

Uwe Meixner

How to Reconcile Non-Physical Causation with the Physical Conservation Laws

Abstract

It is widely assumed (especially in the philosophy of mind) that the non-physical causation of physical events (by non-physical substances or non-physical mental events) is incompatible with physics. The paper clarifies in what sense there is such an incompatibility. The paper shows that there is no incompatibility of non-physical causation of the physical with the physical laws, that is, with physics properly speaking, by describing a physically consistent and even plausible situation of non-physical causation of the physical. In the very same situation, however, the Strong and also the Weak Principle of Causal Closure of the physical world are violated. But, it is argued, this violation does not show an incompatibility of the non-physical causation of the physical with physics, but only with the metaphysical “Überbau” of physics, physicalism; physics and physicalism should be carefully distinguished, just as physics and metaphysics. Nevertheless, in the appendix of the paper an attractive way of preserving the Weak Principle of Causal Closure while admitting non-physical causation of the physical is also explored.

There are two types of non-physical causation: causation by non-physical events, for example physically non-reducible mental events, and causation by non-physical agents, for example Cartesian substances, *res cogitantes*. *Since* non-physical causation of physical events is widely assumed to be irreconcilable with the physical conservation laws, in particular with the law of the conservation of energy, and *since* the physical conservation laws are very well confirmed items of physics, the only ontological status that remains for non-physical entities, like non-physical events and non-physical substances, is widely regarded to be, at least with respect to the physical world, a causally totally inefficacious one. And this, in the eyes of most philosophers, is a very uncomfortable ontological status indeed, since, according to them, it suggests or even implies non-existence.

I will not here discuss whether physical inefficaciousness is ontologically negative, although I do believe that the opinion that it is is unwarranted. Rather, I will focus on the basic assumption of naturalistically minded philosophers that non-physical causation of the physical is irreconcilable with the physical conservation laws. This assumption is false, or to phrase it more carefully: there is a plausible general view of causation according to which non-physical causation of the physical is reconcilable with the physical conservation laws. That does not yet mean that non-physical causation of the physical does in fact occur, without any detriment to the conservation laws; but I will also attempt to make plausible that it does in fact so occur.

Sometimes (and even mostly) when philosophers speak of non-physical causation of the physical being irreconcilable with *physics*, they actually have in mind something else than an alleged irreconcilability with the physical conservation laws: they mean that non-physical causation of the physical is irreconcilable with the so-called *Principle of Causal Closure of the Physical World*. Now, that Principle, which actually has two different versions, as we shall presently see, has to be distinguished carefully from the physical conservation laws. Indeed, it is logically independent from them in both versions:

PCC1 Everything that is physical and that has a cause also has a physical cause.

In symbols: $\forall x(\Phi x \wedge \exists y C y x \supset \exists z(\Phi z \wedge C z x))$.

PCC2 Everything that is a cause of something that is physical is physical.

In symbols: $\forall y \forall x(C y x \wedge \Phi x \supset \Phi y)$.

These are the two versions of the Principle of Causal Closure (I henceforth omit the tag "of the Physical World").¹ They are not always carefully distinguished (not for example by Peter Bieri and by Ansgar Beckermann, when they discuss the difficulties of mental causation in *Analytische Philosophie des Geistes*, p. 5f, and *Analytische Einführung in die Philosophie des Geistes*, p. 115f); but they should be so distinguished since they are very different. PCC2 is the *strong version* of the Principle of Causal Closure; PCC1 is the *weak version* of the Principle: PCC1 follows logically from PCC2, but

¹ Other formulations of the Principle of Causal Closure have been proposed. Jonathan Lowe, for example, proposes a temporalized version of PCC1 as "the *strong* ... principle of the causal closure of the physical": "At every time at which a physical state has a cause, it has a fully sufficient physical cause"; and he proposes PCC1 without the condition "and that has a cause" (see PPC below) as "the ... *weak* principle of the causal closure of the physical": "Every physical state has a fully sufficient physical cause" (see *An Introduction to the Philosophy of Mind*, p. 27, p. 30). I do not find his terminology quite appropriate here.

not vice versa. PCC1 allows non-physical causal overdetermination of the physical,² PCC2 does not. PCC1 appears to be well-confirmed; PCC2, in contrast, is confirmed only to the extent that PCC1, its logical consequence, is confirmed. Indeed, there are *prima facie* counterinstances to PCC2 (derived from the straightforward – physicalists say “naïve” – construal of mental causation), and that leaves PCC2 more or less a metaphysical postulate: not something one arrives at, but something one presupposes from the very beginning. As such PCC2 is definitely not a part of the empirical science of physics, but at most a *regulative principle* for it in Kant’s sense. More likely, it is *merely* a part of the metaphysical position of *physicalism*. I will further substantiate these claims in what follows.

