

*forthcoming* in Vassilios Karakostas and Dennis Dieks (eds.) (2013), “Recent Progress in Philosophy of Science: Perspectives and Foundational Problems”, *The Third European Philosophy of Science Association Proceedings*, Dordrecht: Springer.

## **Mechanistic explanation: asymmetry lost**

Samuel Schindler<sup>1</sup>

### Abstract

In a recent book and an article, Carl Craver construes the relations between different levels of a mechanism, which he also refers to as constitutive relations, by means of mutual manipulability (MM). Interpreted metaphysically, MM implies that inter-level relations are symmetrical. MM thus violates one of the main desiderata of scientific explanation, namely explanatory asymmetry. Parts of Craver’s writings suggest a metaphysical interpretation of MM, and Craver explicitly commits to constitutive relationships being symmetrical. The present paper draws attention to this shortcoming of Craver’s account. The paper furthermore explores the option of interpreting MM epistemologically, as a means for individuating mechanisms. It is argued that MM then is redundant. MM should therefore better be abandoned.

### **1 Introduction**

Ever since Machamer, Darden, and Craver (2000)’s landmark article “Thinking about mechanisms”, mechanistic explanations—thought to be the most pervasive kinds of explanation in the biological sciences—have become a major research topic in the philosophy of science. In a nutshell, Machamer et al. characterize mechanisms as being “composed of both entities (with their properties) and activities. Activities are the producers of change. Entities are the things that engage in activities” (3). To provide a mechanistic explanation of a phenomenon, then is “*to explain how it was produced*” by a mechanism (ibid.). The *production* of the phenomenon in question by a mechanism, call it MPP (*m*echanistic *p*roduction of the explanandum *p*henomenon), is thus absolutely central to the mechanistic conception of explanation. Although not explicitly

---

<sup>1</sup> Centre for Science Studies, Department of Physics and Astronomy, Aarhus University, Munkegade 120, Building 1520, 8000 Aarhus, Denmark, samuel.schindler@ivs.au.dk.

highlighted by Machamer et al., MPP thus ensures that the mechanistic account of explanation captures one of the most important desiderata on accounts of explanation: explanatory asymmetry. Mechanisms explain phenomena, but phenomena do not explain mechanism, because mechanisms produce phenomena and not vice versa. The direction of explanation thus follows the direction of a mechanism's production of the relevant phenomenon. This assumption is in fact analogous to an assumption made by large parts of the philosophical literature on causation, perhaps most explicitly put by Salmon (1998, 129): "The asymmetry of explanation is inherited from the asymmetry of causation" (see also Strevens 2008b, 24f. and 76-7). And indeed, although mostly concerned with the descriptive project of drawing to the attention of philosophers the importance of mechanistic explanations, Machamer et al. do express broad and general sympathy with a causal process theory for MPP in the tradition of Salmon's (1984) early work on causation. Process theories of causation, however, have widely been acknowledged to fail on various counts (Hitchcock 1995).

In his recent book (Craver 2007) and an article (Craver and Bechtel 2006), Craver offers important refinements of the original mechanistic account by Machamer et al. Amongst other things, Craver proposes to understand the relation between the mechanism and the explanandum phenomenon (i.e., MPP relation) in terms of "mutual manipulability", which, by a reviewer, has been judged "one of the main achievements of the book" (Levy 2009, 141). It will be the purpose of this paper, to assess this aspect of Craver's account.

This is how I proceed. In Section 2 I introduce Craver's notion of mutual manipulability (MM) as an explication of MPP. I argue that Craver's explication of MPP strips the mechanistic account of explanation of its ability to capture explanatory asymmetry. In Section 3 I explore ways in which this undesirable consequence might be avoided. One option I highlight is the interpretation of MM as a purely epistemological criterion for identifying MPP's. As I argue in Section 4, however, there is clear textual evidence that Craver intends MM as an explication of the *meaning* of MPP, which I argue is a genuinely metaphysical project. Regardless, I show in Section 5 that, if interpreted in pure epistemological terms, MM becomes redundant. In Section 6 I conclude this

paper by recommending the abandonment of MM and by pointing to one characteristic of mechanisms that the proponents of the mechanistic approach might want to focus on in order to justify the need for special accounts of mechanistic explanations.

