

# Mental Causation and the New Compatibilism

Jens Harbecke

Philosophy of Science  
Witten/Herdecke University, Germany  
jens.harbecke@uni-wh.de  
www.jensharbecke.com

## Abstract

Twenty years ago Stephen Yablo developed his original theory of mental causation, which has drawn much attention ever since. By providing a detailed reconstruction of Yablo's approach, this paper first demonstrates that a certain line of critique that has repeatedly been brought forward against Yablo over the last two decades misconstrues the core idea of the model. At the same time, the reconstruction reveals that Yablo's approach is probably the first explicit version of the "new compatibilism" within the philosophy of mind. This fairly young family of theories essentially claims a non-identity as well as a non-distinctness of mental and physical phenomena. The second part of the paper then demonstrates that the new compatibilist approaches in general, and Yablo's theory in particular, even though they can resist much of the popular criticism, fall prey to a new theoretical trilemma once the nature of the respective *analysantia* is made explicit. Finally, a model of the psychophysical relation is developed that may allow the new compatibilists to escape the mentioned trilemma.

## 1 Introduction

Twenty years ago Stephen Yablo (1992b; 1992a; 1997; 2000; 2001) developed his original theory of mental causation, which has drawn much attention ever since.<sup>1</sup> Yablo's most cited article on the topic (1992b) first develops a distinction between determinable and determinate properties as it is known from traditional classifications of colors and geometrical forms. For instance, "crimson" and "bordeaux" are sometimes described as determinates of the determinable "red"; "being rectangular" and "being square" are considered determinates of the determinable "being quadrangular".

In a second step, Yablo suggests that "(...) mental properties stand to their physical realizations in the relation that quadrangularity bears to squareness, or that colors bear to their shades (...)." (1992b, 256) In Yablo's view, this insight holds the key for a convincing solution to the problem of mental causation. Determinable and determinate properties are in no causal competition with each other. This is true even in situations, in which instances of both properties are already causally sufficient for a corresponding effect. Consequently, the widely discussed "causal exclusion principle" (cf. premises (4), (4\*) and (4\*\*) in sec. 2) cannot be applied to cases where the two candidate causes are related by determination.

According to a popular interpretation of Yablo, the following argument nicely captures the line of thought just sketched and is therefore at the heart of the author's approach:

---

<sup>1</sup>A check on GoogleScholar in March 2013 mentions 440 quotations of Yablo (1992b).

- (I) All properties that are related by determination do not compete for causal influence.
- (II) All mental properties are related by determination to their subvenient physical properties.
- ∴ (III) No mental property competes with any of its subvenient physical properties for causal influence.

Whilst the validity of this argument is not in dispute<sup>2</sup>, the two premises have been debated in various ways over the last twenty years. Some authors have challenged the claim that determinates and determinables do not compete for causal influence (cf. McGrath 1998, 171; Gillett and Rives 2005, 496; Walter 2007, 238-240), which in turn has motivated others to defend this claim in more detail (cf. Shoemaker 2001, 432; Wilson 2009, 166-168). Again other authors have denied that the relation between mental and physical properties is that of determination (cf. Ehring 1996, 471-476; Worley 1997, 283/284, McGrath 1998, 170; Funkhouser 2006, 562/563; Walter 2007, 221; Haug 2009, 7-8), to which again others have responded with a defense of Yablo (Wilson 2009, 162-165).

One of the central aims of the present paper is to show that a good portion of the debate represented by these authors actually misconstrues the core idea of Yablo's model. In contrast to the popular opinion, it is not the case that Yablo's theory is based on an argument of the form (I)-(III). Rather, the notion of classical determination plays a merely heuristic role for Yablo. The real work in the theory is done by the essentialist analysis of events developed in Yablo 1992a (407-425, 430-436), 1992b (261-265), and most detailed in his 1987. Consequently, the mentioned criticisms of premises (I) and (II) are of little relevance for Yablo's approach. After two decades of debate on the theory, this point requires emphasis.

The second central aim of this paper is to support the claim that Yablo's theory is a variation of what can be described as a "new compatibilism" within the philosophy of mind, i. e. an approach that assumes both a *non-identity* and a *non-distinctness* of mental and physical phenomena. Depending on various possible *analysanda* presupposed, this approach has received several different formulations by different authors.

The third central aim is to show that the new compatibilism in general, and Yablo's theory in particular, faces a new fundamental trilemma. All existing versions of the new compatibilism either collapse into a strong dualism, a strong monism, or an unsatisfactory neutralism with respect to their *analysantia* if they want to maintain the contention about the non-identity and non-distinctness of the *analysanda*. Since, in light of this trilemma, the new compatibilism is an unsatisfactory approach at this point, a forth hypothesis to be defended investigates a potential way out for the new compatibilists.

The structure of the paper is the following. Sec. 2 briefly sketches the problem of mental causation that Yablo and the new compatibilists aim to solve. Sec. 3 isolates the reasons why Yablo's argument has often been interpreted as equivalent to (I)-(III). Moreover, some popular objections against premises (I) and (II) are reconstructed. Secs. 4 and 5 focus on the actual dialectics of Yablo and show that the generally valid objections against (I) and (II) do not affect the core of Yablo's approach.

Section 6 relates Yablo's version to other variations of the new compatibilism. The mentioned trilemma is highlighted in 7, before a potential solution is explicated in section 8.

<sup>2</sup>A rigorous formalization would be:  $\forall\phi\forall\psi(\mathcal{D}\phi\psi \rightarrow \neg\mathcal{C}\phi\psi)$ ,  $\forall\phi\forall\psi(\mathcal{M}\phi\wedge\mathcal{P}\psi\wedge\mathcal{S}\phi\psi \rightarrow \mathcal{D}\psi\phi)$  ∴  $\forall\phi\forall\psi(\mathcal{M}\phi\wedge\mathcal{P}\psi\wedge\mathcal{S}\phi\psi \rightarrow \neg\mathcal{C}\phi\psi)$ , where  $\mathcal{D}(\_, \_) = \dots$ determines... ,  $\mathcal{C}(\_, \_) = \dots$ causally competes with... ,  $\mathcal{S}(\_, \_) = \dots$ supervenes on... ,  $\mathcal{M}(\_) = \dots$ is a mental property,  $\mathcal{P}(\_) = \dots$ is a physical property.

## 2 The problem of mental causation

This section briefly reviews the so-called “problem of mental causation”. The problem is usually considered as consisting of four jointly incompatible assumptions. Two general formulations of these assumptions are the following, both of which appear frequently in the literature. The first formulation can be termed the “token formulation”.

- (1\*) Some mental events have physical effects.
- (2\*) No mental event is identical to any physical event.
- (3\*) All physical events have a sufficient physical cause.
- (4\*) Physical events are not systematically causally overdetermined.

The second version can be described as the “type-” or “property formulation”.