Non-physical causation of the physical, as I will argue, is reconcilable with the physical conservation laws. It is *in addition* reconcilable with the Weak Principle of Causal Closure if we do not exclude causal overdetermination, that is, the occurrence of two or more *sufficient* causes for one and the same effect. Causal overdetermination, and in particular its psycho-physical variety, may be un-esthetical or un-economical, but that can be no ground for the ontological claim that it does not occur; much less can it be excluded on the mere ground that it would be at best “extremely odd,” as Eugene Mills points out against Jaegwon Kim.³ The denial of causal overdetermination must be argued for differently (that is, it must be *argued* for), and it is very difficult to do this in a non-circular fashion, since there is good *prima facie* evidence for the occurrence of causal overdeterminations. For example, given transitivity of causation, any causal chain of more than one sufficient cause constitutes a case of causal overdetermination. It may be harder to find overdeterminations by *independent* causes – as long as one forgets the one *glaring* *prima facie* example: independent mental and neurophysiological causation of the same effect (a certain behavior, for example). The task is to dislodge the intuition that *such* causal overdeterminations do in fact occur (which cannot be done by assuming dogmatically that they do not, or rather: *must not*, occur).⁴

Note that there is no apparent problem at all, not even the problem of causal overdetermination, in reconciling the Weak Principle of Causal Closure with the non-physical causation of physical events if “cause” is taken to mean

² This is noted by Eugene Mills, since his principle of “physical closure” is no other than PCC1 (see his “Interactionism and Overdetermination”, p. 105f). Mills does not draw into consideration any other version of the Principle of Causal Closure.

³ See Kim, *Supervenience and Mind*, p. 247, and Mills, “Interactionism and Overdetermination,” p. 106, p. 113f).

⁴ *Systematic* causal overdetermination can be used to reconcile prominent psychophysical intuitions in the philosophy of mind that are widely held to be irreconcilable. See for this the Appendix, (2).

sine-qua-non cause, and not *sufficient cause* (what “cause” is here taken to mean).⁵ Deplorably, many philosophers, even many who call themselves “analytic philosophers,” do not bother at all to specify, even roughly, what concept of causation they have in mind when they confidently assert that non-physical causation of the physical is irreconcilable with physics.

Non-physical causation of the physical remains, however, irreconcilable with the Strong Principle of Causal Closure, no matter how causation is explicated. But as long as non-physical causation of the physical is reconcilable with the physical conservation laws that should only be of secondary interest to *physicists*. And as philosophers, we are obliged to honor physics, we are not obliged to honor physicalism. From the point of view of physics, it would even be of secondary interest if the Weak Principle of Causal Closure is violated while the conservation laws are left intact. Physics already had to jettison a lot of quasi-metaphysical baggage in the last one hundred years, among them the time-honored *Physical Principle of Causation*:

PPC Every physical event has a physical cause.

In symbols: $\forall x(\Phi x \wedge Ex \supset \exists y(\Phi y \wedge Cyx))$.

PPC is stronger than PCC1 (in view of the fact that $\exists y Cyx$ – “x has a cause” – is taken to analytically imply Ex – “x is an event”), and physics (or at least the majority of physicists) has already given up PPC: it is ready to face standard physical events that have no physical cause (for example, certain events of radioactive decay, or, for that matter, even the Big Bang).⁶ Given this situation, it would really be no great matter any more for physicists *qua* physicists (and not *qua* metaphysicians) if they gave up PCC1, too, and consequently, to the

⁵ Jonathan Lowe distinguishes (without further explanation, but apparently in the tradition of Mackie’s INUS-conditions) *cause* from *fully sufficient cause* (which is made up of *causes*) and has a conception of causal overdetermination according to which a case of both “mental state M is a cause of P” and “P has a fully sufficient physical cause” being true does not constitute a case of causal overdetermination, even if M is supposed to be a non-physical state. (See *An Introduction to the Philosophy of Mind*, p. 27f, p. 30f). Yet, in the supposed circumstances, since P has a fully sufficient physical cause, M, being non-physical, is clearly no part of that cause and clearly not necessary for causing P, and hence every fully sufficient cause *that involves* M is, quite obviously, causally overdetermining P. And why shouldn’t there be *some* fully sufficient cause of P *involving* M, since M is after all supposed to be a cause of P? So we do have a case of causal overdetermination of P after all, and one for which M is responsible.

