

Accepted for publication in *History and Philosophy of the Life Sciences*

Making Sense of Downward Causation in Manipulationism. With Illustrations from Cancer Research

Christophe Malaterre

*Institut d'Histoire et de Philosophie
des Sciences et Techniques
Paris, France*

ABSTRACT – Many researchers consider cancer to have molecular causes, namely mutated genes that result in abnormal cell proliferation (e.g. Weinberg 1998); yet for others, the causes of cancer are to be found not at the molecular level but at the tissue level and carcinogenesis would consist in a disrupted tissue organization with downward causation effects on cells and cellular components (e.g. Sonnenschein & Soto 2008). In this contribution, I ponder how to make sense of such downward causation claims. Adopting a manipulationist account of causation (Woodward 2003), I propose a formal definition of downward causation, and discuss further requirements (in light of Baumgartner 2009). I then show that such an account cannot be mobilized in support of non-reductive physicalism (contrary to Raatikainen 2010). However, I also argue that such downward causation claims might point at particularly interesting dynamic properties of causal relationships that might prove salient in characterizing causal relationships (following Woodward 2010).

KEYWORDS – Downward causation, manipulationism, levels of causation, causal granularity, cancer research

1. Introduction

Garden-variety examples of causation often include smoking as the alleged cause of lung cancer. For contemporary health science, the search for the causes of cancer goes much deeper into minute biomolecular entities. Because of its complexity, the causal mapping of carcinogenesis is sometimes compared to integrated electronic circuits and to the identification of an intricate molecular circuitry (Hanahan and Weinberg 2000). Over the past decades, the dominant paradigm has been to look for

mutated genes, be they oncogenes or tumor suppressor genes, as the cause of abnormal cell proliferation and cancer tumors (e.g. Weinberg 1998). More recently, some alternative research programs have proposed to look for causes of cancer not at the molecular level of genes but at the level of tissues or groups of cells: the cause of cancer would not be a faulty gene but a disrupted tissue organization; as a result, carcinogenesis would be best understood as a tissue-level phenomenon with downward causation effects onto cells, for instance inducing enhanced cellular reproduction, or even onto genes, inducing genetic mutations at particular loci (e.g. Sonnenschein & Soto 2008). In this paper, my aim is not to argue which “theory” of cancer is right or wrong, but rather to focus on the philosophical notion of downward causation. I propose to analyze downward causation claims in light of a manipulationist account of causation (Woodward 2003). I argue that such claims need to be asserted with caution and that their metaphysical consequences are weaker than asserted. Nevertheless, I also argue that these claims might point to particularly interesting characteristics of causal relationships that have to do with their dynamics, thereby broadening the set of dimensions along which to characterize causation (as in Woodward 2010). The paper is organized as follows. In section 2, I briefly present the scientific context of cancer research and the related downward causation claims. In section 3, adopting Woodward’s manipulationist account of causation, I propose a formal definition of downward causation. I discuss further formal requirements in section 4. In section 5, I argue that downward causation so-construed cannot be used in support of non-reductive physicalism. In section 6, I discuss how downward causation might nevertheless prove useful by pointing at specific dynamic characteristics of causal relationships, and I illustrate this view in section 7.

2. Organicism and Downward Causation in Cancer Research

In the past fifty years, the field of cancer research has produced hundreds of thousands of publications (Downward 2006) and seen a profusion of schools, visions, or paradigms compete in the search for the ultimate cause (and cure) of cancer. Whereas the early vision of cancer as a dysregulation of genetic expression was replaced in the 1980s by that of cancer as resulting from the presence of faulty genes (oncogenes and tumor suppressor genes), the complexity of carcinogenesis has become even more apparent in the past decade, resulting in a profusion of strikingly different approaches that are also typical for biology more generally (Morange 2007). Old ideas concerning the role of autophagy and of senescence are being revived, as are those about the place of genomic instability and of metabolic alteration in carcinogenesis; epigenetic theories are being proposed, that

draw, for instance, on DNA methylation and chromatin modification; evolutionary models of tumor growth, including niche construction strategies, are being developed, and the role of stromal cells, which was once thought to be peripheral compared to that of parenchymal cells, is increasingly being recognized.¹