## 2 Explanatory asymmetry lost?

Just like Machamer et al. (2000), Craver (2007, 6-7) defines mechanisms as “entities and activities organized such that they exhibit the *explanandum* phenomenon”. Craver’s (and Machamer et al.’s) standard example for a mechanistic explanation is the explanation of the neuronal action potential, which “is explained by reference to component parts of the action potential mechanism”, whereby examples for component entities are ions, ion channels, protein chains, etc. and examples for component activities are diffusion processes, and changes in conformation (121-2).

Craver distinguishes between a lower and an upper level in mechanisms (6-7). At the lower level he locates the entities  $X$  and their properties or activities  $\phi$ , and the ‘upper’ level is constituted by the phenomenon to be explained. The ‘mechanism as a whole’, i.e.,  $X$ ,  $\phi$ , and the explanandum phenomenon, Craver denotes as  $S$ . Furthermore he treats the phenomenon to be explained as being equivalent to  $S$ ’s activity  $\psi$ . MPP then, in Craver’s terminology, is  $S$ ’s  $\psi$ -ing (i.e. the explanandum phenomenon) being “exhibited” or “produced” (Craver uses both terms) by the activities of the mechanism’s components ( $X$ ’s  $\phi$ -ing). Furthermore, even though of minor importance in the following, for Craver (as for Machamer et al.) mechanisms often consist of multiple levels. That is, the upper level of one mechanism may be a component of a lower level of another mechanism, and so on. More importantly, Craver sharply distinguishes between intra-level and inter-level relations (Craver and Bechtel 2006; Craver 2007). Whereas intra-level relations are causal relations, inter-level relations are not; they are so-called *constitutive* relations. Constitutive relations—in contrast to causal relations—are symmetric, synchronous, and part-whole relations (153-4). Although not made very explicit by Craver, constitutive relationships are meant to specify MPP, as we shall see in the following.

In order to elucidate inter-level relationships in mechanisms, Craver (2007), following Woodward (2003), adopts the notion of an ideal intervention: “an ideal intervention  $I$  on  $\phi$  with respect to  $\psi$  is a change in the value of  $\psi$  that changes  $\psi$ , if at all, only via the change in  $\phi$ ” (154). Interventions need not be performable by humans, nor need they be physically possible. All that is required is that they be logically possible (see Woodward 2003, 127ff.). Craver (2007)’s explication of inter-level relations in terms of ideal interventions consists of two parts, which together form his *mutual manipulability* criterion (MM):

(CR1): When  $\phi$  is set to the value  $\phi_1$  in an ideal intervention, then  $\psi$  takes on the value of  $f(\phi_1)$ . (155)

(CR2): When  $\psi$  is set to the value  $\psi_1$  in an ideal intervention, then  $\phi$  takes on the value of  $f(\psi_1)$ . (159)

Apparently, both CR1 and CR2 have the structure of Woodwardian active counterfactuals, i.e., counterfactuals whose antecedents are “made true by interventions” (Woodward 2000, 199), which Woodward intends to pick out *causal* relationships. And yet, Craver denies that neither CR1 nor CR2 do so. As mentioned above, the combination of CR1 and CR2 (i.e. MM) is supposed to individuate *constitutive* relations, which, according to Craver, are not causal relations. In accordance with the convention in the contemporary literature on causation to refer to a causal relation between  $X$  and  $Y$  as  $X$  “being causally relevant” to  $Y$  (cf. Woodward 2003, 39), Craver also refers to CR1 and CR2 as criteria for constitutive *relevance*. More specifically, “one can change the *explanandum phenomenon* by intervening to change a component [of a mechanism]”, and vice versa, “one can manipulate the component by intervening to change the *explanandum phenomenon*” (153). Craver concludes that

many, if not most, causal relationships are unidirectional. In contrast, all constitutive dependency relationships are bidirectional (153).