⁶ For David Chalmers, PPC expresses the causal closure of the physical world, and surprisingly he thinks PPC is very well confirmed: “The best evidence of contemporary science tells us that the physical world is more or less causally closed: for every physical event, there is a physical sufficient cause.” (*The Conscious Mind*, p. 125.) Jonathan Lowe, too, is ready to defend PPC (his “weak principle of the causal closure of the physical”): see *An Introduction to the Philosophy of Mind*, p. 30. As should be clear by now, taking this stance is not recommendable: it contradicts what most believe to be our best *physics*.

extent they actually believed in it, PCC2. I do, of course, not deny that giving up PCC1 and PCC2 would be a really great matter for *some philosophers*.

Consider now some situation of *physical indeterminism* such as physics finds acceptable today. At a certain point in time the laws of physics – that is, as far as we know, the laws of nature – and the course of physical events up to that point in time are not sufficient for determining how things will physically go on. There are two or more nomologically possible continuations of the physical past and present, two or more nomologically possible immediate physical futures. How will this situation be resolved? I will consider the a priori possible general ways of resolving this situation in turn:

- (1) The physical world simply stops: none of its nomologically possible immediate futures becomes real. – This simply isn't the case: the physical world goes on (in fact, the laws of physics may require that it go on *in some way or other*). Or as Shakespeare has expressed it: "Come what come may / Time and the hour runs through the roughest day." (*Macbeth*, I, 3.)
- (2) The physical world branches: all, or at least several, of its nomologically possible immediate futures become real. – This means that we get a plurality of physical worlds, not a plurality of merely possible worlds, but a plurality of *real* worlds, that is, a plurality of real spacetimes, existing side by side without any possibility of physical exchange beyond the branching point. Whoever is ready to jump into this metaphysical abyss may do so, I shall not. Let me just add that it is one of the more bizarre features of contemporary physicalism that most physicalists consider the idea of pure possibilities, and in particular: of merely possible worlds, an extravagant metaphysical fancy not worth their serious attention, but that some physicalists, if faced with certain difficulties for their position, are resolutely ready to embrace a plurality of *real worlds*.⁷
- (3) The physical world continues in precisely one way: exactly one of its nomologically possible immediate futures becomes real. This, surely, is the most plausible resolution of a situation of physical indeterminism. But merely stating it is not enough. The plausibility of this third resolution disguises a disturbing question: Given that the physical world could have continued in this or another nomologically possible way, *why* did it continue in precisely *this* way, and not the other?

⁷ Such a difficulty is the difficulty of explaining the "fine-tuning" of the universe. See Bernulf Kanitscheider, *Im Innern der Natur*, p. 126f. I do not wish to suggest, however, that the idea of a plurality of real worlds resulting from a situation of physical indeterminism is intrinsically connected to physicalism. It is not.

This is a demand for *explanation*, and we now have to consider in which way an explanation may be forthcoming. But there is one thing that is already clear even now: Howsoever the real immediate physical future after a situation of physical indeterminism is to be explained, none of the possible explanations can be in conflict with the physical conservation laws. Those laws are preserved *after* the moment of indeterminism just as they were preserved up to and including the moment of indeterminism, since the possible immediate physical future which becomes real is a *nomologically possible* immediate future, that is, an immediate physical future which is compatible with the totality of physical laws, and therefore, in particular, compatible with the physical conservation laws. *No physical fact after an indeterministic situation contradicts physical laws*; this fact, call it "the compatibility fact," is implicit in the very description of a situation of physical indeterminism. Now: any explanation of the real immediate physical future after such an indeterministic situation has to respect the compatibility fact. Therefore: any such explanation will be compatible with the physical conservation laws.

And here is the best way, I submit, to make some headway in obtaining an explanation of the immediate real physical future after an indeterministic situation. Given a situation of physical indeterminism after which the physical world, as it always does, goes on in a single manner, let e^* be the totality of (real) physical events immediately after the indeterministic situation. What is the (or a) cause of this event? If we mean, as I think we should, by "cause" *actualized sufficient cause*, the (or a) cause of e^* cannot be a physical event or anything physical (e.g., particles, electromagnetic fields, etc.), since not even the totality of all events that have happened before e^* , and therefore not the totality of their physical constituents (e.g., actual particles, electromagnetic fields, etc.),⁸ could determine the happening of e^* . Therefore: if e^* has a cause at all, it must be a non-physical cause. But according to the *General Principle of Causation*,

GPC Every event has a cause,
in symbols: $\forall x (Ex \supset \exists y Cyx)$,

⁸ In my view, there are no physical objects that are not constituents or parts of physical events, and I do not believe in an independent causality of the, if I may say so, *non-evently* constituents of physical events (these views are corollaries of the metaphysical theory I expound in my book *Ereignis und Substanz*). In other words, the causal influence of non-evently physical objects is completely reducible to the causal influence of the events of which these objects are constituents. Therefore, if the totality of all physical events before e^* is not sufficient for determining e^* , then neither is the totality of all the physical constituents of these events. This leaves us with no physical cause for e^* .

e^* must have a cause. Therefore: e^* has a non-physical cause,⁹ and e^* having a non-physical cause points us towards a certain kind of explanation of the immediate real physical future after an indeterministic situation: such an explanation had best be a *causal and non-physical* one.