Within this last stream of research, some propose that cancer may result from a disruption of tissue organization: carcinogens would cause malignant tumors by disrupting the normal interactions between neighboring stromal and parenchymal cells (Soto and Sonnenschein 2006). Under normal conditions, cells maintain physical contact and interact with neighboring cells through a combination of junctions: for instance, adherens junctions enable the reciprocal anchoring of cytoskeletons and play a role in inducing cell structural polarity; gap junctions provide a communication mechanism through which small molecules, including signaling molecules, can pass from one cell to another; tight junctions seal the space between cells and prevent the diffusion of solutes through the intercellular space. In addition, cells also interact with the extracellular matrix, which is a complex three-dimensional network of macromolecules that serves as an architectural scaffold to cells while also providing them with contextual cues. The disruption of these interactions is believed to contribute to carcinogenesis by creating a context that promotes tumor growth and protects it from immune attack (Bissell and Radisky 2001). For some, this disruption is more than a contributing factor and ought to be considered as the real cause of cancer: carcinogenesis would not consist in a faulty gene that causes cells to proliferate, but rather in a disruption of the normal interactions and patterns of reciprocal chemical regulations between cells resulting in cells no longer being able to “perceive” their functional positioning and reverting to a default mode of active proliferation (Soto and Sonnenschein 2005). Such alterations in tissue organization would, in turn, through a complex causal chain, induce aneuploidy and mutations.

Of course, since cancer is still one of the biggest challenges of the biomedical sciences, such a vision of carcinogenesis is still in itself a research program and not yet an accomplished theory. One might therefore interpret this organicist stance as a particular heuristic strategy, and its divergence from the more mainstream genetic and molecular vision of cancer as something that will fade away as the different approaches to carcinogenesis are progressively integrated (Malaterre 2007). Yet, some of their proponents claim that the two approaches are truly incompatible or even incommensurable, in so far as they rely on radically different causal sequences (Sonnenschein and Soto 2008, 375) : the cause of cancerous

¹Stromal cells constitute the support tissue of an organ, whereas parenchymal cells are responsible for its function. This distinction, however, blurs numerous reciprocal interactions (e.g. Soto and Sonnenschein 2006).

tumors (associated with a change in tissue organization) would not be DNA mutations but disruptions in cell-to-cell interactions (resulting subsequently in malignant DNA mutations)². The intricacy of such cell-to-cell interactions makes it practically impossible to sort out cause and effect into neat causal chains at the cellular level and leads to construing carcinogenesis as a tissue-level phenomenon. Tissues, it is argued, are the locus of reciprocal causality between cells and of downward causation phenomena³. Carcinogenesis is even interpreted as providing an experimental argument against the causal closure of the physical world (Soto et al. 2008). It seems however that such claims are in need of philosophical explication.

3. Levels and Manipulationism

The very idea of downward causation builds on three presuppositions: (a) that levels can somehow be defined in nature, or, in milder terms, be attributed to the objects of our theories, such levels being orderly arranged so that, for any pair of levels, one might define an upper and a lower level, (b) that causal relata, or variables in manipulationist terms, can be ascribed to these levels, so that one might define upper- and lower-level causal relata, and (c) that some causal relationships run from upper-level causal relata down to lower-level causal relata. Let us examine each of these three presuppositions.

a. Defining Levels

The idea of downward causation goes hand-in-hand with a view that nature is structured in levels of organization, from the most fundamental entities of particle physics up to the most gigantic ones of astrophysics. After all, aren't organisms composed of cells, cells of molecules and molecules of atoms? Nature thus construed appears as fully structured and organized, hosting objects that form neatly nested hierarchies of parts and wholes, each belonging to given levels of organization. Yet, things are not so simple, or so I will argue, for at least two reasons⁴.

First, it is not at all clear that a given level can be ascribed once and for all to any given entity. Rather, level ascription depends on the way one

² One may also consider the extent to which the adoption of a particular “theory” of carcinogenesis might impact the classification of cancer as a disease – see Kutschenko, this volume, on the classification of diseases.

³ See also Bertolaso, this volume, on the organicist stance in cancer research.