- (1\*\*) Some mental properties are causally sufficient for some physical properties.
- (2\*\*) No mental property is identical to any physical property.
- (3\*\*) For every instance  $x$  of a physical property  $P_1$ , there is an instance  $y$  of a further physical property  $P_2$ , such that  $y$ 's being an instance of  $P_2$  is causally sufficient for  $x$ 's being an instance of  $P_1$ .
- (4\*\*) Instances of physical properties are not systematically causally overdetermined.

Each assumption of the formulations seems to be defensible in isolation. In particular, mental events and properties play an indispensable role when we interpret actions of agents. Assuming that mental events never cause anything would amount to the contention that virtually all our interpretive practices are pointless. The inacceptability of this consequence suggests (1\*)/(1\*\*). On the other hand, as empirical science seems to tell us, any physical event has a complete sufficient physical cause, suggesting the truth of (3\*)/(3\*\*).<sup>3</sup>

The conjunction of assumptions (1\*)/(1\*\*) and (3\*)/(3\*\*) implies that mental events must (i) either be redundant causes of physical events or (ii) they must themselves be physical events. The latter option has turned out difficult to defend at least with respect to the type formulation. To be sure, a quasi consensus exists in philosophy of mind that the mind depends on the physical. One way to express such a hypothesis is by the notion of supervenience:

- (S) Necessarily, for all  $x$  and all mental properties  $M$ , if  $x$  instantiates  $M$ , then there is also a physical property  $P$  such that  $x$  instantiates  $P$  and, necessarily, all individuals instantiating  $P$  also instantiate  $M$ .

The problem is that a supervenience relation between sets of properties is not sufficient for an identity of mental to physical properties. In fact, it has turned out difficult to identify a single paradigmatic mental property that is coextensional with some (natural and non-disjunctive) physical property (cf. Putnam 1967). This insight can be expressed by the following hypothesis:

- (M) For all mental properties  $M$  and all physical properties  $P$ , if  $P$  is sufficient for  $M$ , there is always at least one  $x$  that is an  $M$  but not a  $P$ .

Since coextensionality is necessary for identity, (M) implies (2\*\*) (and the falsity of what we described as “option (ii)”; i. e. the claim that mental events are identical to physical events). Moreover, if a fine-grained model of events is presupposed (cf. Kim 1973), (M) also implies (2\*). The reason is that, according to this model, for two events to be identical it is necessary for the properties of the events be identical.

<sup>3</sup>It should be noted, however, that the doctrine of the causal completeness of the physical is less obvious than its acquiescent acceptance in the philosophical literature suggests (cf. Montero 2003; Primas 2007). Bishop and Atmanspacher (2010) have argued that the notion of causation is inconsistent with the fundamental laws of physics insofar as these laws have no direction of time, hence no past and future, hence no cause and effect.

This leaves only the first of the two options (i) and (ii): Mental events, i. e. instances of mental properties, can only be redundant causes of physical events, i. e. instances of physical properties. This claim, however, contradicts assumption (4\*)/(4\*\*), so that the full set of assumptions (1\*)/(1\*\*)–(4\*)/(4\*\*) is inconsistent. The reason for holding on to (4\*)/(4\*\*) is that redundant causes *per definitionem* do not make any difference to what goes on in the world, and it seems extravagant to assume the existence of things that make no causal difference whatsoever.

From the inconsistency it follows that any feasible theory of mental causation has to reject at least one of the assumptions, where this rejection will have to be defended as demanding the lowest overall theoretical costs. In the 1970s, psychophysical functionalism was widely considered a dominant strategy in this respect (cf. Putnam 1967, 1975; Fodor 1974). Functionalism accepted that mental events overdetermine effects with physical causes. After a while, however, and mainly through the work of Jaegwon Kim (1979, 1985) it became clear that functionalism brings much higher costs than initially expected. Kim’s work led many authors to embrace an identity theory and, thereby, to reject premise (2\*)/(2\*\*). In the late 1990’s however, several authors have shown that Kim’s version of an identity theory is not conservative and ultimately eliminates the mental (cf. for instance Bontly 2002 and Harbecke 2008). This consequence has struck many philosophers as unacceptable as well.

Yablo and authors who below will be described as the “new compatibilists” (sec. 6) have sought to evade this general dilemma between functionalism and the identity theory. The next section focuses on Yablo’s proposal for an analysis that resists both option i) and option ii).

### 3 Classical Determination

As mentioned in sec. 1, a popular line of thought interprets Yablo’s approach as essentially based on argument (I)-(III). The reason is that the notion of “determination” plays a central role in Yablo’s presentations of his theory.

In particular, Yablo develops a thought experiment in his 1992b (257) and 1992a (423) that helps to formulate a *reductio* proof against premise (4\*\*) and, thereby, for premise (I). In this thought experiment, a pigeon “Sophie” is conditioned to peck on red triangles. However, the triangles presented to Sophie are not only red, they are scarlet. If the causal exclusion principle applies generally, it may seem that the redness of the triangle is a different and competing cause of Sophie’s pecking than the triangle’s scarletness. But, of course, that’s absurd. Yablo concludes that “(...) the [exclusion] principle does not apply to determinates and their determinables – for we know that they are not causal rivals.” (1992b, 259)<sup>4</sup> Hence, Yablo accepts premise (I).

Premise (II) Yablo considers established on the basis of the widely accepted thesis of psychophysical supervenience (S) and the thesis of a multiple realization of mental properties (M) (cf. sec. 2 and Yablo 1992b, 254/255). Yablo defines determination as follows:

(D) A property *P* determines a property *Q* just in case “for a thing to be *P* is for it to be *Q*, not *simpliciter*, but in a specific way.” (1992b, 252)

In Yablo’s understanding, (D) is basically implied by the conjunction of (S) and (M). To the rhetorical question whether mental properties are in fact determinables of certain physical properties just as colors are determinables of their shades, Yablo answers: “Yes, at least that is my conjecture.” (1992b, 256) Hence, it seems that Yablo also subscribes to premise (II).

<sup>4</sup>Already MacDonald and MacDonald 1986 had made this observation. However, in contrast to Yablo, the MacDonalds presuppose the identity of instances of determinables and determinates.

As mentioned above, various authors have challenged premises (I) and (II). For instance, Gillett and Rives (2005, 491) and Walter (2007, 239) have argued that Yablo's example of the pigeon Sophie shows something important about potentially competing explanations, but little to nothing about ontology. Metaphysically speaking, the exclusion principle applies without exception: Determinable properties are screened-off from their effects by their determinate properties. Hence, if read metaphysically, premise (I) is false (cf. also McGrath 1998, 171).

Premise (II) has been attacked by arguments of the following form (cf. Ehring 1996, 471-476; Worley 1997, 283/284, McGrath 1998, 170; Funkhouser 2006, 562/563; Walter 2007, 221; Haug 2009, 7-8):

- (IV) All determination relations satisfy the higher-order properties in set  $\mathcal{D}$ .
- (V) The relation that mental properties bear to their subvenient physical properties does not satisfy the higher-order properties in set  $\mathcal{D}$ .
- $\therefore$  (VI) *Not*: The relation that mental properties bear to their subvenient physical properties is that of determination.