The argument that has just been deployed, call it “the C-Argument,” is based on certain assumptions: (1) on the assumption of the occurrence of a situation of physical indeterminism with several nomologically possible continuations into the future, but with a single real continuation; (2) on the assumption of a certain conception of causation; (3) on the assumption of the General Principle of Causation. But these assumptions are certainly *compossible*. Therefore the C-Argument establishes the *compatibility* of non-physical causation of the physical with the physical conservation laws. If the assumptions of the C-Argument are not only compossible, but also *complausible*, then the argument establishes something more: namely, the *plausibility* of non-physical causation of the physical in accordance with the physical conservation laws.

Are the assumptions (1) to (3) plausible? This, in their particular case, is really a matter merely of their separate plausibility: if each of them is plausible, then, in their case, they are also plausible together. What about the plausibility of each of the three assumptions?

I start with the General Principle of Causation, since most philosophers nowadays may find this principle implausible. Logical carelessness is one of the vices of philosophers, and therefore many of them apparently confuse PPC – the *Physical* Principle of Causation – with GPC – the *General* Principle of Causation. There is indeed evidence in physics against PPC, but that evidence leaves GPC unscathed.

But perhaps something more than logical carelessness is, after all, involved here. Given that there is a physical event that has *no physical cause*, in contradiction to PPC, how could one derive from this that there is an event that has *no cause at all*, in contradiction to GPC? Answer: By assuming that if there is no physical cause, then there is no cause at all, or in other words:

PCP1 Everything that has a cause also has a physical cause.

In symbols: $\forall x(\exists yCyx \supset \exists z(\Phi z \wedge Cz x))$.

Or by assuming the logically stronger

⁹ A similar, but more specialized argument can be constructed with respect to the Big Bang: The Big Bang is a physical event that, as science is ready to concede, has no physical cause. But GPC demands its having a cause. Hence the Big Bang has a non-physical cause, and obviously (since there is no time before the Big Bang) without hurting the physical conservation laws.

PCP2 Everything that causes something is physical.

In symbols: $\forall y \forall x (Cyx \supset \Phi y)$.

PCC1 and PCC2 are obvious logical consequences of PCP1 and PCP2, respectively. I call PCP1 the *Weak Principle of Causal Physicalism*, PCP2 the *Strong Principle of Causal Physicalism*. These latter principles bear their metaphysical character on their faces. They can never be confirmed within physics, since physics treats only of the physical. Thus, what could *at most* be confirmed within physics are PCC1 and PCC2, the respective logical weakenings of PCP1 and PCP2. However, we have already seen that even PCC1 and PCC2 – PCC2 more so than PCC1 – are much less principles of physics than principles of *physicalism*, of the metaphysical “Überbau” of physics.

Given metaphysical neutrality, there is no reason to assume PCP2 or PCP1, and therefore the statement stands: that physical evidence against PPC leaves GPC unscathed. But given metaphysical neutrality, there is indeed every reason to assume GPC. This latter principle is *not* a biased metaphysical principle; it is something that every philosopher of any metaphysical persuasion can accept. And being a philosopher, he or she rationally ought to accept it, since GPC asserts part of what is the objective basis for a rational view of the world.¹⁰ In case the reader is not convinced (as is likely),¹¹ let me add that for establishing an instance of the non-physical causation of the physical it is indeed not even necessary to assume GPC, but only necessary to assume that *in some* (not necessarily *in every*) situation of physical indeterminism there is a cause of whatever happens immediately after it. It does not seem plausible to deny this latter assumption, although, of course, one *can* deny it; the denial, however, seems gratuitous, or – alternatively – *ad hoc*.¹²

What about the second assumption of the C-Argument, the particular conception of causation made use of in it? This conception was that of causation as *sufficient causation*. There were no further restrictions on the

¹⁰ As such, GPC is coeval with philosophy itself. Since the time of antiquity it was often asserted in the form of the principle “Nothing comes of nothing.”

¹¹ The reason for such skepticism may well be that GPC apparently enables a cosmological argument for the existence of God. Cf. footnote 9. But aside from the fact that GPC *does not* lead straight from the Big Bang to the existence of God, one should consider that a principle may indeed be effectively refuted by a *reductio ad absurdum*, but certainly not by a *reductio ad deum*.