⁴In an essay on reduction, Hempel (1969) made comments in this same direction, and more specifically on the difficulty of characterizing entities as being “physical”, “chemical”, “biological” and so forth. For him, such labels are relative to the conceptual apparatus and vocabulary distinctive of each *discipline*. In the present essay, I show that a similar difficulty arises when it comes to assigning specific *levels* to entities. I thank Werner Callebaut for kindly pointing this reference to me.

chooses to decompose wholes into parts. Consider, in physics, the case of the standard model of the atom: an atom is described as being constituted by a nucleus surrounded by a cloud of electrons; the level below that of the atom therefore seems to be the level nucleus-electron. But if we decompose further, we find out that the nucleus is itself composed of nucleons (protons and neutrons) that are themselves composed of quarks and leptons. Leptons, therefore, are found two levels down below the level nucleus-electron. Still, according to physics, electrons too are types of leptons. Hence the hybrid composition of the nucleus-electron level that includes entities two-levels up from leptons as well as leptons themselves, depending on how one decomposes an atom and ascribes levels to its parts. The consequence is crucial for downward causation⁵: indeed, should a case of causation between a nucleus and an electron be considered a case of downward causation or, rather, a case of same-level causation? Similar situations are also frequently found in biology when, for instance, ribosomes are sometimes considered as molecular entities and therefore placed at the same level as other complex organic molecules (e.g. DNA, RNA, or even proteins), but sometimes also included among cellular organelles like mitochondria, cellular nucleus and endoplasmic reticulum. Ascribing levels to entities is therefore not as straightforward as one might initially expect but depends on the decompositional approach one takes, that is to say on how one decomposes wholes into parts or complex phenomena into simpler ones.

Second, there also exist entities that do not clearly belong to any particular level of decomposition. Consider an electromagnetic field. Such a field can play a causal role at different levels: it can accelerate an elementary particle, be it a lepton (e.g. electron) or a nucleon (e.g. proton); but it can also deviate the needle of a compass or generate huge aurorae. In fact, numerous variables that notably relate to the environment within which particular systems are studied appear to play significant causal roles at a whole range of levels. Take the atmospheric pressure: it can play a causal role in a chemical reaction in so far as intervening on it changes the output of the reaction depending on the vapor pressure of the reactants; yet the atmospheric pressure also plays a causal role in a barometer by displacing a column of mercury, not to mention changes in atmospheric pressure that move clouds across oceans. Similarly, one may consider that geometric constraints also constitute causal factors that play a role at very different levels of organization: for instance, changing the volume of the optical cavity of a laser, in particular the distance of the reflecting walls on which the photons bounce in resonance, changes the number of photons that are emitted; yet also, changing the volume of a gas while keeping the pressure

⁵The focus here is on “downward” (as it appears in “downward causation”). Yet the same analysis shows that “upward” is as problematic as “downward”, making cases of “upward causation” not as straightforward as one would usually treat them.

constant changes its temperature. More generally, one therefore needs to take into account the existence of causal entities or variables that can act at different levels of organization without being tied to any level in particular. I propose to refer to such variables as “level-neutral variables”⁶. Because level-neutral variables are not particularly tied to any given level, it does not make much sense to qualify a causal relation in which they would have a stake as “downward causation”. For level-neutral variables, levels of causation simply are irrelevant.

Nature thus appears much less neatly organized into levels as one may think initially. Things are much more messy, with entities that might be assigned to one level or another, and with others that might interact at multiple levels. The consequence is that downward causation appears ill-defined in numerous cases. One way to go around this problem is to assert that what is characteristic of such causal relationships is not just they point “downward” (from entities at a higher-level down to entities at a lower-level), but that they are “mereologically downward” (from a whole down to its parts)⁷. Accordingly, in order to properly define downward causation one has to refer to entities that are clearly mapped onto a parts-whole set of relationships. Downward causation requires a “mereological context” that specifies a whole, a set of parts and a set of mereological relationships that describes how the parts compose the whole, while ascribing levels to the parts relative to each other and relative to the whole. It is only when this context is given that level membership ambiguities, as in the case of the electron or the ribosome, are removed. And because parts-whole relationships are required, level-neutral entities like an electromagnetic field or the atmospheric pressure are excluded.

b. Levels and Variables

Once levels have been defined, the second presupposition of downward causation is that one can ascribe causal relata -not just entities- to such levels. In manipulationism (as in Woodward 2003), causal relata are variables that can take up multiple values, and whose values can be construed as attributes of entities⁸. The main idea behind such a

⁶I choose this terminology in reference to C.D. Broad who, in his discussion of emergence, proposes to classify the properties of a given “order” or level as “ultimate characteristics” (i.e. emergent), “reducible characteristics” and “ordinally neutral characteristics” (1925, 79).

