Interventionism and Epiphenomenalism[*]

MICHAEL BAUMGARTNER
University of Konstanz
78464 Konstanz
Germany

Abstract

In a recent paper, Shapiro and Sober (2007) defend two claims with respect to the *master argument for epiphenomenalism*, which is designed to rebut non-reductive physicalism: (i) relative to an interventionist account of causation, as most elaborately presented in (Woodward 2003), the master argument turns out to be invalid; and (ii) interventionism provides a means to experimentally uncover micro effects of macro causes. The first part of this paper takes issue with both of these claims by showing that Woodward's interventionism and non-reductive physicalism are incompatible—contrary to Shapiro's and Sober's assessment. The second part then discusses two possible modifications of Woodward's theory, both of which ensure the compatibility of interventionism and non-reductive physicalism. Nonetheless, it shall turn out that neither of those modifications suits the purposes of non-reductive physicalists.

1 Introduction

One of the central objectives Shapiro and Sober pursue in (2007) is to show that what they call the *master argument for epiphenomenalism*, which is a type of causal exclusion argument, fails. Epiphenomenalism, according to the terminology adopted in (Shapiro and Sober 2007), designates the thesis that supervening macro properties (or variables or factors) have no causal influence on micro properties that are caused by the micro supervenience bases of those macro properties. Well-known classical exclusion arguments are designed to yield such macro-to-micro epiphenomenalism along the lines of the following reasoning: subject to the widely accepted principle of the causal closure of the physical, there exists a causally sufficient micro cause for every micro effect; if it is additionally assumed that macro properties supervene on micro properties without being identical (or reducible) to the latter and if—in light of the rareness of cases of causal overdetermination—micro effects are assumed not to be systematically overdetermined, it follows that macro properties are causally inert with respect to effects

[*]I thank Frederick Eberhardt, Jim Woodward, Delphine Chapuis-Schmitz, Derek Turner, and Mehmet Elgin for very helpful comments and discussions. Moreover, two anonymous referees of this journal have contributed valuable comments on an earlier draft. Finally, I am indebted to the Center for Philosophy of Science of the University of Pittsburgh and to the Deutsche Forschungsgemeinschaft (DFG) for generous support of this work (project CAUSAPROBA).

of their micro supervenience bases.[1] This result directly contradicts one of the principal tenets of non-reductive physicalists who, in a nutshell, subscribe both to the causal closure of the physical and to the non-reductive supervenience of macro on micro properties, yet nonetheless insist on the possibility of downward causal dependencies as are e.g. involved in mental-to-physical or biological-to-physical causation.

Shapiro and Sober (2007) reject the traditional line of reasoning against non-reductive physicalism by claiming that an interventionist theory of causation, as most exhaustively developed in (Woodward 2003), which has gained wide acceptance and popularity in recent years does not rule out causal dependencies among supervening macro properties and effects of their supervenience bases. In consequence, they argue that if the master argument for epiphenomenalism is based on an interventionist notion of causation it is rendered invalid. Shapiro and Sober (2007) not only hold that (i) interventionism blocks the epiphenomenalist's master argument and, thus, immunizes non-reductive physicalism against exclusion arguments by paving the way for a causal interpretation of downward dependencies, they moreover maintain that (ii) the interventionist framework provides a methodology or tool to identify the micro effects of macro properties.[2]

In this paper, I intend to challenge both of the claims (i) and (ii). It shall be shown that spelling out causation in interventionist terms in no way blocks the master argument, but, rather, gives rise to a self-contained *interventionist* master argument for epiphenomenalism which, very broadly, involves the following three premises: (I) Woodward's interventionist theory of causation, (II) the assumption that macro properties supervene on micro properties without being identical to the latter, and (III) the assumption that the micro supervenience bases of macro properties are causally connected to the arguable micro effects of those macro properties. These premises—contrary to Shapiro's and Sober's assessment—shall be demonstrated to imply that macro properties are causally irrelevant to effects of their supervenience bases. In sum, thus, the first part of this paper aims to substantiate that, rather than supporting non-reductive physicalism, Woodward's interventionism is incompatible with non-reductive physicalism. The second part then discusses two conceivable modifications of Woodward's theory that both ensure the compatibility of interventionism and non-reductive physicalism. Nonetheless, neither of these modifications shall turn out to suit the purposes of non-reductive physicalists, for neither of them allows for establishing the actual existence of macro-to-micro causation nor for testing the downward causal powers of supervening properties.

Endorsing the epiphenomenalist conclusion of the master argument is, despite the latter's validity, not my aim here. I agree with Shapiro and Sober as regards the

[1] For details cf. e.g. (Kim 1998) or (Kim 2005). While exclusion arguments were originally designed to demonstrate that mental properties do not have physical effects, they have meanwhile been generalized to show the nonexistence of any kind of macro-to-micro causation involving supervening macro properties and effects of their supervenience bases (cf. Bontly 2002).

[2] Related claims have very recently been advanced in (Shapiro forthcoming) and (Raatikainen forthcoming).

falsity of the conclusion of the master argument. What is more, macro-to-micro causation cannot actually be ruled out as impossible on the mere basis of (I) to (III). Whether there exists macro-to-micro causation is a matter that is to be settled by scientific investigation and not by some a priori philosophical argument (cf. Shapiro and Sober 2007, 259). Accordingly, I take the validity of the master argument and the falsity of its conclusion to count against one of its premises. While I do not want to quarrel with the causal closure of the physical which implies the existence of a causally sufficient micro cause for every micro effect and, thus, entails premise (III), both (I) and (II) are candidates for rejection. On the one hand, it might be argued that a theory of causation that rules out causal dependencies among macro properties and effects of their supervenience bases on mere conceptual grounds, as does Woodward's (2003) variant of interventionism, is too restrictive and, thus, constitutes an inadequate account of causation. In consequence, one could settle for a modification of interventionism as proposed in the second half of this paper. On the other hand, the relationship between macro and micro properties might not be spelled out in terms of non-reductive supervenience, but rather in a way that allows for the reduction of the former to the latter.[3] Assessing the details and prospects of rejecting either premise (I) or premise (II), however, is beyond the scope of this paper. Rather I shall content myself with establishing that, first, the version of interventionism which is commonly thought to be most plausible excludes a causal influence of macro properties on effects of their supervenience bases and that, second, obvious modifications of interventionism are unsuited for the purposes of non-reductive physicalists. All in all, existing interventionist theories of causation whose popularity is constantly growing since the early 2000s in no way diminish the challenge that the exclusion argument or—to use Shapiro's and Sober's term—the master argument for epiphenomenalism creates for the position of non-reductive physicalists.

Section 2 briefly reviews the theoretical fundament of interventionism—as most clearly and prominently presented in (Woodward 2003)—and identifies three crucial implications of this theory. In section 3, I then present the interventionist master argument for epiphenomenalism and prove its validity. Finally, sections 4 and 5 discuss the two modifications of interventionism that suggest themselves and find them unsuitable for non-reductive physicalism.