¹² One can also plead *ignorance* with respect to the assumption. This is legitimate (even though somewhat unsatisfactory philosophically) as long as it does not turn out that the only reason for taking this stance is the wish to escape the conclusion that there is non-physical causation of the physical.

concept of causation, except that causes have to be already *actualized* (this in order to exclude backward causation by *future* physical causes) and (in parentheses) that what is caused is always an event; in particular, and appropriately so, there was no legislation on the question of what categories of entities can be causes.

Sufficient causation with actualized causes and with events as effects is as respectable a conception of causation as sine-qua-non or probabilistic causation or any other conception of causation, and, more importantly, it seems to be essentially the conception that is meant when non-physical causation of the physical is held to be irreconcilable with the physical conservation laws.¹³ Now, it would be a blatant *petitio principii* if a proponent of that position considered causation to be intrinsically connected to the transmission of energy or impulse in the physical sense, or even to *be* such transmission.¹⁴ Non-physical causation *in that sense* of the physical, that is, causation of something physical by something non-physical that nevertheless inexorably involves transmission of, say, energy in the physical sense, is of course irreconcilable with the physical conservation laws, no doubt about that (at least no doubt I care to go into here). But this *physicalistic* conception of causation is also at issue in the whole question, and indeed the C-Argument *also* leads to the conclusion, on minimal assumptions, that the physicalistic conception of causation as involving transmission of a conserved physical quantity is simply not adequate.

But are those assumptions all that minimal? How about the first assumption of the C-Argument, the occurrence of a situation of physical indeterminism with several nomologically possible continuations into the future, but with a single real continuation? I have already argued for the *singleness* of real continuation after a situation of physical indeterminism. This leaves us with the question of the occurrence of situations of physical indeterminism. But it is received opinion among physicists that such situations do occur, and philosophers had better not deny this.

In fact, physicalist philosophers usually do not deny the occurrence of situations of physical indeterminism, they usually do not want to revert to 19th-century determinism. For this would all too clearly reveal the metaphys-

¹³ If causes are merely supposed to be probabilistic or sine-qua-non causes, why then, after all, might there not be probabilistic or sine-qua-non *non-physical* causes of a certain physical effect in addition to the probabilistic or sine-qua-non *physical* causes of it (considering, in particular, that causal overdetermination cannot be a problem then)?

¹⁴ Views identical or close to this have been held, notably, by W. V. O. Quine (see *The Roots of Reference*, p. 7), D. Fair, P. Dowe, and W. Salmon. The idea has recently been brought forward again by M. Kistler. (See References for bibliographical data.)

ical nature of their position. Rather, physicalists¹⁵ usually deny the *relevance* of indeterministic situations, which they regard as being confined to the microphysical world, for the question of *free human agency*. The denial is not reasonable, in particular if at the same time, as is usually done, the basic laws of physics are supposed to explain not only microphysical phenomena but also macrophysical ones. This clearly implies that what happens on the microscopic scale is after all held to be *relevant* for what happens on the macroscopic scale. And why should only (nomologically) explainable (that is, determined) microphysical events be *relevant* for macrophysics, and not also *unexplainable*? Why, in other words, should microscopic indeterminism never magnify itself into macroscopic indeterminism?

The question of free human agency brings us to the final considerations in this paper: What does the C-Argument, which has now been shown to rest on plausible assumptions, mean for non-physical agent and mental causation of the physical, in particular if we focus our attention on the human sphere?

The C-Argument has the conclusion that a certain physical event e^* has a non-physical (sufficient) cause, without there being any violation of the physical conservation laws. This cannot mean that that cause is *entirely non-physical*, it can only mean that that cause is *not entirely physical*, if we consider physical events that went before e^* to be not only relevant circumstances for the causation of e^* , but also parts of its non-physical cause. If, however, all the physical events that went before e^* are merely relevant circumstances for the causation of e^* , then indeed the non-physical cause of e^* could be *entirely non-physical*. Moreover, since e^* , however short, is a complete section of the course of the physical world, e^* is a big event, and therefore, *prima facie*, it is more likely than not that its non-physical cause is, in some sense, big too. And therefore, if for example the cause is to be a non-physical agent, then that agent is more likely to be a large group of non-physical substances, acting together, than a single non-physical substance, acting alone. However, if the difference between e^* and the course of events that has gone immediately before is small, then the reality of e^* , if due to the action of a non-physical agent, could after all be due to a *single* non-physical substance.

Yet, all of this is speculation, the offering of possibilities on conditions. The C-argument is silent about the composition and nature of the non-physical cause of e^* . It leaves many open questions, and this makes it difficult to apply to the human sphere. But suppose we hold that human beings are in a certain sense non-physical substances (*entirely non-physical* or *not entirely physical*?)