⁷This is simply to acknowledge that, somehow, “big things can interact with small things”. The key issue is rather to consider cases of causation between a whole and its parts, as also proposed by Kim and his “reflexive downward causation” (1999).

⁸The aim of this paper is to explore the notion of downward causation in light of manipulationism, and not to argue that manipulationism is a better account of causation than other accounts. Note however that manipulationism appears well suited to causation in the health sciences: it fits well with the idea that causes are to be regarded as levers upon which to act in order to bring about changes (as is the case with public health); it also suits

manipulationist account of causation is that, given a set of variables \mathbf{V} and given two variables X and Y belonging to \mathbf{V} , one can find out whether X causes Y simply by intervening on X and looking at the changes on Y , while keeping the other variables in \mathbf{V} at some fixed value. A much more precise definition of manipulationism can be found in (Woodward 2003); a cause is defined as being either a “direct” or a “contributing cause”, each being defined in reference to the notion of “intervention” through the conditions of “manipulationism” **M** (2003, 59) and of “intervention variable” **IV** (2003, 98). I will take these conditions as a starting point to add further requirements in order to define the narrower notion of downward causation. As seen above, the very notion of downward causation presupposes that one specifies the “mereological context” that includes the entities that the causal relata refer to. Since in manipulationism causal relata are causal variables, matching “levels” and “causation” requires to impose a joint condition on the causal variables and on the entities that have, as properties, specific values of such causal variables. Hence the following “mereological context” requirement:

- MC** All variables in \mathbf{V} are such that their values are attributes of entities for which a “mereological context” has been specified (i.e. a whole, a set of parts and a set of mereological relationships that describes how the parts compose the whole, while ascribing levels to the parts relative to each other and relative to the whole).

the experimental approach of the health sciences. Nevertheless, for a more general overview of diverse causal approaches relevant also to the health sciences, see for instance Russo (2009); see also Russo (this volume) for a defense of “epistemic causality” in biomedical contexts.

c. “Downward”

The condition **MC**, via the specification of a mereological context, enables one to attribute levels to the causal variables that are mentioned in the downward causation claim. Yet, according to the third presupposition behind downward causation, namely that causal relationships can run from upper-level causal relata down to lower-level causal relata, a further condition must be added in order to have the “downward” direction of such causal relations: the entities of which the values of the cause variables are attributes must be such that they belong, according to the mereological context, to a higher level than the entities of which the values of the effect variables are attributes. Hence the following “downward” condition: given two variables X and Y such that X is a cause of Y , and given **MC**, the causal relation between X and Y is downward iff:

- D** The values of X are attributes of an entity that is at a higher level than the entity of which the values of Y are attributes.

This downward condition is the one that sets the level-directionality from top to bottom of the causal relationships relatively to a mereological context that itself specifies the levels among which this level-directionality takes place⁹.

Given a manipulationist account of causation, a downward cause must therefore fulfill the following conditions: **M** (manipulationism), **IV** (intervention variable), **MC** (mereological context) and **D** (downward). These conditions are, however, not sufficient: in particular, they may lead to downward causation claims that are ill-defined, depending on how one interprets **M**. It is to this problem that I now turn.

4. Further Requirements for Downward Causation

Despite being defined quite extensively in (Woodward 2003), the manipulationist notion of cause, be it a direct or a contributing one, may lead to different formal interpretations. Baumgartner (2009) shows that the existence or not of an adequate intervention generates problems when one is to assess certain causal claims, in particular when no such intervention, actual or counterfactual, seems possible. Such problems manifest themselves in cases of supervenience of a causal variable, as is typical -but not limited- to discussions about mental causation. Consider a variable X that supervenes on a variable Z , for instance a mental state that supervenes on a set of brain states, and let us ask whether any of these two variables or

⁹One might similarly define “upward” and “same-level” directionality conditions, and thereby construe specific notions of “upward causation” and “same-level causation”.

both can play a causal role with respect to a third variable Y that we may assume to be another set of brain states at a later time (see Figure 1). For the sake of simplicity, let us also assume that each of these variables can take two values. Which truth values can be attributed to the causal claims (A) “ X causes Y ” and (B) “ Z causes Y ”?

As shown by Baumgartner (2009), it all depends on how one interprets Woodward’s manipulationist condition **M**. Because X supervenes on Z , there cannot be any change in X without there being also a change in Z . As a consequence, any intervention I on X will also make Z vary. Yet if one wishes to assess whether “ X causes Y ” is true, per manipulationism, one should be able to identify an intervention I on X that would make Y vary while one would hold fixed at some value the variable Z .