2 Interventionism

As Woodward's interventionist theory of causation is of crucial importance to Shapiro's and Sober's (2007) argument, its core is concisely recapitulated in this section. Woodward's theory centers on two interdependent definitions. First, he defines the notions of a direct and of a contributing cause:

(M) A necessary and sufficient condition for X to be a (type-level) *direct cause* of Y with respect to a variable set $\mathbf{V}$ is that there be a possible intervention

[3] Cf. e.g. (Kim 1998) or (Kim 2005).

> on X that will change Y or the probability distribution of Y when one holds fixed at some value all other variables Z_i in $\mathbf{V}$. A necessary and sufficient condition for X to be a (type-level) *contributing cause* of Y with respect to variable set $\mathbf{V}$ is that (i) there be a directed path from X to Y such that each link in this path is a direct causal relationship (...); and that (ii) there be some intervention on X that will change Y when all other variables in $\mathbf{V}$ that are not on this path are fixed at some value. (Woodward 2003, 59)

Against this background, a variable X is a cause of Y iff there exists at least one set $\mathbf{V}$ with respect to which X is either a direct or a contributing cause of Y.[4] Second, Woodward defines the notion of an intervention variable:

(IV) I is an intervention variable for X with respect to Y iff

1. I causes X;
2. I acts as a switch for all other variables that cause X. That is, certain values of I are such that when I attains those values, X ceases to depend on the values of other variables that cause X and instead depends only on the value taken by I;
3. Any directed path from I to Y goes through X. That is, I does not directly cause Y and is not a cause of any causes of Y that are distinct from X except, of course, for those causes of Y, if any, that are built into the $I - X - Y$ connection itself; that is, except for (a) any causes of Y that are effects of X (i.e., variables that are causally between X and Y) and (b) any causes of Y that are between I and X and have no effect on Y independently of X;
4. I is (statistically) independent of any variable Z that causes Y and that is on a directed path that does not go through X. (Woodward 2003, 98)

Finally, relative to the notion of an intervention variable an (actual) *intervention* can be straightforwardly understood in terms of an intervention variable I for X with respect to Y taking on some value z_i such that $I = z_i$ causes X to take on some determinate value z_j (Woodward 2003, 98).

Before turning to the interventionist master argument for epiphenomenalism, several things need to be noted about a theory that centers on (M) and (IV). First, in several passages, Woodward (2003, 55, 60–61, 98) explicitly refers to (M) and (IV) as *definitions* of the corresponding notions. Since the notion of causation appears prominently in the definiens of (IV) and the notion of an intervention plays a crucial role in (M), causation and intervention are obviously interdefined by (M) and (IV). That, however, is not considered to be problematic by Woodward. He argues that the particular conceptual interdependence of causation and intervention induced by (M) and (IV) is neither vacuous nor viciously circular (Woodward 2003, 104–105). Even though not all authors agree that interdefining causation and intervention is

[4] Removing the relativization of causation to a particular variable set $\mathbf{V}$ along these lines is not itself contained in (Woodward 2003) but suggested in (Woodward 2008b, 207–211).

as unproblematic as Woodward would like to have it (e.g. Strevens 2007, 2008 or Baumgartner 2009a), it is clear that Woodward neither aims nor claims to provide a reductive analysis of causation. His goal is simply to elucidate the conceptual interdependence of two highly causally loaded notions, and the resulting theory is far from being uninformative. On the contrary, this paper is going to show that it has some fairly strong implications.

Second, note that what counts as a causal dependency, according to (M), hinges on a pertaining variable set $\mathbf{V}$. This relativization of causation to $\mathbf{V}$ has given rise to some controversies in the literature, for instance on the question whether a thus relativized notion of causation is monotonic and, hence, adequately captures our pre-theoretic understanding of causation which certainly presumes monotonicity (cf. Strevens 2007, 2008, Woodward 2008b). For simplicity, these issues shall be sidestepped here. Whoever has reservations against such a relativization can, for the purposes of this paper, just consider (M) as supplying a technical term of art that reflects how causation is understood within the interventionist framework. Third, a theory based on (M) and (IV) differs from traditional reductive interventionist accounts—as e.g. advanced in (Menzies and Price 1993)—in not involving the notion of human action. (IV) identifies intervention variables solely based on their causal (and statistical) relations to the other variables in a given structure. Fourth, it is a variant of a counterfactual analysis of causation because the notion of a *possible* intervention contained in (M), according to Woodward (2003, 70-71), is to be understood in terms of counterfactual conditionals of type "if some intervention were to occur on X with respect to Y, Y (or the probability of Y) would change".[5] Woodward's theory, thus, establishes a tight conceptual connection between manipulability (or counterfactual manipulations), difference-making in context, and causality, which Woodward (2003, 61) sums up in the following slogan: No causal difference without a difference in manipulability relations, and no difference in manipulability relations without a causal difference.

Finally and most importantly, Woodward's interventionism provides three necessary conditions for causation that are of crucial importance for our subsequent discussion of the epiphenomenalist's master argument and that, accordingly, must be clearly understood before we can move on to apply (M) and (IV) to cases of macro-to-micro causation.[6] The core of (M) is constituted by an analysis of the notion of a direct cause. What a contributing cause amounts to is then spelled out by drawing on direct causation, to the effect that two variables are determined to be causally connected iff they are related in terms of either direct or contributing causation. Overall, the analysans of causation supplied by (M) stipulates that if X is a (type level) cause of Y, then, on the one hand, there exists a possible intervention

[5] Cf. also (Woodward 2008a).

[6] As I show in (Baumgartner 2009b), definition (M) can be given three significantly different readings, only one of which actually yields all of the necessary conditions discussed in this section. Nonetheless, (Baumgartner 2009b) also exhibits that this strongest of all possible readings of (M) is the by far most plausible one, i.e. the one that most adequately expresses the basic intuitions behind interventionism.

$I = z_i$ on X with respect to Y and, on the other, while this possible intervention is performed on X all the variables in $\mathbf{V}$ that are not located on a causal path from X to Y are fixed at some value, i.e. these other variables in $\mathbf{V}$ *can* be fixed while $I = z_i$ is performed. If Y changes its value or probability distribution in a possible (or counterfactual) scenario that satisfies these two conditions, (M) yields that X is a cause of Y. That is, (M) identifies *manipulability* of X and *fixability* of the variables in $\mathbf{V}$ that are not located on a path from X to Y as two necessary conditions for X to cause Y. For easy reference later on, let us label these conditions:

(MAN) There exists a possible intervention $I = z_i$ on X with respect to Y.

(FIX) The variables in $\mathbf{V}$ that are not located on a causal path from X to Y can be held fixed while $I = z_i$ is performed on X.