¹⁵ And also, at least tentatively, some non-physicalists. Cf. Jonathan Lowe, *An Introduction to the Philosophy of Mind*, p. 30.

that each have one, in a certain sense, non-physical mental life, as so many people have believed in the course of the last 2500 years.¹⁶ This view is not shown to be correct by the C-argument (far from it), but it is certainly *corroborated* by that argument, since the C-argument shows that a certain intuition which is intimately connected to the mentioned view is logically, scientifically and philosophically *coherent*. This is the intuition that the course of human history, with all its glories and terrible crimes, *is not*, not even as far as its purely physical side is concerned, purely due to physical causes, and *is not*, where physical causes give out, purely due to absolute chance, but is more often than not *causally due* to non-physical human substances making decisions on the basis of their non-physical mental lives.

Appendix

(1) General principles of causation and their logical relations

Weak Principle of Causal Closure (of the Physical World)

PCC1 Everything that is physical and that has a cause also has a physical cause.

In symbols: $\forall x(\Phi x \wedge \exists y Cyx \supset \exists z(\Phi z \wedge Czx))$.

Strong Principle of Causal Closure

PCC2 Everything that is a cause of something that is physical is physical.

In symbols: $\forall y \forall x (Cyx \wedge \Phi x \supset \Phi y)$.

Physical Principle of Causation

PPC Every physical event has a physical cause.

In symbols: $\forall x(\Phi x \wedge Ex \supset \exists y(\Phi y \wedge Cyx))$.

General Principle of Causation

GPC Every event has a cause.

In symbols: $\forall x(Ex \supset \exists y Cyx)$.

Weak Principle of Causal Physicalism

PCP1 Everything that has a cause also has a physical cause.

In symbols: $\forall x(\exists y Cyx \supset \exists z(\Phi z \wedge Czx))$.

Strong Principle of Causal Physicalism

PCP2 Everything that causes something is physical.

In symbols: $\forall y \forall x (Cyx \supset \Phi y)$.

¹⁶ And as I have argued in *Ereignis und Substanz*.

GPC does not logically imply, nor is logically implied by, any of the other five principles of causation listed above: it is quite on its own. But for the other five we have:

$$\begin{array}{ccc} \text{PCP2} & \rightarrow & \text{PCC2} \\ \downarrow & & \downarrow \\ \text{PCP1} & \rightarrow & \text{PCC1} \leftarrow \text{PPC} \end{array}$$

- (2) A reconciliation of psychophysical intuitions on the basis of systematic causal overdetermination

The following¹⁷ is for those readers who are not *a priori* afraid of causal overdetermination (it may be helpful for the timid to give the thing a *neutral* name: *multiple causal determination*). (Before beginning, I hasten to say that, in what follows, all quantifiers will be restricted to the entities that are left in the world if we subtract from it all disembodied spirits that may perhaps be found in it. Without that restriction, some of the principles posited might appear to be *too general*.)

Definition 1

Let F and G be properties, x an individual, t a moment of time:

F is at t in x a causal representative of $G := x$ has F at t , and x has G at t , and $\forall p$ (that x has F at t causes p iff that x has G at t causes p) and $\forall p$ (p causes that x has F at t iff p causes that x has G at t).¹⁸

The above defined time- and individual-dependent relation between properties – the relation of *causal representation* for properties – is symmetric and transitive, and, moreover, reflexive with respect to all properties F , individuals x and times t such that x has F at t .

Principle 1a

For all Ψ -properties G , all individuals x , all times t : if x has G at t , then there is a Φ -property F such that F is at t in x a causal representative of G (or equivalently: ... such that G is at t in x a causal representative of F).

Principle 1a states that any mental (or psychical: Ψ -)property, whenever it is instantiated (no matter in which individual or at what time), is, in that instance, causally represented by a physical (or Φ -)property. *Principle 1a* is

¹⁷ It is close in sentiment to ideas Eugene Mills defends in "Interactionism and Overdetermination." The details are rather different.

¹⁸ In what follows, " p ", " q " and " r " are variables for states of affairs.

already to a very high degree empirically confirmed, and cognitive science and neurophysiology are working strenuously to confirm it even better.

Now, *Principle 1b*, following below, is closely related to *Principle 1a* (and like the latter it is very well confirmed). The only major difference is that *Principle 1b* concerns states of affairs instead of properties. But before stating it, we need another definition (which is closely related to *Definition 1*):

Definition 2

Let p and q be states of affairs:

p is a causal representative of $q := p$ obtains, and q obtains, and $\forall r(p \text{ causes } r \text{ iff } q \text{ causes } r)$ and $\forall r(r \text{ causes } p \text{ iff } r \text{ causes } q)$.