Ever since William Ernest Johnson (1921, 174) introduced the distinction of determinates and determinables into philosophy, several authors have aimed to delineate precisely set  $\mathcal{D}$ .<sup>5</sup> From the results of these investigations, it must indeed be concluded that the relation between mental properties and their physical supervenience bases is not that of determination.

As an example, consider the fact that all determinates falling under the same determinable are incompatible with another: no object can be scarlet and bordeaux at the same place at the same time. This incompatibility principle does not seem to apply to the supervenience bases of mental properties. It seems possible in principle that a person can instantiate two neural supervenience bases of the same mental state at the same time. With arguments of this kind, premises (IV) and (V) seem to be confirmed, and therefore the truth of (VI). However, (VI) is just the negation of (II), suggesting that the argument (I)-(III) is not sound.

## 4 Metaphysical Determination

Although the arguments against premises (I) and (II) presented in sec. 3 have probative force, they turn out to be of little importance for Yablo's overall theory. The reason is that in his version of argument (I)-(III) Yablo refers to a different notion of determination than the one presupposed by the argument (IV)-(VI). Moreover, Yablo's theory is not primarily concerned with the causal connection between mental and physical properties, but with mental and physical events.

Yablo's definition of determination (D) seems to match the criterion of "asymmetric sufficiency" that had been identified by Johnson (cf. also Ehring 1996, 470). However, whilst for classical determination this sufficiency describes a relation between concepts, Yablo intends it as a metaphysical relationship as the following example shows (1992b, 253). Assume that for the property of having a temperature of 95° Celsius, there exists a maximally specific micro-physical property  $K$  such that everything that has  $K$  also has a temperature of 95° Celsius, but not vice versa. According to Yablo, the relation between these two properties is a prototypical determination relation. However, no conceptual analysis will reveal that  $K$  is sufficient for having a temperature of 95° Celsius.

"[And] since it is only the metaphysics that matters to causation, we should discount the traditional doctrine's conceptual component and reconceive determination in wholly metaphysical terms." (1992b, 253)

<sup>5</sup>For reviews of the results of these investigations, cf. Harbecke 2008, 183-191; Funkhouser 2006, and Sanford 2008.

This shows that Yablo considers the metaphysical determination relation with which he is primarily concerned as distinct from Johnson's classical determination relation. Consequently, Yablo can accept premises (IV) and (V), which refer to classical determination (see also Yablo 1997, fn. 22). At the same time, he can accept the truth of premise (II), since conclusion (VI) does in fact contradict (II). In short, argument (IV)-(VI) poses no threat to Yablo's position.

Another argument promoted by Gillett and Rives (2005, 491) and Walter (2007, 239) attacks premise (I) (cf. 3). The authors claim that the exclusion principle (4\*\*) applies in the metaphysical sense to properties connected by determination, even if two explanations referring to a determinate and its determinable property may well be accepted out of explanatory purposes at the same time.

One reason<sup>6</sup> why Yablo's analysis is not affected by this argument either is that the author is clearly not concerned with a second-order relation between properties but a first-order relation between events (cf. Yablo 1992a, 407-425, 430-436, Yablo 1992b, 260-280, Yablo 2001, 66/67). That is, Yablo is concerned not with problem (1\*\*)-(4\*\*), but with problem (1\*)-(4\*). As a consequence, whether the exclusion principle (4\*\*) applies to properties related by determination is not a question that Yablo's theory has to concern itself with. At least more argument would be required to show that Yablo has to confront this question.

This suggests that the two kinds of criticism often brought forward against Yablo misunderstand the core aim and content of the author's theory. As a consequence, the original theory is left widely unharmed by the critical arguments, even if these are of great importance in themselves. At the same time, this insight of the immunity of Yablo's theory against the popular objections does not yet answer questions about the positive content of this theory. In particular, it is not clear at this point why Yablo can accept the truth of premise (4\*\*) and still reject premise (4\*). It has also not been clarified what role argument (I)-(III) ultimately plays for Yablo's theory. To provide answers to these questions is the aim of the following section.

## 5 Cumulative Subsumption

According to Yablo, the key for solving problem (1\*)-(4\*) is the insight that mental and physical events stand in a first-order determination relation. The author's definition of this relation is analogous to (D):

(d) “[An event]  $p$  determines [an event]  $q$  iff: for  $p$  to occur (in a possible world) is for  $q$  to occur (there), not simpliciter, but in a certain way.” (1992b, 260)

The central challenge for Yablo now consists in showing that (i) this definition is interpretable in a metaphysical way, (ii) that events related by (d)-determination do not compete for causal influence, and (iii) that mental and (the corresponding) physical events are in fact related in this way. Secs. 5.1, 5.2, and 5.3 discuss these challenges consecutively. For what follows, Yablo's notion of an “individual essence” is central, which he analyzes in terms a “cumulative subsumption of categorical properties”.<sup>7</sup> Both notions are presented by Yablo in his 1992b (sec. 5), and most extensively in his 1987 and 1992a (sec. 2).

<sup>6</sup>A second reason for rejecting arguments similar to those of Gillett & Rives and Walter has been developed by Bennett (2003). Bennett shows that there is a disanalogy between prototypical cases of overdetermination and cases in which one of the causing properties is sufficient for the other.

<sup>7</sup>In 1997 Yablo uses the notion of an “inclusion of influence” and in his 2000 and 2001 that of “intensive parthood”; however, it is clear that the author intends these synonymously as the originally developed notion of “subsumption/refinement/strengthening” explicated in 1992b, 1987, and 1992a.

### 5.1 Definition

According to the ordinary understanding, the essence of a thing is the set of properties without which the thing cannot exist. Not all essential properties are contained in a thing's essence, however. Identity properties such as "being identical to  $x$ " or type-properties such as "being of the same type as  $x$ " are determined by the essence of a thing, but they do not themselves contribute to the way a thing is.

Essences also provide a set of conditions that must be satisfied to be a certain thing. For many things  $x$  and  $y$ , more conditions must be satisfied to be  $x$  than to be  $y$ . For instance, the *Rosetta Stone* should be distinguished from the *Rosetta Granite*, which has just the same properties as the *Rosetta Stone* but lacks the properties of having been discovered by Pierre-Francois Bouchard and of having provided the key for the deciphering of hieroglyphic writing. Similarly for events, the blocking of the valve should be distinguished from the sudden blocking of the valve, not the least because the two may have different causal effects.

In Yablo's view, the inclusion relation between essences of events helps to make sense of (d). If  $q$ 's essence is included in  $p$ 's essence, then  $p$  "subsumes"  $q$  ( $p \geq q$ ). Event  $p$  "determines" event  $q$  ( $p > q$ ) if the inclusion is proper (1992b, 262).<sup>8</sup> Since only categorical properties can be contained in essences, something must be said about what distinguishes categorical from hypothetical properties. Yablo makes the following first attempt to launch the distinction:

"[A] property is *categorical* if its possession by a thing  $x$  at a possible world is strictly a matter of  $x$ 's condition in that world, without regard to how it would or could have been; other properties, for example counterfactual and modal properties, are *hypothetical*." (Yablo 1992b, 261/262)

Despite its intuitive appeal, the definition turns out to be circular. The problem is that it ultimately describes a property as categorical if it is possessed by thing  $x$ , without regard to how  $x$  would or could have been in other possible worlds *in categorical respects*. This makes the circularity obvious. To solve this problem, Yablo uses the inclusion relation itself to distinguish categorical from hypothetical properties.