Moreover, (IV)—more precisely, the conjunction of (IV.1) and (IV.4)—stipulates that an intervention variable on X with respect to Y must cause changes in the values (or the probability distribution) of Y and must be (statistically) independent of all causes of Y that are not located on a path that goes through X. If there does not possibly exist a variable I that meets this *independence requirement*, there does not exist a possible intervention on X with respect to Y which implies that (MAN) cannot be satisfied which, again, implies that X does not cause Y. Therefore, in view of the fact that (MAN) is necessary for causation, according to Woodward's interventionism, so is (IND):

(IND) There possibly exists a variable I that causes changes in the values (or the probability distribution) of Y and is (statistically) independent of any variable Z that causes Y and that is not located on a causal path from X to Y.

Woodward's theory determines that if either (IND), (MAN), or (FIX) cannot be satisfied by two variables X and Y relative to a variable set $\mathbf{V}$, then X and Y are not causally connected relative to $\mathbf{V}$. While (IND) and (FIX) are logically independent, (IND) and (MAN) are not.[7] For all variables X, Y and all sets $\mathbf{V}$ it holds that if (IND) cannot be satisfied, then (MAN) cannot be satisfied, then, trivially, the conjunction of (MAN) and (FIX) cannot be satisfied, then X does not cause Y with respect to $\mathbf{V}$:[8]

$$\begin{aligned}&\text{for all } X, Y, \mathbf{V}:\\ &\neg(\text{IND}) \rightarrow \neg((\text{MAN}) \wedge (\text{FIX})) \rightarrow \neg(X \text{ causes } Y \text{ with respect to } \mathbf{V})\end{aligned} \tag{Φ}$$

[7] Prima facie, it might be thought that there also exists some form of logical dependence between (FIX) and (IND): if not all variables in $\mathbf{V}$ that are not located on a causal path from X to Y can be held fixed while $I = z_i$ is performed on X, then it might appear to follow that I is not independent of all other causes of Y apart from X, i.e. ¬(FIX)→ ¬(IND). However, as (FIX) is relativized to the set $\mathbf{V}$, while (IND) is not, that logical dependence in fact does not hold. If $\mathbf{V}$ is chosen thus that it contains no other causes of Y apart from X, violations of (FIX) have no implications for the satisfiability of (IND).

[8] Note that (MAN), (FIX), and (IND) are modal statements. For simplicity, this modality is not explicitly formally expressed in schema (Φ).

Let us now look at the consequences a theory of causation that subscribes to (Φ) has for macro-to-micro causation.

3 The Master Argument

Shapiro and Sober (2007) do not present what they call the *master argument for epiphenomenalism* in schematic form. Rather, they introduce the argument in the following very informal passage:

> How could believing or wanting or feeling cause behavior? Given that any instance of a mental property X has a physical microsupervenience base MSB(X), it would appear that X has no causal powers in *addition* to those that MSB(X) already possesses. The absence of these additional causal powers is then taken to show that the mental property X is causally inert. We call this the master argument for epiphenomenalism (...). (Shapiro and Sober 2007, 241)

In virtue of this informal presentation, the master argument only seems to have two premises: first, a mental or, more generally, macro property X has a physical microsupervenience base MSB(X) and, second, MSB(X) has causal powers on a behavioral or, more generally, micro effect Y. These premises are then said to apparently imply that X is causally inert. Put in this truncated way, the argument is obviously invalid, for those two premises are perfectly compatible with X having causal powers on Y in addition to those of MSB(X). What is required to get the epiphenomenalist conclusion is a further premise which, in one way or another, excludes the overdetermination of Y. Hence, I read the passage quoted above as an abbreviation of the argument Shapiro and Sober really have in mind, which is a sort of exclusion argument, whose premises are somewhat dispersed through their paper. This section, thus, assembles the premises and conclusion of the interventionist master argument as transparently as possible and presents my reconstruction of the argument.

As a first premise, Shapiro and Sober (2007, 237) presuppose Woodward's interventionist notion of causation as expressed in (M). As a second premise, they then (241) assume a macro property X to supervene on a physical micro supervenience base MSB(X), without being identical to MSB(X) (247–249), which, as third premise, is throughout the paper (especially 237–239) presumed to be a cause of a physical effect Y. According to Shapiro and Sober (2007, 241), these three premises are commonly *falsely* taken to imply that X does not cause (or is causally irrelevant to) Y. They argue that the premises of the master argument do not in fact imply the downward causal inertness of X.

In order to spell out the three premises, the intermediate inferential steps, and the conclusion of the master argument in more detail and to assess its validity, let us draw on the classical case of downward mental causation as illustrated in figure 1. Suppose, thus, we are looking at the variable set $\mathbf{V} = \{M_1, M_2, P_1, P_2\}$, where M_1 and M_2 represent two types of mental phenomena and P_1 and P_2 represent

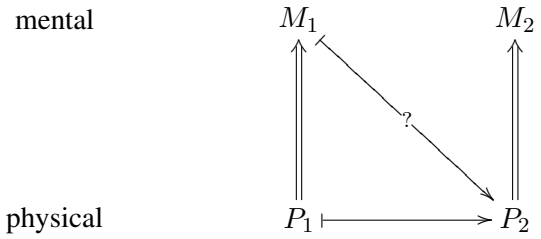

Fig. 1: The classical question of macro-to-micro causation: if the mental phenomenon M_1 supervenes on the physical phenomenon P_1 (symbolized by "$\Longrightarrow$"), which, in turn, is a cause of the physical phenomenon P_2 (symbolized by "$\longmapsto$"), in what sense—if at all—can M_1 then be said to be a cause of P_2?

the two types of physical phenomena that realize particular values of M_1 and M_2, respectively. M_1 and M_2 are taken to supervene on P_1 and P_2. That is, the sets of possible values of P_1 and P_2 constitute the physical supervenience bases of M_1 and M_2, i.e. $\text{MSB}(M_1) = \{P_1 = y_1, P_1 = y_2, \ldots, P_1 = y_n\}$ and $\text{MSB}(M_2) = \{P_2 = z_1, P_2 = z_2, \ldots, P_2 = z_m\}$. It furthermore holds that $M_1 \neq P_1$ and $M_2 \neq P_2$. Finally, P_1 shall be assumed to be a cause of P_2. Relative to this context, the question as to the possibility of macro-to-micro causation amounts to the question whether M_1 can be said to cause P_2. Is it in fact the case—as claim Shapiro and Sober—that interventionism provides a positive answer to that question?

