for a property to possess the right kind of explanatory relevance. True counterfactuals are a reliable guide to explanatory relevance (sive causal efficacy) since one knows one is dealing with a causal explanation when it affords counterfactual control over the event to be explained: "When we can produce or prevent a phenomenon at will, we know that we have found a cause. So if control of some property yields control of some phenomenon, we have a causal explanation of the phenomenon" (Baker, 1995, p. 122; emphasis S.W.). Baker therefore offers the following 'Control Test' (CT) as "a sufficient condition for a causal explanation" (Baker, 1995, p. 121):

Let C be circumstances in which someone can produce or prevent an occurrence of a certain type of phenomenon G, by producing or preventing an occurrence of another type of phenomenon, F, where an occurrence of F is temporally distinct from an occurrence of G and an occurrence of F does not itself entail an occurrence of G. Then:

(CT) An occurrence of F in C causally explains an occurrence of G in C if: (i) If an F had not occurred in C, then a G would not have occurred in C; and (ii) given that an F did occur in C, an occurrence of G was inevitable. (Baker, 1995, p. 122)

Note that CT (correctly) does not count pseudo-explanations in terms of containing phlogiston and paradigmatic cases of non-causal explanations as causal explanations: the former violate CT because it is not true that had something not contained phlogiston, it would not have been combustible, the latter violate CT because in non-causal explanations F's occurrence is not temporally distinct from G's occurrence. CT thus offers a notion of explanatory relevance substantial enough to give (1) at least a chance of being true.<sup>viii</sup> However, if CT works, it does not work because metaphysics is discarded in favor of epistemology, but because the crucial epistemological notion—that of a 'causal explanation'—is explicated in a certain way. And the way this notion is explicated, viz., by appeal to counterfactual dependencies, is entirely compatible with the Standard View. It may be objected that counterfactual dependence seems to be a paradigmatically epistemological notion, so that the fact that the epistemological approach appeals to counterfactual dependencies is not an objection.<sup>ix</sup> Yet, it seems unclear to me that this really is so. Why exactly does something like CT count as an 'epistemological' account of mental causation in Baker's view, while the account offered by someone who talks, say, about 'causal powers,' 'causal inheritance' etc. could be deemed 'metaphysical'? According to Baker, one of the central tenets of the metaphysical background behind the Standard View is the claim that "causation is ... 'objective,' a 'real relation' in nature, instances of which are independent of anyone's explanatory interests" (Baker, 1993, p. 75). But why couldn't someone hold that the counterfactual dependencies CT appeals to are exactly that—objective facts of nature which are independent of anyone's explanatory interests?

The same holds for Jackson and Pettit's program explanation account. The key idea is that a macroproperty which does not figure in the production process leading to an event e can nevertheless be causally responsible (or relevant) for e's occurrence in a weaker sense if its instantiation ensures, or programs for, the instantiation of a physical property which is causally efficacious in e's production. When addressing the obvious question of what makes an explanation a program explanation, Jackson and Pettit eventually argue that program explanations give rise to what they call "invariance of effect under variation of realization" (Jackson & Pettit, 1990b, p. 202):

We can express the basic idea behind a programme explanation in terms of what remains constant under variation. Suppose state a caused state b. Variations on a, say, a', a'', ... would have caused variations on b, say b', b'', ..., respectively. It may be that if the a<sup>i</sup> share a property P, the b<sup>i</sup> would share a property Q: keep P constant among the actual and possible causes, and Q remains constant among the actual and possible effects. ... [I]n such a case P causally explains Q by programming it, even though it may be that P does not produce Q. (Jackson & Pettit, 1988, p. 394)

For the current purpose the details of this suggestion do not matter. The important point is the following: When spelling out their initial suggestion that to make a causal difference is to figure in a program explanation, i.e. when answering the question what makes it the case that 'P causally explains Q by programming it,' there is nothing characteristically epistemological in Jackson and Pettit's explication. Like Baker, they resort to a criterion that could equally well be used by a proponent of the Standard View who holds that all these variances and invariances are objective facts of our world, independent of anyone's explanatory interests, and suggests that a macrolevel property P is causally efficacious for the instantiation of a property Q if it is the case that if we 'keep P constant among the actual and possible causes ... Q remains constant among the actual and possible effects.'<sup>x</sup>