(K) "[A property]  $C$  is categorical only if: necessarily, for all  $p$  and  $q$  such that  $p \geq q$ ,  $p$  has  $C$  iff  $q$  does." (Yablo 1992b, 263)

This definition excludes all identity properties from the set of categorical properties. Moreover, with the set of cumulative properties it is now possible to explain the content of (d). If for  $p$  to occur (in a possible world) is for  $q$  to occur (there), not simpliciter, but in a certain way, then the essence of  $p$  subsumes that of  $q$ .

Applied to events, this idea implies that, if  $p$  determines  $q$ , then  $q$  occurs in more worlds than  $p$ . Suppose, for example, that  $p$  and  $q$  are coincident events that instantiate property set  $\{F, G, H, I\}$  in the actual world. Suppose further that  $\mathcal{P} := \{F, G, H\}$  is the essence of  $p$  and that  $\mathcal{Q} := \{F, G\}$  is the essence of  $q$ . Event  $q$  now can occur in some possible worlds in which  $p$  does not occur, because its essence is poorer, whilst the opposite is false.

### 5.2 Causal competition

With the example of two events  $p$  and  $q$  developed in the last section, it becomes clear why Yablo rejects the claim of a causal competition between events related by determination. If  $p$  and  $q$  have essences  $\mathcal{P}$  and  $\mathcal{Q}$ , then it is impossible for  $p$  to screen off  $q$  and vice versa, since this would amount to  $p$  ( $q$ ) screening off itself.

<sup>8</sup>What Yablo 1992b calls "subsumption" what is described as "refinement" in his 1987 and as "strengthening" in his 1992a. The intended meaning is the same, however. Cf. fn. 7 also.

Despite the fact that the coincident events  $p$  and  $q$  do not screen each other off, they are not identical, since  $q$  occurs in some possible worlds in which  $p$  does not occur. With these points in place, Yablo can show how two non-identical events that both have influence on the same physical effect, need not be in causal competition with each other in the sense of a “causal overdetermination” (cf. premise (4\*)).

### 5.3 Mental Determination

The solution for problem (1\*)-(4\*) now seems straightforward. The author assumes that mental events are typically determined by physical events (note that the two claims (s) and (m) are analogous to (S) and (M)):

(s) “Whenever a mental event  $m$  occurs, there occurs also a subsuming physical event  $p$ , that is, a physical event whose essence includes  $m$ ’s essence.” (Yablo 1992b, 268)

(m) “For every mental event  $m$ , and every physical event  $p$  which subsumes  $m$ ,  $p$  subsumes  $m$  properly and so determines it.” (1992b, 270)

If mental and physical events are in fact related in this way, then they are not identical. But as sec. 5.2 showed, they can also not compete for causal influence. With this conclusion, problem (1\*)-(4\*) is solved.<sup>9</sup>

The general picture that emerges from this hypothesis is that mental events necessarily coincide with physical events, i. e. each mental event always instantiates the same categorical properties as some physical event. In this sense, mental events are *not distinct* from their subvenient physical events. However, since mental events are to be distinguished from physical events in hypothetical respects, they are also *not identical* to them. Mental events have “impoverished” essences relative to physical events.

### 5.4 Evaluation

The main claims of Yablo’s theory are now clearer and it is more obvious why it is not threatened by premise (4\*). This result complements the claims made in sec. 4, according to which Yablo can accept the conclusiveness of argument (IV)-(VI) and he is not affected by the truth of premise (4\*\*).

However, if determination in terms of cumulative subsumption is the relation that does the main work for Yablo’s theory, it may seem confusing that Yablo spends much time discussing examples of classical determination and an argument similar to (I)-(III) especially in 1992b (252-260). The reasons for this dialectical step become clearer once the whole of Yablos theory as represented by 1987; 1992a; 1992b; 1997; 2000, and 2001 is taken into account. Yablo’s reply to the question whether “(...) mental properties stand to their physical realizations in the relation that rectangularity bears to squareness, or that colors bear to their shades?” (1992b (256) is intended *heuristically* and not in the sense of a factual equivalence. The goal of introducing the “Sophie” thought experiment is merely to show that there is at least one second-order relation whose *relata* do not compete causally with respect to further events. This is then taken as a reason to search for a first-order relation with the same features.<sup>10</sup>

<sup>9</sup>A question that cannot be answered here is whether Yablo actually rejects a premise of the argument or whether he assumes an ambiguity in the premises. It may seem at first that the author rejects premise (4\*) thereby claiming an overdetermination of physical events. The definition of causal proportionality (cf. 1992b, 273-279) suggests, however, that Yablo accepts an “overdetermination” of causal influence but rejects an overdetermination in terms of genuine causality. For a discussion of this distinction, cf. Harbecke 2008, 300-309.

<sup>10</sup>Already in the introduction to his 1992b, Yablo is very cautious in asking: “What if mental phenomena are determinables of physical phenomena in *something like* the traditional sense (...)?” (250, emphasis added) This makes it

## 6 A new compatibilism

By describing mental and physical events as *non-identical* and *non-distinct*, Yablo initiated a theoretical tradition that can be described as a “new compatibilism” in the philosophy of mind.<sup>11</sup> It is a “compatibilism”, because this theory claims a compatibility of mental causation with a psycho-physical non-identity. It is “new” because it should be carefully distinguished from older compatibilist approaches such as functionalism à la Fodor and Putnam, a pre-established harmony à la Leibniz, or a Cartesian dualism.

What unites several theoretical approaches under this term is the fundamental notion of an inclusion, or containment, relation between certain entities. For the notion of an inclusion, recall some simple definitions of set theory. Two sets  $M$  and  $N$  are identical iff:  $M \subseteq N$  and  $N \subseteq M$ . Two sets  $M$  and  $N$  are distinct iff:  $M \cap N = \emptyset$ . If  $M \subset N$ , then  $M$  and  $N$  are neither identical nor distinct. The new compatibilists use this simple idea in one way or other to describe the relation between mental and physical phenomena.