And since inter-level relationships are symmetrical relationships, they are therefore “only uncomfortably viewed as causal” (p. 153). Another reason that Craver mentions for constitutive relationships being distinct from causal relationships, which I will not be concerned with in much detail here, is that in the former, “ $\phi$ ’s taking on a particular value is not temporally prior to  $\psi$ ’s taking

on its value” (151-2), in other words, constitutive relationships are ‘synchronic’, whereas causal relationships are not. However, contrary to what Craver seems to suggest, a relationship being symmetrical does not imply that it cannot be causal. In fact, the account of causation Craver explicitly allows for ‘cyclic’ causal relations where a change in the cause variable brings about a change in the effect variable and vice versa (Woodward 2003, 396). Furthermore, as Leuridan (2012, fn. 27 and 29) points out, many relations of interest in neurobiology (Craver’s subject) are causal feedback loops, i.e., symmetrical causal relationships.

Craver’s explicit commitment to inter-level relationships being symmetrical relationships raises the following concern: if inter-level relationships really are symmetrical, what is it in Craver’s account that ensures that the desideratum of explanatory asymmetry is respected? Recall, on the original mechanistic account by Machamer et al., explanatory asymmetry is respected, because mechanisms *produce* phenomena, but not vice versa. The direction of explanation simply follows the direction of production. So what happened to the production relationship in Craver’s account, which I referred to as MPP, and which Craver, like Machamer et al., characterizes as a central feature of mechanisms? At one point in his book, Craver *seems* to say that MPP is to be spelled out in terms of Woodwardian counterfactuals picking out *causal* relations:

*to say that one stage of a mechanism is productive of another (as I suggest in Machamer et al. 2000; Craver and Darden 2001) is to say, at least in part, that one has the ability to manipulate one item by intervening to change another. (93-4; added emphasis)*

As mentioned above, however, for Craver inter-level relations are *not* causal relations. So if the above quotation *were* to refer to MPP (i.e. an inter-level relation) Craver would clearly contradict himself. So despite speaking of “production” in this one passage, suggesting reference to MPP, Craver is better read as spelling out *intra*-level causal relations here. This still leaves us with the question of how MPP is to be understood.

In personal communication,<sup>2</sup> Craver is ready to give up on a 'literal' reading of MPP; he suggests MPP be interpreted metaphorically instead. That is, whenever we say that the explanandum phenomenon is "produced" by the mechanism, what we should say more carefully is that the phenomenon is constituted or "made up" by the mechanism, very much in accordance with a constitutive understanding of inter-level relations. This is in line with the one taken in Craver and Bechtel (2006):

The causal claims, when made explicit, are all intra-level. But we continue to talk about bottom up causal relation [from one to another level] when we are being quick or informal as long as we understand that the change at the higher level is mediated by, or explicable in terms of, a mechanism [and its constitutive relations]. (ibid., 557)

In this context it is furthermore interesting to note that the idea of MPP, i.e., the idea of the mechanism's producing the explanandum phenomenon, is altogether absent from Craver and Bechtel's joint paper. But again, construing MPP as a symmetrical relation (namely as CR) subjects the mechanistic account to the problem of explanatory symmetry. If the relation between the mechanism and the phenomenon is symmetrical, what is it that prevents us from saying that the phenomenon also explains the mechanism? In response to this question, Craver (personal communication) is ready to embrace a deflationary "explanatory pluralism", implying that phenomena might as well explain mechanisms. But perhaps there are more appealing options for Craver.

### **3 Explanatory asymmetry saved?**

There is a much simpler response to the problem of explanatory symmetry available to Craver. He could point out that MM is only part of how constitutive relationships are to be understood. Another important aspect of constitutive relationships is that they are *part-whole* relationships (see above and Craver 2008, 153-4). Part-whole relationships are asymmetrical relationships: if  $\phi$  is a part of  $\psi$  then  $\psi$  cannot be part of  $\phi$ . Explanatory asymmetry between  $\phi$  and  $\psi$

---

<sup>2</sup> Email communication with the author of this paper on 02-22-2012.

would thus be secured.<sup>3</sup> The problem with this response, however, is that it stands in outright contradiction with Craver's assertion that constitutive relationships are symmetrical: either constitutive relations are part-whole relations or they are symmetrical. Craver cannot have it both ways. Although the former option seems much more plausible in the face of the problem of explanatory asymmetry, part-whole relations are underdeveloped in Craver's account. In fact Leuridan (2012) argues that under the perhaps most intuitive definition of part-hood, cases of mutual causation cannot be ruled out by CR. This is contrary to Craver's insistence that constitutive relations are different in kind from causal relations. What is more, Craver puts a lot of stress on the symmetry property of constitutive relations and characterizes it as one of the crucial differences between constitutive and causal relations (Craver and Bechtel 2006; Craver 2007). At any rate, there appear to be only two options for Craver: he either gives up on the idea that constitutive relations are symmetrical or he tries to save explanatory asymmetry whilst holding onto constitutive relationships being symmetrical (and giving up on part-hood). The former option appears much more plausible. It would allow Craver to embrace the asymmetry property of part-whole relations and thereby explanatory asymmetry. But because Craver is so adamant about constitutive relations being symmetrical, let us briefly consider the latter option, before exploring the former option.