As we have seen in the previous section, spelling out "M_1 is a cause of P_2 with respect to the variable set $\mathbf{V} = \{M_1, M_2, P_1, P_2\}$" in interventionist terms amounts to this:

(1) M_1 is a cause of P_2 with respect to the variable set $\mathbf{V} = \{M_1, M_2, P_1, P_2\}$ iff there possibly exists an (IV)-defined intervention $I_1 = z_1$ on M_1 with respect to P_2 such that all other variables in $\mathbf{V}$ that are not located on a causal path from M_1 to P_2 are held fixed and the value or the probability distribution of P_2 changes. [*assumption*]

Section 2 has also shown that (1) implies, among other things, that if it is impossible that there exists a variable I_1 that causes changes in M_1 while being statistically independent of all causes of P_2 that are not located on a path from M_1 to P_2, M_1 cannot be manipulated with respect to P_2. This, in turn, entails that M_1 does not cause P_2. Or contrapositively put, if M_1 is a cause of P_2 in the sense given by (M) and (IV), it follows that there possibly exists a variable I_1 that satisfies (IND) with respect to M_1 and P_2. Assumption (1), thus, implies the following conditional:

(2) If M_1 is a cause of P_2 with respect to the variable set $\mathbf{V} = \{M_1, M_2, P_1, P_2\}$, then there possibly exists a variable I_1 that causes changes in the values (or the probability distribution) of M_1 and is statistically independent of any variable Z that causes P_2 and that is not located on a causal path from M_1 to P_2. [*from (1)*]

Next, we follow Shapiro and Sober in assuming that macro properties non-reductively supervene on micro properties and that all micro effects are causally determined by corresponding micro causes. Applied to our exemplary case, these assumptions amount to:

(3) M_1 supervenes on MSB$(M_1) = \{P_1 = y_1, P_1 = y_2, \ldots, P_1 = y_n\}$ without being identical to P_1. [*assumption*]

(4) P_1 is a cause of P_2. [*assumption*]

The exact meaning and implications of (3), of course, hinge on the notion of supervenience presupposed. As is well known, supervenience has been cashed out in a number of very different ways in the literature.[9] However, there are at least two features shared by all of these notions: first, supervenience is a non-causal relation and, second, every change in the supervening property or variable is necessarily accompanied by a change in the supervenience base. Thus, whatever version of supervenience is taken to be presupposed by (3), the latter implies that M_1 and P_1 differ, that M_1 and P_1 are not causally related, and that changes in M_1 are necessarily accompanied by changes in P_1. Or more specifically:

(5) $M_1 \neq P_1 \wedge \neg(M_1 \text{ causes } P_1) \wedge \neg(P_1 \text{ causes } M_1)$ [*from (3)*]

(6) Every change in the values of M_1 is necessarily accompanied by a change in the values of P_1. [*from (3)*]

Now, if P_1 is a cause of P_2, as stated in (4), and if the relationship between the two different variables P_1 and M_1 is not of causal nature, as advanced in (5), we get:

(7) P_1 is on a causal path to P_2 that does not include M_1. [*from (4),(5)*]

In other words, P_1 is one of those causes of P_2 of which possible intervention variables on M_1 with respect to P_2 must be independent according to (IND). That is, from the conjunction of (2) and (7) it follows that:

(8) If M_1 is a cause of P_2 with respect to the variable set $\mathbf{V} = \{M_1, M_2, P_1, P_2\}$, then there possibly exists a variable I_1 that causes changes in M_1 while being statistically independent of changes in P_1. [*from (2),(7)*]

The consequent of (8), however, is determined to be unsatisfiable by (6), which rules that the values of every variable that induces changes in M_1 will necessarily be correlated with changes in the values of P_1. That means there cannot possibly exist an intervention variable for M_1 with respect to P_2. A straightforward application of modus tollens to (8) then leads to:

(9) $\neg(M_1$ is a cause of P_2 with respect to the variable set $\mathbf{V} = \{M_1, M_2, P_1, P_2\})$ [*from (6),(8)*]

[9] Cf. e.g. (McLaughlin 1995) or (Bennett 2004).

Hence, the assumptions in lines (1), (3), and (4) which correspond to the three premises of Shapiro's and Sober's master argument indeed yield the epiphenomenalist conclusion. Furthermore, the downward causal inertness of M_1 resulting from this argument does not depend on the choice of a particular set $\mathbf{V}$. (7) entails that there exists a cause of P_2 that is not on a path that includes M_1, *viz.* P_1. Since the notion of an intervention provided by (IV) is not relativized to a specific variable set, every possible intervention on M_1 must be independent of P_1—irrespective of whether P_1 is contained in $\mathbf{V}$ or not. (6), however, excludes the possible existence of such an intervention. Or differently, because (IND) cannot be satisfied for M_1 and P_2, there does not possibly exist a set $\mathbf{V}$ relative to which M_1 could be (IV)-manipulable with respect to P_2. If (IND) cannot be satisfied for two variables, (MAN) cannot be satisfied either, which entails that these two variables are not causally connected. That is, (1), (3), and (4) imply that M_1 does not cause P_2 *simpliciter*.[10]

As can easily be seen from the above considerations, our exemplary case of figure 1 not only violates (IND) and (MAN) but also the third necessary condition for causation the interventionist framework provides, *viz.* (FIX). (7) states that P_1 is located on a causal path to P_2 that does not include M_1, which according to (FIX) requires that P_1 be fixed while M_1 is manipulated. Yet (6) excludes just that fixability, i.e. (6) excludes that P_1 can possibly be held fixed while M_1 is manipulated. Therefore, M_1, P_1, and $\mathbf{V}$ also violate (FIX).

The violations of (IND), (MAN), and (FIX) establish the causal irrelevance of M_1 to P_2 in virtue of (M) and (IV). Thus, the interventionist master argument for epiphenomenalism as reconstructed above is clearly a valid argument—contrary to Shapiro's and Sober's assessment. The master argument shows that M_1 does not cause P_2 in terms of an interventionist notion of causation as advanced in (Woodward 2003) and, since M_1 and M_2 represent any properties that supervene on the causally connected variables P_1 and P_2, the conclusion as to the causal irrelevance of M_1 to P_2 can be generalized for all macro properties that supervene on causally connected micro supervenience bases. The argument reveals that (M) and (IV) rule out the existence of causal dependencies among supervening macro properties, on the one hand, and micro properties that are causally connected to the supervenience bases of the corresponding macro properties, on the other. It is not consistent, first, to spell out causation in Woodward's terms, second, to claim that macro properties non-reductively supervene on micro properties, and third, to nonetheless insist on the causal efficacy of macro properties on effects of their supervenience bases. Woodward's variant of interventionism and non-reductive physicalism are *incompatible*.

As indicated in the introduction, Shapiro and Sober, however, reject the master argument as invalid. In order for them to come to their different validity assess-

[10] Note that this does entail that M_1 cannot be claimed to cause M_2 within the interventionist framework. As long as P_1 is only seen to cause P_2 but not M_2 (due to supervenience which is non-causal), P_1 may vary in accordance with (IND) while M_1 is manipulated with respect to M_2. Cf. also (Macdonald 2007).

ment they must spell out at least one of its premises differently from the argument reconstruction presented here. The next section will show that Shapiro and Sober explicitly endorse assumptions (3) and (4). Hence, (1) turns out to be the only remaining candidate for disagreement. Any interventionist account of causation which, as the one developed in (Woodward 2003), entails that (IND), (MAN), and (FIX) are necessary conditions for causation gives rise to the master argument. Every version of interventionism that entails (Φ) is incompatible with non-reductive physicalism. I therefore take Shapiro's and Sober's rejection of the master argument to indicate that they tacitly have a variant of interventionism in mind that is significantly weaker than Woodward's—notwithstanding their endorsement of Woodward's theory.