In all these cases, the claim that causal efficacy reduces to explanatory relevance, allegedly the key insight required for a quick and easy solution to the problem of mental causation, does in fact add nothing to a solution. If the epistemological approach works at all,<sup>xi</sup> it does not do so because of some characteristically epistemological features, but only because of features that are entirely compatible with a metaphysical perspective on mental causation. In order to give (1) a chance of being true, the proponent of the epistemological approach thus has to go beyond the purely epistemological. It is therefore misguided to think that the slogan 'epistemology trumps metaphysics' alone could solve the problem of mental causation. Reversing the priority of explanation and causation and making causation an explanatory concept is at best half the battle when it

comes to understanding mental causation. The other half requires answering the question what the explanations are that work, and it is here that notions crop up again—counterfactual dependence, for instance, or invariance of effect under variation of realization—for which it is not clear in what sense they are ‘epistemological’ as opposed to ‘metaphysical’ notions.

If we want to square our manifest image of the world with our conviction that we are fundamentally physical beings in a fundamentally physical world in which physical forces are the only forces, we should concentrate on a joint effort to explicate adequate sufficient conditions that capture our intuitive notion of causal efficacy and that can be fulfilled by mental and other macrolevel properties, rather than engaging in distracting and unnecessary trench battles about the explanatory value of mentalistic explanations and their potential pay-off vis-à-vis the problem of mental causation. Perhaps—probably maybe—a successful solution to the problem of mental causation can be achieved only if epistemology and metaphysics can be carefully ‘attuned’ to each other.<sup>xii</sup> That is, it may turn out that a successful account of mental causation requires us to invoke notions which are both epistemological and metaphysical or notions from the penumbra. That is, it might be that I have no objections to that. I would just interpret it as one more indication of the fact I have stressed throughout this paper: When it comes to a solution to the problem of mental causation, arguing over whether an approach is ‘purely epistemological’ or ‘gorged with metaphysics’ is a mute business.

---

<sup>i</sup> First, causation seems to require laws, while there are grounds for thinking that there are no laws of the appropriate kind connecting the mental and the physical (the Argument from the Anomaly of the Mental; see Davidson, 1970). Second, causation seems to be a local or intrinsic affair, while those aspects constitutive of mental phenomena like beliefs and desires qua mental seem to be relational or extrinsic (the Argument from Anti-Individualism; see Fodor, 1987, ch. 2). Third, we do not understand how the mental can be causally efficacious without conflicting with physical parts of the causal structure that we suspect to play an indispensable causal role in the production of physical effects (the Argument from Causal Exclusion; see Kim, 1998, 2005).

---

<sup>ii</sup> Which is not to say that Baker accepts (1.), (3a.) and (3b.): (3a.) must be given up once (2.) is revised, (3b.) is a “gratuitous bit of metaphysics” (Baker, 1993, p. 91), and since she denies that mental states are brain states, she would presumably reject (1.), too.

<sup>iii</sup> Paul Raymont, another proponent of the epistemological approach, adopts a Davidsonian conception of events as datable, non-repeatable, structured particulars and argues that “on [such] a coarse-grained model of events, the causally relevant properties of an event c are those that help us to explain why c caused an event e,” concluding that “[f]or properties, then, ‘causal relevance’ is wholly a matter of possessing a certain type of explanatory relevance” (Raymont, 2003, p. 234), which allegedly resolves the problem of mental causation (Raymont is concerned with the Argument from Causal Exclusion only): “On this model the causally relevant or ‘efficacious’ properties are just those properties that are salient in the light of the above-mentioned explanatory project (viz., the project of explaining why the cause caused the effect), and we have no reason to believe that these explanatory items compete with or exclude each other” (Raymont, 2003, p. 234; see also Pietroski, 1994 and Thomasson, 1998).