Take one of Woodward's preferred examples, the ideal gas law. This law relates variables of pressure (P), the volume of a gas (V) and temperature (T) in the formula  $PV=RT$  (R is the gas constant). Clearly this is a symmetrical relationship. We can intervene on P to change T, and conversely, we can intervene on T to change P (by holding fixed V, respectively). Now, assuming (with Woodward) that the ideal gas law is an explanatory generalization, in each of the above scenario, explanatory asymmetry is preserved despite the relationship being a symmetric relationship. In the one context T is the cause variable and P the effect variable, and in another context, P is the cause variable and T the effect variable. In the first scenario a change in T explains the change in

---

<sup>3</sup> Craver did not make this reply in the abovementioned email correspondence. I thank an anonymous referee for making me consider this option.

P (but not vice versa), and in the second scenario P explains T (but not vice versa).<sup>4</sup> There is of course no *a priori* reason why this insight could not be extrapolated to the context of mechanistic explanations. However this extrapolation is only of a limited sort. It extends only to intra-level but not to inter-level relationships in mechanisms. That is, on one particular level of a mechanism it might make sense to say, as in the above example, that a change in  $X_1$ 's  $\phi_1$ -ing explains (in a minimal sense) a change in  $X_2$ 's  $\phi_2$ -ing, and vice versa (just in case, of course,  $X_1$  and  $X_2$  are related as P and T above). Further, it also makes sense (now between different levels of mechanisms) to say that if we can intervene on  $\phi$ 's to change  $\psi$  (the explanandum phenomenon), then  $\phi$ 's explain  $\psi$ . However, crucially, it makes little or no sense to say the reverse, namely that  $\psi$  explains  $\phi$ , *even if CR2 were satisfied*. After all, the explanandum phenomenon cannot explain the mechanism, at least not in the standard sense of the term. It therefore seems that there is no straightforward way in which Craver could stick to his symmetry thesis and save explanatory asymmetry.

Again, Craver could simply give up on the symmetry thesis. Indeed, MM implies the symmetry of inter-level relations *only if* it is to be understood as a *metaphysical* explication of inter-level relations, i.e., as specifying the meaning of inter-level relations. If read in epistemological terms, that is, as a criterion for *identifying* inter-level relations (in contrast to, say, mere correlations between a mechanism and some phenomenon), MM has no implications for the directionality of inter-level relations. At least *a priori*, there is no contradiction between inter-level relations being asymmetrical and our means for identifying them being applicable in both directions of this relation (i.e., bottom-up *and* top-down). An epistemological interpretation of MM would also allow Craver to reconcile MM with the asymmetry of part-hood. Unfortunately for Craver, however, there is strong textual evidence that Craver aims for a metaphysical explication of constitutive relevance in terms of MM. So before we can consider a pure epistemological interpretation of MM, we need to consider that textual evidence in more detail.

---

<sup>4</sup> This was suggested to me by Bert Leuridan.

## 4 The metaphysics of mechanisms

First and foremost, as pointed out above, part of Craver's motivation to amend the original mechanistic account by Machamer et al. is clearly owed to the wish to explicate the *meaning* of MPP. Such endeavor is generally regarded as a genuinely metaphysical. Second, Craver's concession that MM implies the symmetry of inter-level relations clearly presupposes that MM is interpreted metaphysically. Again, if MM were a mere epistemological criterion, nothing would follow for the direction of inter-level relations. Furthermore, Craver makes clear that he wishes to provide a 'normative' account that can "demarcate [mechanistic] explanation from other kinds of scientific achievements", and that can "reveal criteria for *assessing* explanations", in other words it "should prescribe norms of explanation" (20; original emphasis). All this he could not do if he were to interpret MM epistemologically rather than metaphysically. One e.g. cannot assess whether an explanation is a good explanation without having provided at least a partial answer to the question of what constitutes an explanation, i.e., a question about the meaning of explanation.