Definition 2 gives us a symmetric and transitive relation of causal representation between state of affairs; it is, moreover, reflexive with respect to all obtaining states of affairs. And we are all set to posit

Principle 1b

For every obtaining Ψ -state of affairs q : there is a Φ -state of affairs p such that p is a causal representative of q .

Note that if a Ψ -state of affairs q and a Φ -state of affairs p are causal representatives of each other, then they are indeed closely related to each other. *But* that relation cannot be equivalent to causation. We cannot but conclude that it is neither the case that q causes p , nor the case that p causes q , for otherwise either p or q would cause itself (according to *Definition 2*, since p and q are causal representatives of each other), and that is absurd.

Further we have

Principle 2a

No Ψ -property is a Φ -property.

Principle 2a is the expression of a deeply entrenched dualistic intuition (the one that is denied by the so-called type-identity theory, which nowadays even hard-core physicalists find somewhat hard to believe). *Principle 2a*, just like *Principle 1a*, has a counterpart for states of affairs, which is, however, from the epistemological point of view not entirely its equal (because states of affairs are, in a way, *tokens*, and many people find so-called token-identity theories so much more plausible than type-identity theories):

Principle 2b

No Ψ -state of affairs is a Φ -state of affairs.

Nevertheless, strong common sense intuitions also support *Principle 2b*, common sense being dualistic in sentiment. Continuing, we also have:

Principle 3

There are Ψ -states of affairs p and Φ -states of affairs q such that p causes q .

Principle 3 affirms the existence of the causation of physical states of affairs by mental (psychical) states of affairs.¹⁹ *Principle 3*, too, is the expression of a deeply entrenched intuition, one that interactionist dualists and (non-eliminativist) physicalists (hence the majority of philosophers) share, an intuition that is apparently confirmed by experience at every turn. Finally we posit

Principle 4

Every Φ -state of affairs that is caused at all is caused by a Φ -state of affairs.

Principle 4 simply is the Weak Principle of Causal Closure *intended for states of affairs as effects and causes*. (PCC1, on the other hand, taken as we left it, is the Weak Principle of Causal Closure *intended for events as effects*. Merely given that some physical events have causes, PCC1 is *compatible* with assuming that some *or even all* of these causes are not events. But merely given that some physical states of affairs have causes, *Principle 4* is *compatible only* with assuming that *some* of these causes are not states of affairs.)

What is remarkable about the above six principles is that there is *no* logical contradiction in their conjunction, that each of them is plausible in itself, and that, taken together, they give a rather a satisfactory picture of the relationship between mental and physical entities, one that reconciles the demands of science and of common sense: Common sense assumes, in accordance with *Principle 3*, that some mental state of affairs q – for example, that Jim feels ashamed at t_0 – causes a physical state of affairs p – for example, that blood rushes to Jim's cheeks at t_1 . Science demands, following *Principle 4*, that p also have a physical cause (let us *here* neither question that science demands this, nor question that the demand is in every case in fact fulfilled, even though we have seen above that physics might dissociate itself even from the Weak Principle of Causal Closure, and even though we have argued above for the occurrence of violations of even the Weak Principle of Causal Closure). But *Principle 1b* already provides for the required physical cause. According to it, there must be a physical state of affairs r which is a causal representative of q – which state of affairs r , therefore, causes p , since q causes p .²⁰ Hence science, as far as the relation of the mental to the physical is concerned, may

¹⁹ Note that if F is a Ψ -property (alternatively: Φ -property), then *that x has F at t* is here taken to be a Ψ -state of affairs (alternatively: Φ -state of affairs). This allows it to easily generate *prima facie* examples for the causation of physical states of affairs by mental ones.

²⁰ In fact, *Principle 1b* alone suffices to prove the following corollary of *Principle 4*: Every Φ -state of affairs *that is caused by a Ψ -state of affairs* is also caused by a Φ -state of affairs.

as well follow *Principle 1b* instead of *Principle 4*. Guided by *Principle 1b*, it becomes the task of science to *specify* (actually *find*) some such physical state of affairs r whose existence as a causal representative of q is already predicted by *Principle 1b* (and actual scientific practice shows that this task is taken up, and that it is worthwhile to take it up, since the prospects of completing it are promising). Here *Principle 1a* is helpful: Given (as assumed) that q is the state of affairs that Jim feels ashamed at t_0 and that this state of affairs obtains, Jim has at t_0 the mental property of *feeling ashamed*, and therefore, according to *Principle 1a*, there must be a physical property F such that F is at t_0 in Jim a causal representative of that property of feeling ashamed. Science merely needs to specify F ,²¹ for then it has automatically given a full specification of a physical state of affairs r that is a causal representative of the mental state of affairs q : because, clearly, r can be taken to be the state of affairs that Jim has F at t_0 .