<sup>iv</sup> What about the suggestion that every property that figures in a true statement that is de facto offered and accepted as an explanation is thereby ipso facto causally efficacious? The idea could be that incorrect explanations that appeal to non-existing things like phlogiston will not pass this criterion, so the above counterexample to (1) is avoided (thanks to an anonymous referee for raising this concern). There are two issues in this suggestion that must be kept separate. The first is the idea that ‘real’ explanations may not appeal to non-existing entities, the second is the idea that ‘real’ explanations should be true. The first idea successfully deals with the phlogiston-example, but it does not solve the problem. Suppose I say “Sarah has fever because she has the measles.” Here the ‘explanation’ talks about something that does exist—Sarah’s measles—and still Sarah’s having measles is not causally responsible for her having fever. A proponent of the second idea would agree but would insist that this just shows that “Sarah has fever because she has the measles.” is not true and thus no explanation in the relevant sense. Yet, we know that “Sarah has fever because she has the measles.” is not true only because we know that Sarah’s having measles is not causally responsible for her

---

having fever. And quite obviously, it will not do to say that a property is causally efficacious if it figures in a true explanation, if in order to know whether an explanation is true we first must determine the causal status of the property in question.

<sup>v</sup> The obvious response that causal explanations differ from pseudo-explanations and from non-causal explanations in that they appeal to causally efficacious properties is of course of no avail to the proponent of the epistemological approach, for she cannot hold that a property is causally efficacious in virtue of figuring in a causal explanation, only to add that a causal explanation is an explanation which appeals to a causally efficacious property.

<sup>vi</sup> In fact, this is rather unsurprising: To say that a property is causally efficacious in virtue of figuring in a causal explanation successfully replaces the question ‘Which properties are causally efficacious?’ with the question ‘Which properties are appealed to in causal explanations?’, but it obviously does not provide an answer the question which properties are appealed to in causal explanations.

<sup>vii</sup> Suppose someone advances the hypothesis that  $x$  is an even number just in case  $x$  is a POES. Is that an adequate account of what makes a number even? It all depends on how the class of POESs is characterized. If it turns out that it is characterized in such a way that there clearly are even numbers who do not meet the characterization, then the hypothesis is false. If it turns out that it is characterized in such a way that all and only the numbers that are divisible by two are members of the class of POESs, then the account may be extensionally adequate, but the notion of a POES does no real work. Saying that  $x$  is an even number just in case  $x$  is a POES and that  $x$  is a POES just in case  $x$  is divisible by two is just a lengthy and cumbersome way to say that  $x$  is an even number just in case  $x$  is divisible by two. The problem is not that the account is incomplete unless we are told what the POESs are, the problem is that the appeal to the notion of a POES does no work in the proposed account of what it is for a number to be even—and similarly so for the epistemological approach.

<sup>viii</sup> As a matter of fact, I doubt that CT provides adequate sufficient conditions for a property’s being explanatorily relevant/causally efficacious. CT faces the same notorious problem encountered by all counterfactual accounts of causation, viz., what David



---

Robb has called ‘fork cases’—cases “in which an effect depends nomologically or counterfactually on a property, but only because that property is itself a mere result, an epiphenomenon, of the properties that do the real causal work” (Robb, 1997, p. 181). Suppose Jack and Jill are having a picnic where they eat poisonous mushrooms and get sick. In that case, CT would erroneously render Jack’s getting sick causally responsible for Jill’s getting sick—had Jack not got sick, Jill would not have gotten sick, and given that Jack got sick, Jill’s getting sick was inevitable. (Baker has argued that this scenario does not pass CT, because “if an experimenter had prevented Jack’s getting sick by not allowing him to eat the mushrooms, and not interfering in any other way, Jill (who ate the mushrooms) would still have gotten sick” [Baker, 1995, p. 124]. Given that Baker accepts the Stalnaker-Lewis truth conditions of counterfactual conditionals [Baker, 1995, p. 155n5], this response is surprising, for the possible world relevant for the evaluation of the counterfactual in question, i.e. the closest possible world in which Jack doesn’t get sick but Jill does, certainly doesn’t involve an interfering experimenter.)