In Yablo’s theory, this idea occurs in the definition of a proper subsumption of *event* essences. Interestingly, Kim in his 1993 claimed that events such as Sebastian’s stroll and his leisurely stroll are “different, if not entirely distinct, events. Not entirely distinct since the latter *includes* the former.” (45)

Several further authors have argued analogously for the claim that mental *properties* are neither identical nor distinct from physical properties. According to this idea, the set of causal powers constitutive of a mental property is always contained in the set of causal powers constitutive of some physical property (cf. Worley 1997; Wilson 1999, 2009; Shoemaker 2001; Clapp 2001; Ehring 2003).<sup>12</sup> According to the causal theory of properties, every natural property is associated with a set of conditional powers (cf. Shoemaker 1980, 212/13). A property  $P$  equips objects, that instantiate  $P$  with a conditional power  $K(U, E)$  if instances of  $P$  in circumstances  $U$  bring about instances of  $E$ , whilst realizations of  $U$  in isolation do not bring about instances of  $E$ . Powers whose circumstantial condition  $U$  is empty are called “powers simpliciter”.

In some cases, the conditional powers that are bestowed by a property are a subset of a further property that are bestowed by a second property. According to authors such as Shoemaker (2001), Clapp (2001), Ehring (2003) and Wilson (1999, 2009), the relation between mental and subvenient physical properties should be understood in this way: the powers associated with any mental property are a subset of powers associated with some physical property. Since the two properties are not identical, but also not distinct, this theory should be viewed as versions of the new compatibilism.

Another approach claims that mental phenomena are “wholly constituted” by physical phenomena, without the constitution relation implying identity (Pereboom and Kornblith 1991; Pereboom 2002; Gillett 2003). More concretely, it proposes that “(...) mental causal powers are wholly constituted of physical causal powers; they are neither identical to (nor are they necessary and sufficient for) them, nor wholly independent of them. That’s why

---

clear that the example of classical determinables and determinates plays merely a heuristic role. In his 1992a, the author even develops the theory of cumulative subsumption first, and then discusses the notion of determination.

<sup>11</sup>To my knowledge, the term “new compatibilism” is not in use so far in the philosophy of mind. I adapt the term “compatibilism” from Terence Horgan, who uses the notion of “causal compatibilism” to describe the hypothesis that “(...) mental causation via nonphysical properties can co-exist with physical causation even if the physical realm is causally closed (...).” (Horgan 1997, 166) See also Karen Bennett, who uses the same term to describe the hypothesis that “(...) no effect can have more than one sufficient cause unless it is overdetermined.” (Bennett 2003, 473) The way “compatibilism” is used in these contexts is to be distinguished from its use in the debate on free will.

<sup>12</sup>The original idea of an inclusion of sets of causal powers definitional of properties is already found in Fales 1982 and 1990, 239-243. However, Fales did not intend this idea as a theory of mental causation.

they don't compete." (1991, 143) This "constitutionalist" version of the new compatibilism presupposes mental causal *powers* as the *analysandum*.<sup>13</sup> Mental powers are taken to be wholly constituted by physical powers, which reflects the inclusion idea and which seems to rule out a distinctness of the two. Nevertheless, the former are not identical to the latter, as they can be multiply constituted (cf. *op. cit.* 138). It follows that mental powers do not compete with causal powers and both have a place in the causal nexus (cf. *op. cit.* 143).

A more recent approach developed by Dardis (2008) is based on the claim that instantiations of mental properties as "form properties" constitute mental events always in conjunction with physico-material properties. According to Dardis, every mental event necessarily instantiates a complex material property  $P_M$ . This is the property of being of the kind of matter that underlies the mental event. This material property, i. e. typically the property of being made of neurons, is not sufficient for the corresponding mental state, however. Additionally, a form property is required that determines a certain structure of the matter (cf. *op. cit.* 158-161). The form property is a mental property  $M_F$ . "Mental properties dovetail with matter properties custom-made to supplement their work." (173). Only the synthesis of  $P_M$  and  $M_F$  makes a corresponding mental event.

Since Dardis accepts a psychophysical supervenience hypothesis (cf. *op. cit.*, 135/136), any mental event  $m$  has to be coinstantiated with a physical event  $p$ . Such a physical event presumably instantiates also a matter property  $P_M$ , but additionally a physical form property  $P_F$ . Since  $m$  and  $p$  are different with respect to  $M_F$ , resp.  $P_F$ , they cannot be identical. However, since they share the instantiation of  $P_M$  they are not distinct and they cannot compete causally in a strong sense. This again reflects the inclusion idea and it shows that Dardis' approach is a version of the new compatibilism.

## 7 A new trilemma

The new compatibilism can escape the general dialectical dilemma that was diagnosed in sec. 2 for the debate on mental causation. It can embrace premise (2\*) or (2\*\*) respectively. At the same time, the new compatibilists typically disarm exclusion principles such as (4\*) and (4\*\*) by showing that the relevant mental and physical phenomena are not distinct. If mental and physical phenomena are not distinct, then they cannot causally overdetermine physical effects in the strong sense of (4\*) and (4\*\*). Rather, they overdetermine physical effects in an unproblematic way. Consequently, (1\*)-(4\*) and (1\*\*)-(4\*\*) can all be true and the new compatibilists are able to avoid the classical criticisms against both reductive and non-reductive theories of mental causation (cf. sec. 2).<sup>14</sup> This makes the new compatibilism a serious candidate for a theory of mental causation.

Nevertheless, the following arguments show that the new compatibilism faces a new kind of trilemma which applies to the different versions of the theory in different ways. For instance, recall that in Yablo's proposal, the essence  $\mathcal{M}$  of any mental event  $m$  is a subset of an essence  $\mathcal{P}$  of some physical event  $p$ . Both essences  $\mathcal{M}$  and  $\mathcal{P}$  contain only categorical properties. The question is whether these categorical properties can be characterized more specifically.

<sup>13</sup>In the original text of Pereboom and Kornblith, it is not always clear what the *analysandum* is supposed to be. In another paragraph, the authors discuss mental states (1991, 134): "[W]e hold that token mental states are physically constituted, but not identical to, the token physical states which constitute them." However, this does not seem to change anything about the fundamental idea of the constitutionalist theory.

<sup>14</sup>Just as for Yablo, it is not immediately clear whether the new compatibilists in general reject premise (4\*) or (4\*\*), or whether they claim that the two formulations of the problem of mental causation contain an ambiguity. Both interpretations are possible.

If all the categorical properties contained in  $\mathcal{M}$  are mental properties (if there be such), then the essence  $\mathcal{P}$  of the physical event  $p$  contains as a subset at least some genuinely mental properties. This, however, questions the physical character of  $p$ , for how can  $p$  be physical if it is partly essentially mental? Secondly, this reading of Yablo's theory would seem to postulate a strong property dualism that claims an independent existence of mental and physical properties. With such a strong property dualism it may no longer be clear that Yablo can actually accept the completeness premise (3\*). This would only be possible if he also accepted that the instantiation of property set  $\mathcal{P} \setminus \mathcal{M}$  is already sufficient to bring about the physical effect of  $p$ . In this case, the subset of  $\mathcal{P}$  that is equivalent to  $\mathcal{M}$  would be redundant for  $p$ 's causal role with respect to its physical effects. This redundancy once again calls into question whether those aspects of  $p$  that are mental have anything to do with the causal effects of  $p$ .