There may indeed be a *unique* way of specifying F . For the following additional principle seems to be highly plausible (it has never been refuted):

Principle 5

For all Φ -properties F and F' , Ψ -properties G , individuals x and moments of time t : if F is at t in x a causal representative of G , and F' is at t in x also a causal representative of G , then $F = F'$.

Principle 5 makes it possible to speak of the physical property which is a causal representative of G in x at t , if x has a Ψ -property G at t . We can abbreviate the definite description just used by " $\phi_{G,x,t}$ ", and thus can formulate the following psychophysical law (as a corollary of *Principles 1a* and 5):

For all Ψ -properties G , all individuals x , all times t : if x has G at t , then x has $\phi_{G,x,t}$ at t .

In other words, there is no instance of a mental property which is not also an instance of the physical causal representative, relative to the instance, of that property. In this (new) sense mental properties can be said to *depend* on physical ones.

In view of the rather satisfactory epistemological situation created by the seven principles now stated – a situation which serves common sense and science alike, *preserving* common sense and *furthering* the progress of science – why should one want to give up any one of the principles stated

²¹ It is worth remarking that if Jim is again ashamed at t_2 , one of the physical properties which are at t_2 in Jim causal representatives of the property of feeling ashamed *may, but need not* be one of the physical properties that are at t_0 in Jim causal representatives of the property of feeling ashamed.

above, especially in view of the fact that none of them is contradicted or disconfirmed by experience?²² The only "reason" seems to be a *mere dislike* of causal overdetermination, or a *mere dislike* of dualism. But mere dislikes are no reasons.

References

- Beckermann, A.: *Analytische Einführung in die Philosophie des Geistes*, de Gruyter, Berlin 1999.
- Bieri, P.: *Analytische Philosophie des Geistes*, Beltz/Athenäum, Weinheim³ 1997.
- Chalmers, D.: *The Conscious Mind. In Search of a Fundamental Theory*, Oxford University Press, New York/Oxford 1997.
- Dowe, P.: Wesley Salmon's Process Theory of Causality and the Conserved Quantity Theory, *Philosophy of Science* 59 (1992), 195–216.
- Fair, D.: Causation and the Flow of Energy, *Erkenntnis* 14 (1979), 219–250.
- Kanitscheider, B.: *Im Innern der Natur. Philosophie und moderne Physik*, Wissenschaftliche Buchgesellschaft, Darmstadt 1996.
- Kim, J.: *Supervenience and Mind: Selected Philosophical Essays*, Cambridge University Press, Cambridge 1993.
- Kistler, M.: Reducing Causality to Transmission, *Erkenntnis* 48 (1998), 1–24.
- Lowe, J.: *An Introduction to the Philosophy of Mind*, Cambridge University Press, Cambridge 2000.
- Meixner, U.: *Ereignis und Substanz*, Schöningh, Paderborn 1997.
- Mills, E.: Interactionism and Overdetermination, *American Philosophical Quarterly* 33 (1996), 105–117.
- Quine, W. V. O.: *The Roots of Reference*, Open Court, La Salle, Illinois 1973.
- Salmon, W.: Causality Without Counterfactuals, *Philosophy of Science* 61 (1994), 297–312.

²² *Principle 4* and *PCCI* is a subtle case. There is empirical evidence that some physical events / (obtaining) states of affairs have no physical cause. But this, by itself, does not disconfirm these two principles: we need empirical evidence that some physical events / states of affairs *that have a cause* have no physical cause. No such empirical evidence is forthcoming. One reason for this is that there is no evidence for causation by mental states of affairs / events *without* a physical representation of it, and lots of evidence for such causation *with* a physical representation; precisely this fact of the matter makes the (relevant version of) *Principle 1b* empirically well-confirmed. We frequently *do* have the feeling that *we* (and not our body or parts of it, or current conditions of them, nor our current mental life or parts of it or facts about it) cause a physical event / state of affairs that is otherwise entirely undetermined; but this feeling is surely not empirical evidence in the strict sense, its importance for our self-understanding as *freely acting* human beings notwithstanding. Given that some physical events / (obtaining) states of affairs have no physical cause, it may, indeed, be *plausible* that some physical events / states of affairs *that have a cause* have no physical cause. But this plausibility is best seen as supported by the (relevant version of the) General Principle of Causation, which for all its metaphysical neutrality is a *metaphysical* principle nonetheless.