The rest of this paper will be concerned with two possible ways to weaken Woodward's interventionism such that the interventionist master argument presented in this section is blocked and interventionism is rendered compatible with non-reductive physicalism. The first of these conceivable modifications—considered in section 4—seems to underlie Shapiro's and Sober's dismissal of the master argument, the second—discussed in section 5—is a suggestion of my own.

4 *Non-definitional Interventionism*

Shapiro's and Sober's rejection of the master argument ultimately springs from their conviction that which downward causal dependencies exist, if any, "is something for science, not armchair philosophy, to determine" (Shapiro and Sober 2007, 259). However, if the reconstruction of the master argument proposed in the previous section is accepted, the possibility of macro-to-micro causation can be ruled out without having to conduct scientific studies. On the face of it, any of the assumptions (1), (3), and (4) is a candidate for rejection. Nevertheless, the three assumptions differ greatly as to how plausible or consequential their respective rejections are. (4) is a direct consequence of the causal closure of the physical which is an (almost) uncontroversial principle in the pertinent literature. Shapiro and Sober are far from wanting to quarrel with (4). An alternative to (3) that suggests itself would be to somehow reduce the macro (mental) property to its physical realizer. Shapiro and Sober (2007, 246–247) discuss this option by drawing on Kim's functional model of reduction as an example. Kim (1998, 106–120) suggests, first, to functionalize macro properties by defining them in terms of their functional roles and, second, to reduce instances of these functional properties to their physical realizers by identifying them with the latter relative to particular causal structures under consideration. Clearly, if (3) is replaced by an assumption stating the reducibility of M_1 to P_1 along the lines of Kim, (5) can no longer be derived which is required to get (7). If the functionalized values of M_1 are identical to particular values of P_1, M_1 and P_1 can no longer be seen to be distinct variables which implies that their statistical dependence and the impossibility to manipulate one while holding the other fixed no longer violate (IND), (MAN), and (FIX). That is, replac-

ing (3) by an assumption in the vein of Kim's functional model of reduction would indeed block the master argument. For the following reasons, however, Shapiro and Sober do not opt for a Kimian way around the master argument:

> In the first place, Kim intends his solution to work only for macroproperties that can be functionalized. Kim believes that qualia cannot be functionalized and concludes that they cannot be causes (1998, 116). In contrast, from the empirical standpoint we have advocated, whether a macroproperty is functionalizable makes no difference to whether it has causal powers. Intervention provides a means by which to test which causal powers a macroproperty has. (Shapiro and Sober 2007, 247)

Hence, Shapiro's and Sober's empirical standpoint according to which the causal powers of a macro property are a matter for science to determine not only accounts for why they reject the master argument in the first place but also for why they discard a Kimian way around that argument. No macro property, irrespective of whether it is functionalizable and, thus, reducible to a micro property or not, can be ascribed downward causal inertness on a priori philosophical grounds. Even macro properties that, in the end, do not turn out to be reducible to micro properties could be found to have downward causal powers in pertaining scientific studies. For later reference, let us furnish this non-reductionist and empiricist foundation of Shapiro's and Sober's rejection of the master argument with a label:

(A) Which causal dependencies among macro and micro variables exist, if any, is to be determined by scientific investigation, independently of whether macro variables are reducible to micro variables. Rather than ruling out macro-to-micro causation, the interventionist framework provides a means by which to test which micro effects a macro variable has.

As Shapiro and Sober do not want to dispute the causal closure of the physical nor to plead for the general reducibility of macro to micro variables, they are left with (1) as only remaining locus for disagreement with the argument reconstruction proposed in the previous section. Their rejection of the master argument must be based on an interventionist account of causation that differs significantly from a theory that centers on (M) and (IV)—notwithstanding the fact that they explicitly base their discussion on Woodward's interventionism (cf. Shapiro and Sober 2007, 237, 256).

In the following passage, in which Shapiro and Sober indicate their reasons for rejecting the master argument as fallacious, the differences between their tacit understanding of interventionism and Woodward's theory become most transparent:

> The crucial mistake in this line of reasoning [i.e. in the master argument] is that it requires one to consider a counterfactual situation that is in fact impossible and is, in any case, irrelevant to the question of whether the mental property X, or any other supervening property, is epiphenomenal with respect to an effect term Y. To see if X has an effect on Y that is additional to whatever effect MSB(X) has on Y, one would have to compare what would

> happen to Y if both MSB(X) and X were present with what would happen to Y if MSB(X) were present and X were absent. The master argument purports to evaluate this counterfactual and then concludes that the mental property X makes no contribution to Y additional to the effect that MSB(X) has. The conclusion is then drawn that mental properties are causally inert. The principal fallacy is the thought that if X causes Y, then X must have an impact on Y additional to the impact on Y that MSB(X) has. (...) The lesson we draw from the master argument is that one must be very careful in deploying "holding-fixed arguments." To assess whether X causes Y, the common causes of X and Y must be held fixed, but not the microsupervenience base of X. (Shapiro and Sober 2007, 241)

Shapiro and Sober here point to two reasons for their dismissal of the master argument that reflect an interventionist account of causation that substantially deviates from a theory that entails (Φ):[11]

(B) To assess whether X causes Y, the common causes of X and Y must be held fixed, but not the micro supervenience base of X. X does not have to have an impact on Y additional to the impact on Y that MSB(X) has.

(C) To see if X has an effect on Y that is additional to whatever effect MSB(X) has on Y, one would have to compare what would happen to Y if both MSB(X) and X were present with what would happen to Y if MSB(X) were present and X were absent. This counterfactual situation is impossible and is irrelevant to the question of whether the mental property X is epiphenomenal with respect to an effect term Y.

Applied to the context of our exemplary instance of the master argument, reason (B) amounts to the claim that, contrary to what has been stipulated in the previous section, only the common causes of M_1 and P_2 need to be held fixed, not however P_1. Yet, relative to the setup of the structure in figure 1, P_1 is clearly located on a causal path to P_2 that does not include M_1. The master argument reproduces this property of P_1 by the fact that the conjunction of (3) and (4) implies (7). Furthermore, from (7) and (FIX) it follows that to assess whether M_1 causes P_2, P_1 must be held fixed while M_1 is manipulated which contradicts (B). As Shapiro and Sober do not challenge (3) and (4), reason (B) for their rejection of the master argument suggests that they do not take (FIX) to be a necessary condition for a causal dependency between M_1 and P_2.