On the other hand, if the categorical properties in  $\mathcal{M}$  are physical properties exclusively, then it is unclear why  $m$  should be considered a mental event at all and not just another, and less comprehensive, physical event  $p^*$ . If all of the properties contained in the essence of  $m$  are physical, there is nothing particularly mental about  $m$  any more. It is clear that, in this way, a causal exclusion of mental events by physical events is avoided. However, at the same time, mental causation is explained away, rather than explained. The nature of the cause is completely physical, and mental causation has given way to physical causation.

Finally, if the properties contained in both  $p$  and  $m$  should be considered neither mental nor physical, but perhaps "neutral", a similar problem occurs. Now the mental event  $m$  is completely made up of neutral properties, and so is the physical event  $p$ . As a consequence, the only thing that remains "mental" about  $m$  is its denomination. In the metaphysical sense, mental causation, and indeed physical causation, has disappeared. Some events cause other events, but characterizing these events as "mental" or "physical" has no ground in how things are in themselves. The distinction becomes merely pragmatic. Again, it seems as though in Yablo's theory mental causation is ultimately explained away.

Note that this conclusion does not depend on the distinction of categorical vs. hypothetical properties. Rather, it is induced by the notion of the mental or physical *nature* of an event. The nature of an event is directly dependent on the (categorical) properties it instantiates and, hence, on its essence. It is due to this that mental causation disappears once only physical and/or neutral properties are allowed into its essence.<sup>15</sup>

The trilemma can be developed analogously for the inclusion theory of mental and physical properties. This theory explained a supervenience relation between mental and physical properties, and the multiple realization of mental properties by physical properties, through the fact that the conditional powers associated with a mental property  $M$  constitute a subset of several sets of conditional powers associated with physical properties  $P_1, P_2, \dots$ .

However, if all conditional powers associated with the mental property  $M$  are mental powers, then it is unclear why these can be contained in sets of powers  $P_1, P_2, \dots$ . Either the powers associated with  $M$  are causally redundant in  $P_1, P_2, \dots$  or premise (3\*\*) is false. If,

<sup>15</sup>Yablo seems to anticipate this trilemma already when he says: "Someone might of course ask, why any physical [event]  $p$  should have the mentally consequential *kind* of physical property, but this is easily explained. Consider the bearing of supervenience on *mental* events: for each of  $m$ 's mental properties, supervenience assigns it a necessitating physical property. But it is hard to think what  $m$ 's physical properties could be if not those of some physical event  $p$  which subserved it." (1992b, 267/267) In this quote, Yablo suddenly introduces a supervenience relation between *properties* to explain the subsumption/inclusion between essences of mental and physical events (note the difference to the supervenience formulation (S) in sec. 2). The problem is that now a critic may demand an explanation as to why this supervenience relation between properties holds. Does it hold in virtue of another inclusion relation, perhaps? Yablo does not answer this question.

however, all conditional powers associated with  $M$  are physical powers, then there seems to be no reason to describe  $M$  as a mental property. Finally, if the powers associated with  $M$  and  $P_1, P_2, \dots$  are neutral, i. e. neither mental or physical powers, then it also becomes unintelligible why they should be contrasted with another in the first place.

For the constitutionalist approach, which had declared mental powers to be wholly constituted by physical powers without identifying the former with the latter, the trilemma takes the following form. Either a mental power  $M$  is actually constituted *completely* by physical powers  $P_1, P_2, \dots, P_n$ . Then  $M$  just is the conjunction of powers  $P_1 \wedge P_2 \wedge \dots \wedge P_n$ . However, this conjunction just seems to pick out a complex physical power, so that it remains quite unclear why it should also be described as mental. Or a mental power  $M$  is *not* completely constituted by physical powers  $P_1, P_2, \dots, P_n$  because there is something more to it than these powers can provide. The result is a dualist ontology once again, since that which is “more to  $M$ ” is clearly non-physical. Moreover, unless a corresponding premise analogous to (3) is to be false, that which is “more to  $M$ ” is likely to turn out redundant with respect to the effects of physical powers  $P_1, P_2, \dots, P_n$ . Finally, if both mental and physical powers are neutral, the question is why they should be distinguished in the first place.

Dardis’ (2008) approach is affected by the trilemma in the following way. He describes events as essentially constituted by a matter property and a form property. The form property of mental events is always a mental property, so that a mental event  $m$  is always constituted by an instantiation of a matter property  $P_M$  and a mental form property  $M_F$ . Supervenience implies that  $m$  is always accompanied by a physical event that in addition to  $P_M$  instantiates a physical form property  $P_F$  as well. Since  $m$  and  $p$  differ with respect to  $M_F$ , resp.  $P_F$ , they are not identical. But since they share the instantiation of  $P_M$  they are also not distinct and they cannot compete causally, at least in the strong sense. Nevertheless, Dardis can accept premise (3\*) only if he claims that the instantiation of  $P_M$  and  $P_F$  is already sufficient for physical effects of  $p$ . But this again leaves the impression that  $M_F$  does not do much causal work, and hence, that  $m$  is not a cause. Once again, mental causation would be explained away. This problem would, of course, vanish if  $M_F$  could be considered a physical form property or if perhaps both  $M_F$  and  $P_F$  would be somehow considered neutral. However, the consequences would be just as disastrous for the project of explaining how mental causation works.

With these implications of the general trilemma, the project of the new compatibilism is doubtlessly under threat. It may be granted, that the presented theories all make a promising move to avoid the problem of mental causation on one level by describing mental and physical phenomena as non-identical and non-distinct. However, any deeper analysis of the postulate reveals a new trilemma. The new compatibilism will have to offer a way out of this trilemma if it wants to circumvent defeat. The next section develops a potential solution to this particular challenge of the new compatibilism.

## 8 A potential way out

The central challenge that Yablo and the new compatibilists face is to provide a credible analysis of the relation between mental and physical events (which they deem to be somewhere in between identity and distinctness) that can avoid the trilemma explicated in sec. 7. This section aims to characterize an idea that may help the new compatibilists in this respect. The idea makes reference to the heuristic notion of a “pattern” that has a certain correspondence to a concept presented by Daniel Dennett (1991).

From an information-theoretic point of view, a pattern can be defined as follows. A series (of dots, numbers etc.) is random if, and only if, the information required to describe the series

accurately (by picking out each element along with its position in the series) is incompressible. In this case, only the verbatim bit map will preserve the series. A series contains a pattern if, and only if, there is a more efficient, compressible description of the series (cf. Chaitin 1975, 48).

Dennett offers a first step to a metaphysical interpretation of a pattern. In his view, a pattern does not only exist relative to an actual description. The only condition required for the existence of a pattern within a series, structure, or configuration of objects is that this series, structure, or configuration is amenable to an information-compressing description (cf. 1991, 34). As long as this condition is fulfilled, the relevant pattern is “real”.

Dennett’s proposal must be credited general intuitive force. For instance, a description of figure 1 that mentions question marks of a certain shape and size is much more efficient than a detailed list mentioning each heart and ring along with their size and relative positions. Moreover, it seems true that the two question marks are “there” and “real” in some sense.