<sup>ix</sup> Thanks to an anonymous referee for raising this issue.

<sup>x</sup> The same applies to Raymont who also analyses explanatory relevance (sive causal efficacy) in terms of “nomological or counterfactual dependency relations” (Raymont, 2003, p. 238).

<sup>xi</sup> In note viii above I have briefly indicated why I don’t think that Baker’s (or any other, for that matter) counterfactual account of mental causation works. In Walter, 2005 I argue that Jackson and Pettit’s program explanation account also fails to offer adequate sufficient conditions of causal efficacy that can be met by mental properties.

<sup>xii</sup> Thanks again to an anonymous referee.

### Acknowledgments

An early predecessor of this paper was written in 2004 while I was a visiting fellow at the department of philosophy at the Ohio State University, thanks to a DAAD grant. Support from both institutions is gratefully acknowledged. A later version was presented at the GAP.6 conference of the German Society for Analytic Philosophy 2006 in Berlin (Germany). I would like to thank the members of the audience at this occasion, in

---

particular Sanne Lodahl, who provided helpful criticism, and two anonymous referees for Erkenntnis for various valuable comments.

### References

- Baker, L. (1993). Metaphysics and mental causation. (In J. Heil & A. Mele (Eds.), Mental causation (pp. 75-95). Oxford: Clarendon Press.)
- Baker, L. (1995). Explaining attitudes. (Cambridge: Cambridge University Press)
- Burge, T. (1993). Mind-body causation and explanatory practice. (In J. Heil & A. Mele (Eds.), Mental causation (pp. 97-120). Oxford: Clarendon Press.)
- Davidson, D. (1970). Mental events. (In L. Foster & J.W. Swanson (Eds.), Experience and theory (pp. 79-101). Amherst, MA: The University of Massachusetts Press and Duckworth.)
- Fodor, J. (1987). Psychosemantics. (Cambridge, MA: MIT Press)
- Jackson, F. & Pettit, P. (1988). Functionalism and broad content. Mind, 97, 381-400.
- Jackson, F. & Pettit, P. (1990a). Program explanation: A general perspective. Analysis, 50, 107-117.
- Jackson, F. & Pettit, P. (1990b). Causation and the philosophy of mind. Philosophy and Phenomenological Research, 50, 195-214.
- Jackson, F. & Pettit, P. (1992). Structural explanation in social theory. (In D. Charles & K. Lennon (Eds.), Reduction, explanation, and realism (pp. 97-131). Oxford: Clarendon Press.)
- Kim, J. (1998). Mind in a physical world: An essay on the mind-body problem and mental causation. (Cambridge, MA: MIT Press)
- Kim, J. (2005). Physicalism, or something near enough. (Princeton, NJ: Princeton University Press)
- Pietroski, P. (1994). Mental causation for dualists. Mind and Language, 9, 336-366.
- Raymont, P. (2003). Kim on overdetermination, exclusion and nonreductive physicalism. (In S. Walter & H.D. Heckmann (Eds.), Physicalism and mental causation: The metaphysics of mind and action (pp. 225-242). Thoverton: Imprint Academic.)

- 
- Robb, D. (1997). The properties of mental causation. Philosophical Quarterly, 47, 178-194.
- Thomasson, A. (1998). A nonreductivist solution to mental causation. Philosophical Studies, 89, 181-191.
- Walter, S. (2005). Program explanations and mental causation. Acta Analytica, 20, 32-47.
- Walter, S. (2006a). Causal exclusion as an argument against non-reductive physicalism. Journal of Consciousness Studies, 13, 67-83.
- Walter, S. (2006b). Mentale Verursachung. Eine Einführung. (Paderborn: mentis)
- Walter, S. (2007). Epiphenomenalism. (In J. Fieser & B. Dowden (Eds.), The Internet Encyclopedia of Philosophy. <http://www.iep.utm.edu/e/epiphen.htm>)