Adapted to our exemplary context, (C) essentially states that manipulating M_1 independently of P_1 is impossible and irrelevant to whether M_1 causes P_2. It is not only impossible to hold P_1 fixed while intervening on M_1, it is even impossible to intervene on M_1 in the first place. As shown in the previous section, there does not possibly exist an (IV)-defined intervention variable for M_1 with respect to P_2, for (IND) cannot be satisfied for M_1 and P_2. In (C), Shapiro and Sober implicitly

[11] A very similar critique of epiphenomenalist argumentative strategies can be found in (Shapiro forthcoming).

acknowledge the impossibility to (IV)-manipulate M_1 with respect to P_2. Yet, instead of concluding from this observation that M_1 does not cause P_2, as would be implied by a variant of interventionism that entails (Φ), they continue to claim that the non-manipulability of M_1 is *irrelevant* to whether M_1 causes P_2. That is, reason (C) suggests that they do not have an interventionist account in mind according to which (IND) and (MAN) are necessary for causation.

In sum, Shapiro and Sober apparently conceive of interventionism in such a way that it does not imply (Φ). For them, neither (IND) nor (MAN) nor (FIX) are necessary conditions for two variables to be causally dependent. Indeed, if (Φ) is not taken to follow from an interventionist analysis, the first premise of the master argument is weakened to such a degree that it no longer implies (2) which, in turn, blocks the master argument. This finding raises the question as to the exact details of the interventionist theory Shapiro and Sober have in mind. Unfortunately, there is no evident answer to this question because they do not make their understanding of interventionism explicit. In what follows, I shall therefore try to reconstruct their weakened version of interventionism on my own.

As long as a biconditional is taken to be the main operator governing (M)'s truth conditions, violations of (IND), (MAN), and (FIX) will turn out to be relevant—in one way or another—to whether X causes Y or not. The relevance of such violations, however, disappears as soon as an interventionist theory is no longer conceived of as providing a conceptual analysis or definition of causation. An interventionist theory that does not claim to provide both necessary and sufficient conditions for any two variables X and Y to be causally connected can confine itself to analyzing what causation amounts to in contexts where interventions on X are possible and the other variables in an investigated structure can be fixed at some value. Against such a background, (IND), (MAN), and (FIX), rather than being necessary for causation, constitute *preconditions* for applying a corresponding interventionist analysis in the first place. As it turns out, *non-definitional* interventionism, thus construed, indeed backs up both Shapiro's and Sober's rejection of the master argument and their claim that violations of (IND), (MAN), and (FIX) are of no relevance to whether two variables X and Y are causally connected. These considerations induce a modification of (M) along the following lines:

(M*) If there possibly exists a scenario s such that, in s, an (IV)-defined intervention $I = z_i$ is performed on X with respect to Y while all variables in a corresponding variable set $\mathbf{V}$ that are not located on a path from X to Y are held fixed, then X is a (type-level) cause[12] of Y with respect to $\mathbf{V}$ iff Y changes its value or its probability distribution in s.

[12] As the distinction between direct and contributing causes is of no relevance to the master argument, it can be neglected here. (M*) (just as (M**) below) covers both direct and contributing causes. If X is a direct cause of Y, the path from X to Y includes no other variables in $\mathbf{V}$, which hence all have to be fixable in order to apply (M*), if X is a contributing cause of Y, a proper subset of $\mathbf{V} \setminus \{X, Y\}$ has to be fixable.

(M*) is weaker than (M), i.e. (M) implies (M*) but not vice versa. Contrary to (M), (M*) is only applicable if there possibly exists a scenario of type s, i.e. if (IND), (MAN), and (FIX) are satisfiable. In a recent paper, Woodward himself suggests a modification of his theory along the lines of (M*):

> I also assume that if a candidate causal claim is associated with interventions that are impossible for (or lack any clear sense because of) logical, conceptual or perhaps metaphysical reasons, then that causal claim is itself illegitimate or ill-defined. In other words, I take it to be an implication of (M) that a legitimate causal claim should have an intelligible interpretation in terms of counterfactuals [i.e. claims about the possible existence of interventions] the antecedents of which are coherent or make sense. (...) Thus if we have two apparently competing claims, the first contending some mental state is causally inert and the other contending that it causes some outcome, it must be possible to specify some set of (coherent, well-defined) interventions such that the two claims make competing predictions about what would happen under those interventions. If we cannot associate such an interventionist interpretation with one or both of the claims, the claim(s) in question lack a clear sense. (Woodward 2008a, 224–225)

Plainly, if interventions on X with respect to Y are impossible, (M) and (IV) do not yield that "X causes Y" is ill-defined—contrary to what Woodward claims in this passage. Rather, as (M) and (IV) entail (Φ), they render causal claims false whenever interventions on X are impossible. In contrast, (M*) indeed does not assign a truth-value to "X causes Y" when it is impossible to intervene on X with respect to Y. Subject to (M*), the possibility to intervene on X with respect to Y can—in the vein of the passage quoted above—be argued to be a precondition of "X causes Y" being a well-defined interventionist causal claim. Or put differently, violations of (IND), (MAN), and (FIX) can—as required by Shapiro and Sober—be said to be irrelevant to whether X causes Y in virtue of (M*). When (IND), (MAN), or (FIX) are violated, the existence of a causal dependency between X and Y must be assessed on the basis of some other account of causation.

If we modify assumption (1) of the master argument along the lines of (M*), we get:

(1^*) If there possibly exists a scenario s such that, in s, an (IV)-defined intervention $I_1 = z_1$ is performed on M_1 with respect to P_2 while all variables in the set $\mathbf{V} = \{M_1, M_2, P_1, P_2\}$ that are not located on a path from M_1 to P_2 are held fixed, then M_1 is a (type-level) cause of P_2 with respect to $\mathbf{V}$ iff P_2 changes its value or its probability distribution in s.

(1^*) does not imply (2) and, hence, in combination with (3) and (4) does not allow to derive the conclusion of the master argument. That means (M*) has the two features that Shapiro and Sober require for their rejection of the master argument: it blocks the master argument and it does so because—as to (M*)—(IND), (MAN), and (FIX) are not necessary and violations thereof irrelevant for causation. On the other hand, of course, (M*) is significantly weaker than (M). (M*) does not

provide an analysis of causation. Whenever one of the conditions mentioned in its antecedent is violated, (M*) does not shed any light whatsoever on whether X causes Y. In cases of violations of (IND), (MAN), and (FIX), (M*) is simply inapplicable.