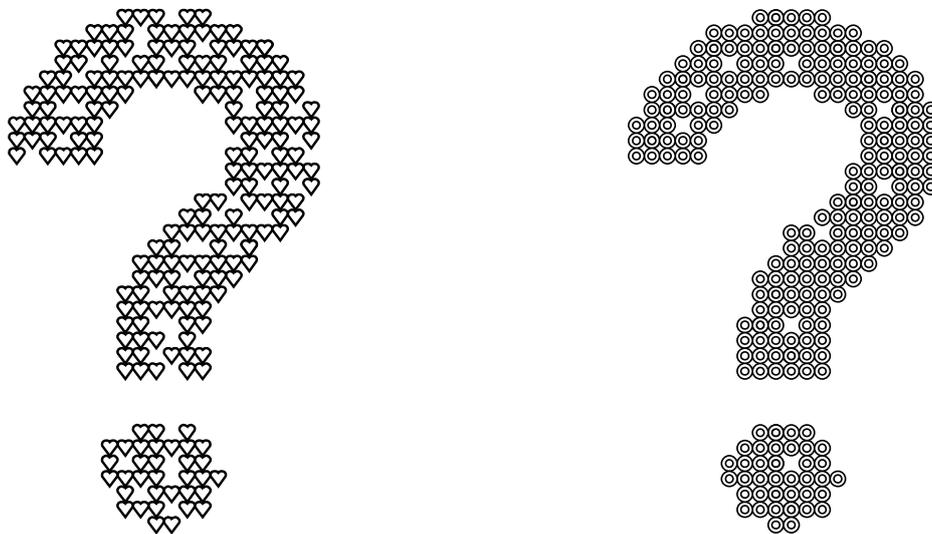


Figure 1: Question marks as real patterns.

One problematic aspect of Dennett’s model, however, is the fact that he remains vague as to what “real” and “being there” means in this context. It remains unclear whether patterns are metaphysically real in a substantial sense or only epistemically real, i. e. in the sense that they are inevitably considered as real by a suitable observer. As a consequence, the precise nature of the relation between patterns and the elements constituting them stays opaque. Dennett may consider this actually a virtue of his theory, as it is only intended to provide a framework for more detailed accounts of mental phenomena. However, since being robustly real is a prerequisite for having causal effects in the world, the vagueness is problematic for a satisfactory theory of mental causation. Or in other words, patterns can play a role for a theory of mental causation only if they can be shown to be real in a metaphysically robust sense.

How could such a demonstration of the robustness of patterns be achieved? One strategy is to bring out more explicitly the unfiltered intuitions directly provoked by colloquial examples such as the one of figure 1. From an unprejudiced perspective, it seems quite odd to say that the question marks of figure 1 exist only in situations in which someone observes or describes them. They are clearly not illusionary as are ghosts walking down the staircase or mirages in the desert. Ghosts and mirages disappear when their observer disappears, but black question

marks printed on a white sheet of paper do not. They continue to be there even if in an instance all minds should cease to exist.

Moreover, already two-dimensional patterns such as question marks have various kinds of typical effects. For instance, they emit light waves of a specific wave length and with a particular structure, they cause the production of characteristic shapes if laid on a xerox machine (where the produced shapes do not necessarily involve hearts and rings), they may also sometimes leave certain kinds of imprints on fingertips and other highly sensitive objects.

Observations of this kind are also induced by three-dimensional patterns. The shape of a cube and the shape of a pop can are both patterns in a sense reminiscent of the question marks of figure 1. Widely independently from the material they are made of, they have various typical causal effects such as characteristic air currents if situated in a wind tunnel, certain imprints they leave in soft sand etc. In short, many two- and three-dimensional patterns are associated with typical causal effects that are widely independent of the stuff constituting them. Intuitions such as these are always disputable, of course. However, from a coherentist perspective, they may offer at least *prima facie* reasons to believe that patterns are real in a metaphysically robust sense.

A further evident feature of patterns is the fact that they necessarily require something of which they are made, and that many different, but not all, kinds of matter can serve to constitute, or realise, a given pattern with all its typical causal effects. For instance, the question mark pattern determines that it can be made up of small printed hearts and rings in many different configurations, but not of oxygen molecules (at least not at room temperature). Only if constituted by particular kinds of stuff, the question marks can perform the causal effects associated with them. Note that the hearts and rings, i. e. the stuff that makes up the question marks in the above illustration, are again patterns themselves. In short, patterns seem to have a dual character by combining a causal (“what kinds of effect does the pattern have?”) with a constitutive aspect (“what kinds of material can make up the pattern?”). Both aspects are obviously interdependent, but they cannot be reduced to one another.

If these conclusions, and the intuitions on which they are based, are accepted, a particular picture of the nature of patterns emerges. Actual patterns are not really distinct from their realizing elements as they could never be realized without the elements. At the same time, the multiple realizability of the causal profile of patterns suggests that a pattern is not identical to the various configurations realizing it. In particular, the causal behaviour of the question mark on the right of figure 1 would have been the same in virtually all relevant circumstances even if one or two rings would have been replaced by hearts.

The relationship between patterns and the matter constituting them is probably characterized best as a “sufficient approximation” of the constituting elements and their configuration to the boundary conditions definitional to the patterns in question. For two-dimensional patterns, this sufficient approximation is determined by spatio-temporal and material aspects. For instance, it is fairly easy to draw lines defining the outer boundaries of the question marks of figure 1 and, with a suitable procedure, to calculate the degree of the approximation of various configurations of hearts to these lines. A similar procedure for defining the outer boundaries of cubes and pop cans is conceivable.

The upshot of these observations and of the intuitions underpinning them is that, if mental and neural events are like real patterns, the trilemma described in sec. 7 for the relationship between mental and neural events loses its bite. But are mental and neural events pattern-like? Mental events surely involve a causal dimension in the sense that a set of typical inputs and outputs is associated with the mental properties they instantiate. Moreover, they involve a constitutional aspect: They require physical material to be instantiated. And according to a

widespread assumption, many different kinds of matter, but not all, are in principle able to realise a mental event (cf. principle (M) in sec. 2). In these respects, there is a certain analogy between two-dimensional and three-dimensional patterns and mental events/patterns.

Consequently, just as the question marks are not identical to, but also not entirely distinct, from the hearts and rings constituting them, mental events would seem to be not identical with, but also not distinct from, the neural events that realize them. However, this non-identity and non-distinctness is not grounded in the kinds of inclusion relations that lead to the trilemma described in sec. 7. Rather, it imitates what with respect to two- and three-dimensional patterns was described as a “sufficient approximation”. Notwithstanding, it excludes a competition between mental patterns and the neural stuff constituting them. In this sense, interpreting mental and neural events as patterns may help to develop a more satisfactory version of the new compatibilist’s approaches to mental causation, including Yablo’s theory.