But again, there are indeed also a number of passages in Craver's book that suggest that Craver views MM as an epistemological criterion. Primarily, this is suggested by the context in which Craver explicates inter-level relations. This context is formed by Craver's pointing to the various inter-level experimental strategies that can be used, inter-level bottom-up and top-down, to *establish* certain entities and activities as being part of a certain mechanism. Verbatim, Craver says:

I build my positive account [of mechanisms] by considering the *experimental* strategies that neuroscientists *use to test* whether a given entity, activity, property, or organizational feature is relevant to the behavior of the mechanism as a whole [...] (140; added emphasis).<sup>5</sup>

Furthermore, Craver appears to think that a non-reductive analysis of causation (or constitution) based on active counterfactuals implies that the analysis is non-metaphysical. Explicitly, he says in his summary of Woodward's manipulationist

---

<sup>5</sup> Couch (2011, fn. 6) also reports that Craver explicitly embraces an epistemological interpretation of MM in personal communication.

account, which he then goes on to use to define constitutive relevance in mechanisms, that

I do not discuss here whether such metaphysics [of causation] is required or what the available metaphysical options are. Even if the manipulationist view does not identify the truth-maker[s] for causal claims, it is nonetheless an illuminating analysis of the causal truths themselves [...] (Craver 2007, 105-7).

This, however, is a misapprehension. The fact that a philosophical analysis is non-reductive does not imply that it is non-metaphysical. To see this requires a brief excursion into Woodward's account.

Many philosophical analyses seek to reduce the concept of causation to another concept. Humeans, for instance, reduce causation to mere empirical regularities. David Lewis reduces causation to counterfactual dependence. Woodward's analysis of causation, in contrast, is decidedly non-reductive. Woodward defines causal relationships as generalizations that remain invariant under interventions. As Woodward acknowledges, the notion of an intervention is itself a causal notion. In a sense, Woodward's account is thus circular. However it is not *viciously* circular, as Woodward (2003, 20ff.) points out, because the causal relation that is being appealed to (*I* causing a change in *X*) is different from the causal relation that the analysis seeks to illuminate (namely, *X* causes *Y*). Woodward goes on to show that such a non-reductive account can very well be insightful. It for instance delivers markedly different verdicts on a number of important issues (e.g. action-at-a-distance, causation by prevention or absence) than Salmon (1984)'s classical causal process theory of causation does.

The crucial question now is of course: is a non-reductive analysis of causation non-metaphysical (as Craver would have it)? This is not an uncontroversial matter. In a slightly heated exchange with Strevens (2007, 2008a), Woodward (2008) explicitly denies that his own account is metaphysical. However Strevens (2008) offers a number of convincing reasons why Woodward might in fact be mistaken about the aims of his own book. First, Woodward (2003), throughout his book, presents his analysis as a superior rival to Lewis's metaphysical analysis of causation. It would be hard to see why Woodward does that if the aim of his project were entirely different from Lewis's. Second, Woodward seeks to provide an account according to which

causation is mind-independent, in other words, an account of the *nature* of causation which is independent of how we get to know about this relation (2003, 118ff.). This clearly is a metaphysical endeavor. Third, Woodward states that “my aim is to provide an account of the meaning or content of various locutions, such as X causes Y” and that “my project is semantic or interpretive” (38). As Strevens (2008) points out

In modern times, such a project is invariably interpreted as aiming to provide truth conditions for the sentences or thoughts in question, and therefore as aiming to specify those representations’ truthmakers. It may look like semantics, but it is also a kind of metaphysics [...] it is generally agreed that a word with an explicit definition has as its extension whatever stuff satisfies that definition. If Woodward’s causal semantics is a truth-conditional semantics, he is inevitably, unavoidably, ineluctably committed to producing an account of the truthmakers for causal talk, a metaphysics of causal facts, whatever his protestations. (184)

Back to Craver. If Craver wishes to spell out the *meaning* of the inter-level relation in terms of constitution then, by Strevens’s lights, he inextricably commits himself to a metaphysical project. Since there are clear signs that Craver does wish to elucidate the meaning of inter-level relations (see above),<sup>6</sup> his project, contrary to what he says, does commit him to a metaphysics of constitutive relations. And since Craver’s spelling out of inter-level relations implies a symmetrical inter-level relation, to which he explicitly commits, the mechanistic account no longer captures explanatory asymmetry. So let us now finally consider the possibility of interpreting MM, contrary to much of what (is implied by) what Craver says about it, as a merely epistemological tool for individuating constitutive relationships.