Indeed, the thus restricted applicability of (M*) constitutes such a substantial weakness that non-reductive physicalists cannot really consider (M*) a viable alternative foundation for interventionism—notwithstanding the fact that non-definitional interventionism does not give rise to the master argument and is, hence, compatible with non-reductive physicalism. Non-reductive physicalists do not want to claim that causal statements expressing downward causation are ill-defined or meaningless. On the contrary, they hold that supervening macro properties are *causally relevant* to effects of their supervenience bases, i.e. they hold that "M_1 causes P_2" is meaningful and moreover true. Furthermore, as we have seen above, non-reductive physicalists with sympathies for interventionism, as Shapiro and Sober, additionally maintain that the truth of "M_1 causes P_2" can be assessed by interventionist means. Yet, although Shapiro and Sober can do justice to their reasons (B) and (C) for rejecting the master argument by endorsing a non-definitional variant of interventionism along the lines of (M*), the latter does not provide a means to determine which causal dependencies, if any, subsist among supervening macro properties and effects of their supervenience bases. Rather, variables as M_1 and P_2 violate the preconditions for applying (M*). In consequence, even though a causal dependency between M_1 and P_2 is compatible with (M*), the latter is of no help whatsoever when it comes to determining whether that dependency in fact exists or not. On the basis of (M*) it is impossible to determine whether a supervening macro property is causally connected to effects of its supervenience base because supervening properties cannot be (IV)-manipulated with respect to effects of their supervenience bases. Hence, the non-definitional interventionism Shapiro and Sober tacitly plead for by advancing (B) and (C) is unsuited to meet their claim that interventionism allows to identify downward causal dependencies, i.e. unsuited to meet claim (A).

More generally put, the existence of macro-to-micro causation cannot be assessed on the basis of non-definitional interventionism. Non-reductive physicalists, however, insist on the existence of macro-to-micro causation despite the causal closure of the physical. Non-definitional interventionism provides no theoretical backing whatsoever for that position. In consequence, notwithstanding the fact that non-definitional interventionism does not rule out non-reductive physicalism, it is far from constituting a theoretical foundation for the latter.

5 *Relaxing Independence and Fixability*

Non-definitional interventionism is not the only conceivable modification of Woodward's (2003) theory which ensures the compatibility of interventionism and non-reductive physicalism. Instead of denying independence, manipulability, and fix-

ability the status of necessary conditions for causation it might also be attempted to weaken the conditions themselves. One such weakening clearly suggests itself. At the core of the interventionist exclusion problem introduced in section 3 lies the non-causal dependence of macro on micro properties induced by supervenience. Macro variables cannot be manipulated independently of their supervenience bases. This non-causal dependence precludes the manipulability of macro properties while all other variables in pertaining structures are held fixed. That is, interventionism could be rendered compatible with non-reductive physicalism by simple exemption clauses allowing interventions on macro variables to be correlated with changes in corresponding micro supervenience bases, notwithstanding the fact that the causal closure of the physical determines those micro variables themselves to be causally sufficient for the physical effects under investigation. More concretely, (M) and (IV) could be modified as follows:

(M**) X is a cause of Y with respect to the variable set $\mathbf{V}$ iff there possibly exists an (IV*)-defined intervention $I_1 = z_1$ on X with respect to Y such that all other variables in $\mathbf{V}$ that are not located on a causal path from X to Y and that are not part of the supervenience base of X are held fixed and the value or the probability distribution of Y changes.

(IV*) I is an intervention variable for X with respect to Y iff I satisfies (IV.1), (IV.2), (IV.3), and (IV.4*):

(IV.4*) I is (statistically) independent of any variable Z such that Z is a cause of Y, Z is not located on a causal path from X to Y, and Z is not part of the supervenience base of X.

The exemption clauses for supervenience relationships contained in (M**) and (IV*) yield relaxed independence and fixability requirements. (M**) and (IV*) entail that two causally connected variables X and Y and the pertaining set $\mathbf{V}$—apart from (MAN)—satisfy (IND*) and (FIX*):

(IND*) There possibly exists a variable I that causes changes in the values (or the probability distribution) of Y and is (statistically) independent of any variable Z such that: Z is a cause of Y, Z is not located on a causal path from X to Y, and Z is not part of the supervenience base of X.

(FIX*) The variables in $\mathbf{V}$ that are not located on a causal path from X to Y and are not part of the supervenience base of X can be held fixed while $I = z_i$ is performed on X.

A modified version of interventionism that relaxes independence and fixability in this manner still provides a definition of causation—contrary to a theory that centers on (M*). Instead of (Φ), (M**) and (IV*) imply:

$$\begin{array}{l}\text{for all } X, Y, \mathbf{V}: \\ \neg(\text{IND}^{*}) \rightarrow \neg((\text{MAN}) \wedge (\text{FIX}^{*})) \rightarrow \neg(X \text{ causes } Y \text{ with respect to } \mathbf{V})\end{array} \qquad (\Phi^{*})$$

Against the background of (M**) and (IV*), the interventionist master argument for epiphenomenalism does not go through any more, for, although P_1 is located on a path to P_2 that does not include M_1 in the structure of figure 1, (IV*)-defined interventions on M_1 do not need to be independent of P_1. As P_1 represents the supervenience base of M_1, P_1 may vary arbitrarily while M_1 is (IV*)-manipulated. Or differently, the structure in figure 1 does not violate (IND*) and (FIX*), which means that M_1 may well be manipulable as required for it to cause P_2 in virtue of (M**). The master argument presented in section 3 does no proceed beyond line (7) if causation is understood in terms of (M**) and (IV*).

Is interventionism with thus relaxed independence and fixability not only compatible with non-reductive physicalism—as is non-definitional interventionism—but does it moreover ground the existence of downward causal dependencies between supervening macro properties and effects of their supervenience bases? Do (M**) and (IV*), as Shapiro and Sober would like to have it, provide a means to identify micro effects of supervening macro causes? In order to answer that question, let us reconsider the standard case of mental-to-physical causation depicted in figure 1. Assume we perform an (IV*)-defined intervention on the mental property M_1 and assume furthermore that we find this intervention to be followed by a change in the value of P_2. Does this test result reveal that M_1 is a cause of P_2? Certainly not. For by (IV*)-manipulating M_1 we explicitly allowed for changes in P_1 which the non-reductive physicalist takes to be *another* cause of P_2. This other cause is not located on a path from M_1 to P_2 and, above all, is determined to be causally *sufficient* for P_2 by the causal closure of the physical. In consequence, our test result significantly underdetermines a causal inference. At least two structures can generate the result of our hypothetical test: either (i) the change in the value of P_2 is only caused by a change in the value of P_1 which necessarily accompanied our intervention on M_1 or (ii) the change in the value of P_2 is overdetermined by P_1 and M_1. Of course, this ambiguity does not only arise due to a misguided intervention in one particular experimental context, rather, (IV*)-defined interventions, in general, are not required to be independent of *all other* causes of an effect under investigation. Supervenience bases of macro variables may vary and thereby causally influence investigated effects at will when those macro variables are (IV*)-manipulated. Hence, all empirical data that result from (IV*)-interventions and that could stem from macro-to-micro causation might just as well stem from a structure that only features micro-to-micro causation. (IV*)-manipulations never induce an unambiguous inference to macro-to-micro causation. Or differently: to every causal structure $\mathcal{S}_1$ that involves at least one macro-to-micro dependency in the sense of non-reductive physicalists there exists a causal structure $\mathcal{S}_2$ that is only composed of micro-to-micro dependencies such that $\mathcal{S}_1$ and $\mathcal{S}_2$ generate the exact same (IV*)-manipulability relations, notwithstanding the fact that they differ in causal respects. That is, somebody who subscribes to (M**) and (IV*) and conceives of the relationship between macro and micro properties in terms of non-

reductive supervenience renounces one of the core principles behind interventionism, *viz.* "no causal difference without a difference in manipulability relations".[13]