The central challenge of this analysis of mental events as pattern instantiations concerns the notion of a “sufficient approximation” that has been described as the criterion for the instantiation of a pattern by the underlying matter. If applied to mental events, the notion would have to involve a complex set of spatial, temporal and material dimensions. Spelling out these dimensions in detail is beyond the scope of this paper. The notions of robust mental patterns and the relation of sufficient approximation at this point only constitute a framework for a more detailed account.

## 9 Conclusion

This paper first developed a reconstruction of Stephen Yablo’s widely discussed theory of mental causation. On the basis of this reconstruction, it was possible to show that many of the popular criticisms brought forward against Yablo misconstrue the core ideas of the approach. Moreover, it was argued that Yablo’s theory should be considered a variation of the “new compatibilism” in the philosophy of mind – a theory which describes mental and their underlying physical phenomena as non-identical but non-distinct. However, it was also shown that Yablo’s theory falls prey to a new kind of trilemma that huffs virtually all of the approaches within the new compatibilism. In a final step it was suggested that analyzing mental properties as patterns, and mental events as instantiations of mental patterns, would allow the new compatibilists to avoid the trilemma and remain faithful to their most central claims. To provide a plausible specification of the notion of a “sufficient approximation” and the indication of its dimensions turned out to be the central challenge to this strategy of the new compatibilism. Its precise analysis was relegated to future research within the new compatibilist framework.

## References

- Bennett, K. (2003). Why the exclusion problem seems intractable, and how, just maybe, to tract it. *Noûs* 37(3), 471–497.
- Bishop, R. and H. Atmanspacher (2010). (manuscript).
- Bontly, T. (2002). The supervenience argument generalizes. *Philosophical Studies* 109(1), 75–96.
- Chaitin, G. (1975). Randomness and mathematical proof. *Scientific American* 232(5), 47–52.
- Clapp, L. (2001). Disjunctive properties: Multiple realizations. *The Journal of Philosophy* 98(3), 111–136.
- Dardis, A. (2008). *Mental causation: the mind-body problem*. Columbia University Press.
- Dennett, D. (1991). Real patterns. *The Journal of Philosophy* 88(1), 27–51.

- Ehring, D. (1996). Mental causation, determinables and property instances. *Noûs* 30(4), 461–480.
- Ehring, D. (2003). Part-whole physicalism and mental causation. *Synthese* 136(3), 359–388.
- Fales, E. (1982). Generic universals. *Australasian Journal of Philosophy Kensington* 60(1), 29–39.
- Fales, E. (1990). *Causation and universals*. London: Routledge.
- Fodor, J. (1974). Special sciences (or: the disunity of science as a working hypothesis). *Synthese* 28(2), 97–115.
- Funkhouser, E. (2006). The determinable-determinate relation. *Noûs* 40(3), 548.
- Gillett, C. (2003). The metaphysics of realization, multiple realizability, and the special sciences. *The Journal of Philosophy* 100(11), 591–603.
- Gillett, C. and B. Rives (2005). The non-existence of determinables: or, a world of absolute determinates as default hypothesis. *Noûs* 39(3), 483–504.
- Harbecke, J. (2008). *Mental Causation. Investigating the Mind's Powers in a Natural World*. Frankfurt a. M.: Ontos.
- Haug, M. (2009). Realization, determination, and mechanisms. *Philosophical Studies*, 1–18.
- Horgan, T. (1997). Kim on mental causation and causal exclusion. *Noûs*, 31, Supplement: *Philosophical Perspectives* 11, *Mind, Causation, and World*, 165–184.
- Johnson, W. (1940/1921). *Logic-Part I*. Cambridge: Cambridge University Press.
- Kim, J. (1973). Causation, nomic subsumption, and the concept of event. *The Journal of Philosophy* 70(8), 217–236.
- Kim, J. (1979). Causality, identity, and supervenience in the mind-body problem. *Midwest Studies in Philosophy* 4(1), 31–49.
- Kim, J. (1985). Supervenience, determination, and reduction. *The Journal of Philosophy* 82(11), 616–618.
- Kim, J. (1993). Events as property exemplifications. In J. Kim (Ed.), *Supervenience and Mind*, pp. 33–53. Cambridge: Cambridge University Press.
- MacDonald, C. and G. MacDonald (1986). Mental causes and explanation of action. *The Philosophical Quarterly* 36(143), 145–158.
- McGrath, M. (1998). Proportionality and Mental Causation: A Fit? *Noûs* 32, 167–176.
- Montero, B. (2003). Varieties of causal closure. In S. Walter and H. Heckmann (Eds.), *Physicalism and mental causation: the metaphysics of mind and action*, pp. 173–187. Exeter: Imprint Academic.
- Pereboom, D. (2002). Robust nonreductive materialism. *Journal of Philosophy* 99(10), 499–531.
- Pereboom, D. and H. Kornblith (1991). The metaphysics of irreducibility. *Philosophical Studies* 63(2), 125–145.
- Primas, H. (2007). Non-boolean descriptions for mind-matter problems. *Mind and Matter* 5(1), 7–44.
- Putnam, H. (1975). Philosophy and our mental life. In H. Putnam (Ed.), *Mind, Language and Reality: Philosophical Papers, vol. 2*, pp. 291–303. Cambridge: Cambridge University Press.
- Putnam, H. (1980/1967). The nature of mental states. In N. Block (Ed.), *Readings in philosophy of psychology. Vol. 1*, pp. 223–231. Cambridge: Harvard University Press.
- Sanford, D. H. (2008). Determinates vs. determinables. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2008 ed.).

- Shoemaker, S. (2001). Realization and mental causation. In C. Gillett and B. Loewer (Eds.), *Physicalism and its Discontents*, pp. 74–98. Cambridge/New York: Cambridge University Press.
- Shoemaker, S. (2003/1980). Causality and properties. In *Identity, Cause, and Mind. Philosophical essays*, pp. 206–233. Cambridge: Cambridge University Press.
- Walter, S. (2007). Determinables, determinates, and causal relevance. *Canadian Journal of Philosophy* 37(2), 217.
- Wilson, J. (1999). How Superduper does a Physicalist Supervenience need to be? *The Philosophical Quarterly* 49(194), 33–52.
- Wilson, J. (2009). Determination, realization and mental causation. *Philosophical Studies* 145(1), 149–169.
- Worley, S. (1997). Determination and mental causation. *Erkenntnis* 46(3), 281–304.
- Yablo, S. (1987). Identity, essence, and indiscernibility. *The Journal of Philosophy* 84(6), 293–314.
- Yablo, S. (1992a). Cause and essence. *Synthese* 93(3), 403–449.
- Yablo, S. (1992b). Mental causation. *The Philosophical Review* 101(2), 245–280.
- Yablo, S. (1997). Wide causation. *Noûs* 31, *Supplement: Philosophical Perspectives 11, Mind, Causation, and World*, 251–281.
- Yablo, S. (2000). The seven habits of highly effective thinkers. In *Proceedings of the 20th World Congress of Philosophy*, Volume 9, pp. 35–46.
- Yablo, S. (2001). Superproportionality and mind-body relations. *Theoria* 16(40), 65–75.