## **5 Individuating mechanism boundaries**

The main epistemological function of MM I alluded to briefly above is this: it concerns the delineation of the *boundaries* of mechanisms (Craver 2007, 141ff). In other words MM is supposed to specify (i) which entities and activities are, and which ones are not, part of the lower level of a mechanism with respect to a particular explanandum phenomenon (this function is performed by CR2), and,

---

<sup>6</sup> This is also what Craver told me in personal communication.

conversely, (ii) which phenomenon is picked out by a particular mechanism (this function is performed by CR1). Craver gives the following example (152).

The cognitive capacity of word-stem completion, in which a subject is presented with a list of words and afterwards asked to complete the word stems of the words presented previously, is affected by changing the heart rate of the subject. That is, if one were to change the heart rate of the subject (by e.g. torturing the subject), the subject's capacity to complete word stems would invariably change. According to CR1 *alone*, however, the heart rate would *erroneously* be deemed as constitutively relevant for the phenomenon of word-stem completion; the heart rate normally would not be considered a part of the mechanism of word stem completion, or so Craver reasons. Here is where CR2 comes in. Engaging subjects in word-stem completion will *not*, under normal circumstances, result in a different heart rate. Hence CR2 is not satisfied by the example and the heart rate is therefore not to be deemed part of the mechanism of word stem completion.

It is questionable whether MM really fulfills the purpose Craver foresees for it. To see this, note that the persuasiveness of the above example rests on the implicit assumption that the relevant mechanism for the capacity of word stem completion is a *cognitive* mechanism. Only then the heart rate appears irrelevant. But not in all contexts need this be so. In contexts in which one is interested in, for instance, the capacity of completing word stems as compared to the absence of any cognitive capacity, the heart rate appears to be indeed a part of the relevant mechanism. In other words, whether or not a mechanism (or part thereof) is relevant to the explanation of a phenomenon is subject to pragmatic considerations. In fact Craver is very well aware of this. He for instance highlights the importance of contrast classes in the specification of the explanandum phenomenon (2007, 202ff.). And

Once the explanandum phenomenon has been specified [...] it is then an *objective matter* whether or not it is possible to make a difference to that effect by manipulating certain antecedent variables and which manipulations of that variable in fact make a difference. My account is thus no more pragmatic than any view of causation or explanation would have to be. (204)

In a more recent publication, Craver (2009) states even more explicitly that

[...] the spatial and causal boundaries of mechanisms depend on the epistemologically prior delineation of relevance boundaries. But relevance to what? The answer is: relevance to the phenomena that we seek to predict, explain, and control. Within the boundaries of a mechanism are *all and only* the entities, activities, and organizational features relevant to the phenomenon selected as our explanatory, predictive, or instrumental focus. (591; added emphasis)

The epistemological function CR2 is supposed to perform, namely the picking out of a particular mechanism, *given a particular explanandum phenomenon*, is therefore not necessary. It is plausibly carried out by pragmatic considerations: we choose a particular phenomenon we want to explain, predict, etc. and then we ask, by reverse engineering, as it were, what causes are responsible for the phenomenon in question. MM as an epistemological criterion in Craver's account looks therefore redundant.