Non-reductive physicalists with sympathies for interventionism might insist that weakening Woodward's theory in the vein of (M**) and (IV*) nevertheless supports their position. For whenever (IV*)-manipulations of M_1 are followed by changes in P_2, (M**) rules that M_1 causes P_2. Yet, what evidence justifies such a ruling in light of the fact that there exists an epiphenomenalist causal structure that equally accounts for the data in our (IV*)-manipulation test? The existence of a structure of type $\mathcal{S}_2$ for every structure of type $\mathcal{S}_1$ demonstrates that, in principle, there cannot exist any (IV*)-evidence for that ruling. (M**) hence rules that M_1 causes P_2 without evidential basis. The (IV*)-equivalent epiphenomenalist structure is simply excluded or disregarded by fiat. The dispute between the epiphenomenalist and the non-reductive physicalist is certainly not argumentatively decided in this manner. (M**) does not theoretically back up non-reductive physicalism, but simply begs the question against the epiphenomenalist. As a result, (M**) and (IV*) are of no use to non-reductive physicalists when it comes to answering the challenge that the classical problem of causal exclusion poses for their position—a challenge which Kim formulates in the following often cited passage:[14]

> To be a cause of P_2, M_1 must somehow ride piggyback on physical causal chains—distinct ones depending on which physical property subserves M_1 on a given occasion, in the same world or in other possible worlds. And we may ask: In virtue of what relation it bears to physical property P_1 does M_1 earn its entitlement to a free ride on the causal chain from P_1 to P_2 and to claim this causal chain to be its own? Obviously, the only significant relation M_1 bears to P_1 is supervenience. But why should supervenience confer this right on M_1? The fact of the matter is that there is only one causal process here, from P_1 to P_2, and M_1's supposed causal contribution to the production of P_2 is totally mysterious. (Kim 2005, 48)

Rendering interventionism and non-reductive physicalism compatible by relaxing independence and fixability deprives interventionism of the capacity to provide a theoretically or empirically grounded answer to this challenge. Manipulating any macro variable M_1, that is seen to differ from its supervenience base, by means of (IV*)-defined interventions does not allow to establish, in even one single case, that M_1 in fact has a downward causal influence on effects of its supervenience base. The claim of non-reductive physicalists that such macro-to-micro causation indeed exists, thus, appears no less mysterious if causation is spelled out in the vein of (M**) and (IV*)—at best, it appears dogmatic.

[13] Cf. section 2 above.

[14] In this quote, the indices of the variables are adjusted to suit the conventions adopted in this paper.

6 Conclusion

To sum up: First, the version of definitional interventionism which is broadly taken to be the most plausible one, i.e. the one presented in (Woodward 2003), yields the master argument and, thus, universal macro-to-micro epiphenomenalism. It is incompatible with non-reductive physicalism. And second, even though the two modifications of Woodward's original theory discussed in sections 4 and 5 block the master argument and, hence, ensure the compatibility of interventionism and non-reductive physicalism, none of them suits the purposes of non-reductive physicalists. While non-definitional interventionism is inapplicable to supervening properties with respect to effects of their supervenience bases, relaxing the independence and fixability requirements of Woodward's original theory gives rise to uneliminable ambiguities in contexts of causal discovery, to the effect that every case of macro-to-micro causation might just as well be a case of mere micro-to-micro causation. When it comes to theoretically accounting for macro-to-micro causation and to testing the downward causal powers of non-reductively supervening properties a theoretical framework is called for that significantly differs from all the versions of interventionism discussed in this paper. All in all, therefore, the interventionist framework which has been gaining considerable popularity in recent years is far from reducing the challenge that the problem of causal exclusion creates for the position of non-reductive physicalists. Whoever wants to causally interpret the dependencies among supervening macro and subvening micro properties either has to renounce existent versions of interventionism or has to conceive of the relation between macro and micro variables in a way that allows for the identification of the former and the latter. As non-reductive physicalists reject the second path around the master argument, they are bound not to spell out causation in terms of the prevalent variants of interventionism—notwithstanding the remarkable attractivity those theories apparently have for many non-reductive physicalists. That, of course, does not mean that there might not exist an as yet unconceived of version of interventionism that both paves the way around exclusion arguments for non-reductive physicalists and provides a methodology to uncover macro-to-micro causation. Such an interventionist theory, however, is yet to be developed.

References

Baumgartner, M. 2009a. 'Interdefining Causation and Intervention.' *Dialectica* **63** 175–194.

——-. 2009b. 'Interventionist Causal Exclusion and Non-Reductive Physicalism.' *International Studies in the Philosophy of Science* **23** 161–178.

Bennett, K. 2004. 'Global Supervenience and Dependence.' *Philosophy and Phenomenological Research* **68** 510–529.

Bontly, T. D. 2002. 'The Supervenience Argument Generalizes.' *Philosophical Studies* **109** 75–96.

Kim, J. 1998. *Mind in a Physical World*. Cambridge: MIT Press.

———-. 2005. *Physicalism or Something Near Enough*. Princeton: Princeton University Press.

Macdonald, G. 2007. 'Emergence and Causal Powers.' *Erkenntnis* **67** 239–253.

McLaughlin, B. P. 1995. 'Varieties of Supervenience.' In *Supervenience: New Essays*, E. Savellos and U. Yalcin, eds., 16–59. Cambridge: Cambridge University Press.

Menzies, P. and H. Price 1993. 'Causation As a Secondary Quality.' *British Journal for the Philosophy of Science* **44** 187–203.

Raatikainen, P. forthcoming. 'Causation, Exclusion, and the Special Sciences.' *Erkenntnis*.

Shapiro, L. forthcoming. 'Lessons From Causal Exclusion.' *Philosophy and Phenomenological Research*.

Shapiro, L. and E. Sober 2007. 'Epiphenomenalism. The Dos and Don'ts.' In *Thinking About Causes: From Greek Philosophy to Modern Physics*, G. Wolters and P. Machamer, eds., 235–264. Pittsburgh: University of Pittsburgh Press.

Strevens, M. 2007. 'Review of Woodward, Making Things Happen.' *Philosophy and Phenomenological Research* **74** 233–249.

———-. 2008. 'Comments on Woodward, Making Things Happen.' *Philosophy and Phenomenological Research* **77** 171–192.

Woodward, J. 2003. *Making Things Happen*. Oxford: Oxford University Press.

———-. 2008a. 'Mental Causation and Neural Mechanisms.' In *Being Reduced: New Essays on Reductive Explanation and Special Science Causation*, J. Hohwy and J. Kallestrup, eds., 218–262. Oxford: Oxford University Press.

———-. 2008b. 'Response to Strevens.' *Philosophy and Phenomenological Research* **77** 193–212.