## 6 Conclusion

When interpreted metaphysically, MM, in violation of a central desideratum of explanation, implies explanatory symmetry and is inconsistent with the part-hood characterization of constitutive relevance (as I argued in Section 2 and Section 3). On the other hand, when interpreted epistemologically, MM does not give us any extra purchase on the individuation of mechanisms (as argued in the last section). I therefore believe that MM is better to be abandoned altogether. I thus disagree with Leuridan (2012) who concludes his detailed discussion of MM by suggesting that mechanistic inter-level relations be interpreted, contrary to Craver, as relations of *mutual causation*. But such a proposal is of course just as much subject to my criticism of explanatory asymmetry being lost (see also Section 3). Rather I think that inter-level relations, i.e. the productive relation MPP, are perhaps best understood in terms of *unidirectional* Woodwardian counterfactuals picking out *causal* relations, without invoking top-down counterfactuals (contra Craver). But of course, there would then be no need for a specific mechanistic account of explanation (see Woodward 2002). So is there? A crucial feature of mechanisms is the *organization* of entities and their spatio-temporally *concerted* interaction that produces a phenomenon. In order to accommodate this feature of mechanisms, Woodward (2011) suggests that counterfactuals describing causal relations in mechanisms possess

„characteristic spatio-temporal signatures“ (427). I take it that this suggestion translates into mechanisms being specifiable in terms of conjunctions of active counterfactuals with complex antecedents of the following form: <If an appropriate intervention *I* had changed the value of an “entity variable” *X* at time *t*<sub>1</sub> and, if an appropriate intervention *I* had changed the value of an “entity variable” *Y* at time *t*<sub>2</sub>, and ... etc., then the explanandum phenomenon would have been produced by the mechanism comprising entities *X*, *Y*, etc.>. Since it seems awkward to say that variables have a “location”, spatial information might be a little harder to accommodate than this. At any rate, barring concerns about the modularity assumption that Woodward makes (Cartwright 2002), such an amendment of Woodwardian counterfactuals to accommodate the genuine mechanistic feature of organization and concerted interaction of mechanism-components seems to be unproblematic in principle.

### **Acknowledgements**

This article has profited significantly from discussions with Bert Leuridan, Kareem Khalifa, and Petri Ylikoski. Marcel Weber, and Alexander Reutlinger. An anonymous referee of this journal provided helpful comments. Carl Craver kindly clarified his views for me in personal communication. I’m grateful to all of these individuals. I furthermore want to thank the audience at the EPSA 2011 conference in Athens (Greece) for feedback.

## **7 References**

- Cartwright N (2002) Against modularity, the causal Markov condition, and any link between the two: Comments on Hausman and Woodward. *The British Journal for the Philosophy of Science* 53 (3):411-453.
- Couch MB (2011) Mechanisms and constitutive relevance. *Synthese* 183 (3):375-388.
- Craver CF (2007) *Explaining the brain : mechanisms and the mosaic unity of neuroscience*. Oxford: Oxford University Press.
- — — (2009) Mechanisms and natural kinds. *Philosophical Psychology* 22 (5):575-594.
- Craver CF, W Bechtel (2006) Top-down Causation Without Top-down Causes. *Biology & Philosophy* 22 (4):547-563.
- Hitchcock CR (1995) Salmon on explanatory relevance. *Philosophy of Science* 62 (2):304-320.
- Leuridan B (2012) Three problems for the mutual manipulability account of constitutive relevance in mechanisms. *The British Journal for the Philosophy of Science* 63 (2):399-427.

- Levy A (2009) Carl F. Craver, Explaining what? Review of explaining the brain: mechanisms and the mosaic unity of neuroscience. *Biology and Philosophy* 24 (1):137-145.
- Machamer P, L Darden, CF Craver (2000) Thinking about mechanisms. *Philosophy of Science* 67 (1):1-25.
- Salmon W (1984) *Scientific explanation and the causal structure of the world*. Princeton, N.J.: Princeton University Press.
- — — (1998) *Causality and explanation*. New York: Oxford University Press.
- Strevens M (2007) Review of Woodward, Making Things Happen. *Philosophy and Phenomenological Research* 74 (1):233-249.
- — — (2008a) Comments on Woodward, Making Things Happen. *Philosophy and Phenomenological Research* 77 (1):171-192.
- — — (2008b) *Depth : an account of scientific explanation*. Cambridge, Mass.: Harvard University Press.
- Woodward J (2003) *Making things happen: A theory of causal explanation*: Oxford University Press, USA.
- — — (2008) Response to Strevens. *Philosophy and Phenomenological Research* 77 (1):193-212.
- — — (2011) Mechanisms revisited. *Synthese*:1-19.
- Woodward J. (2002) What is a mechanism? A counterfactual account. *Philosophy of Science* 69 (3):366-377.