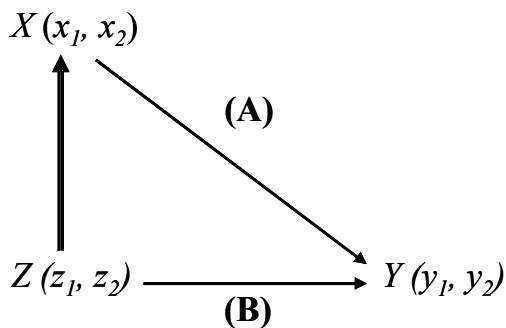


Fig. 1 – A classic illustration of downward causation. Causal relationships (A) and (B) are represented by a single arrow, while the supervenience relationship between X and Z is represented by a double arrow.

On a strict interpretation of **M**, because it is impossible to identify a proper intervention I on X , causal claims such as (A) are deemed to be false (see Baumgartner 2009). Yet, in some recent papers, Woodward asserts that causal claims that are associated with interventions that are impossible for logical, conceptual or metaphysical reasons, ought to be qualified as “illegitimate or ill-defined” (Woodward 2008, 224) rather than false. Following this line of thought, Baumgartner proposes a slightly weaker formalization **M'** of **M** that renders causal claims such as (A) not false but simply ill-defined¹⁰. I would argue that such weaker reading, though

10According to this reading **M'**, causation is defined in the following way: “If there possibly exists an intervention $I = z_i$ on X with respect to Y relative to a variable set \mathbf{V} such that $X, Y \subseteq \mathbf{V}$ and such that all other variables in \mathbf{V} that are not located on a path from X to Y are held fixed at some value while $I = z_i$ is performed on X , then X is a (type-level) cause of Y with respect to \mathbf{V} iff Y changes its value or its probability distribution when $I = z_i$ is performed on X ” (Baumgartner 2009)

justified, is accompanied by a major drawback, namely that of leading to numerous cases of ill-defined causal claims, contrary to our causal intuitions.

Indeed, any case of supervenience-related causal claim, whereby a supervening property might be claimed to be causally relevant, would appear as being ill-defined as soon as one would simultaneously consider the supervenience basis of the supervening property in question. What's more, adding a variable whose values are attributes of the supervenience basis of a supervening property to the set \mathbf{V} of causally-relevant variables turns a previously legitimate causal claim (related to the supervening property) into an ill-defined one. Consider the following set of variables $\mathbf{V} = \{X, Y\}$ where X relates to a traffic light and can take two values: $x_1 = \text{"red"}$, $x_2 = \text{"not red"}$, and Y relates to my driving behavior with $y_1 = \text{"stop the car"}$, $y_2 = \text{"continue driving"}$. Obviously, the traffic light turning to "red" changes my driving behavior and makes me stop, thereby justifying the causal claim " X causes Y ". Yet, consider the set of variables $\mathbf{V}' = \{X, Y, Z\}$, where X and Y are the same as above, and Z relates to the wavelength of the light waves emitted by the traffic light, and can take, for the sake of simplicity, two values $z_1 = \text{"700 nm"}$, $z_2 = \text{"a different value than 700 nm"}$. Because $x_1 = \text{"red"}$ is realized by $z_1 = \text{"700 nm"}$ but also by other wavelength values (typically between 620 and 780 nm), X supervenes on Z . And when one considers \mathbf{V}' , the causal claim " X causes Y " relative to \mathbf{V}' becomes ill-defined since one cannot possibly intervene on X while holding fixed Z at a given value. Such a strategy of introducing, in the variable set \mathbf{V} , a well-chosen variable Z that relates to the supervenience basis of any given variable X might turn any causal claim of the form " X causes Y " into an ill-defined claim. Suffice it to say, this raises questions, not so much about the very account of causation in manipulationist terms, but about one of its modalities, namely the choice of the variables that enter \mathbf{V} . What these examples show is that supervening variables and variables that relate to their supervenience basis ought not be placed in the same set. In other words, such situations call for an additional condition to be imposed on the variables of \mathbf{V} , roughly speaking that of being independent if they are not causally related. In this respect, it appears relevant to go back to a particular "modularity" condition **MOD*** proposed by Hausman and Woodward (2004) which is precisely about such relative independence of non-causally related variables and reads as follows:

MOD* When X_i does not cause X_j , then the probability distribution of X_j is unchanged when there is an intervention with respect to X_i (2004, 149)

As Hausman and Woodward show, this modularity condition is crucial in establishing the Causal Markov condition **CM** within a manipulationist account of causation that takes as premises such conditions as **M** and **IV**¹¹. Furthermore, **MOD*** is crucial in removing known cases where the Causal Markov condition **CM** does not obtain such as cases when **V** omits common causes, when variables are not distinct, when ‘wrong’ variables are defined and measured, or when population is selected by conditioning on a common effect of variables in **V** (Hausman and Woodward 2004, 148).

As a matter of fact, **MOD*** is also crucial in removing the downward causal claims that are ill-defined if one follows Baumgartner’s weaker reading **M'** of manipulationism. Indeed, consider the set of variables **V** that includes both a supervening causal variable and a variable from its supervenience basis. One can see in the examples above that refer to the set **V**= $\{X, Y, Z\}$ (Figure 1), that $\Pr(Z|X) \neq \Pr(Z)$ since X supervenes on Z , and that therefore **CM** fails to obtain. And, similarly, because the probability of Z changes when there is an intervention on X , while X is not a cause of Z , **MOD*** does not obtain either. In sum therefore, making sense of downward causation in manipulationism requires one to add an extra condition such as **MOD*** on the set **V** of variables in order not to end-up with ill-defined claims¹². Hence I propose the following definition of downward causation (**DC**):

11The Causal Markov condition **CM** can be formulated as follows: “For all distinct variables X and Y in the variable set **V**, if X does not cause Y , then $\Pr(X|Y \& \text{Parents}(X)) = \Pr(X|\text{Parents}(X))$ ” (e.g. Hausman and Woodward 2004, 147). (Woodward 2003) does not assume **CM** (hence imposes no such constraints on **V**), and defines “causation” on the basis of “intervention”. On the contrary, (Spirtes et al. 1993) and (Pearl 2000) assume **CM** (hence impose constraints on **V**), and define “causation” and “intervention” on this basis (yet, somehow need to justify **CM**).

12One might rightly argue that adding **CM** would have the same effect; however, in this contribution, I propose to stick to Woodward’s foundational premises for manipulationism and to this incremental addition of **MOD***. One might also rightly argue that it has not been proved that **MOD*** is necessary for removing the ill-defined causal claims related to downward causation, and that a weaker condition might be enough: my point here is that **MOD*** is indeed sufficient for no longer having such ill-defined claims, and that I propose to define downward causation in such a way.

DC X is a downward cause of Y relative to \mathbf{V} iff:

- (i) **M** (Manipulationism) and **IV** (Intervention Variable), i.e. in short X is a cause of Y relative to \mathbf{V} (**M** being interpreted in its weaker reading **M'**),
- (ii) **MC** (Mereological Context), i.e. in short all variables in \mathbf{V} are such that their values are attributes of entities that fit a mereological context,
- (iii) **D** (Downward) i.e. the values of X are attributes of an entity that is at a higher level than the entity of which the values of Y are attributes,
- (iv) **MOD*** (Modularity), i.e. in short given two variables X_i and X_j in \mathbf{V} , if X_i does not cause X_j , then the probability of X_j is unchanged by an intervention on X_i .

Jointly together, these four conditions make it possible to define precisely downward causation within Woodward's manipulationism. So construed, downward causation claims are well-defined (i.e. no longer false as in Baumgartner 2009 or ill-defined as in Woodward 2008) and they are actually ubiquitous¹³. Yet, as we will see, they cannot perform the job that some non-reductive physicalists would like them to do.

5. Downward Causation and Non-reductive Physicalism

A classic argument proposed by some non-reductive physicalists has to do with the causal efficacy of higher-level entities. The argument is not just that higher-level entities may be endowed with downward causal powers, but also that such downward causal powers are somehow novel and cannot be accounted for when one sticks to lower-level entities. This argument has received much support in some domains of the philosophy of mind literature, in particular when it is argued that mental states have real and novel downward causal powers onto physical brain states, and engender effects that brain states -upon which they supervene- are argued not be able to account for. This argument has been termed the argument of "the causal efficacy of the emergents" by Kim and others, and formulated as the claim that "emergent properties have causal powers of their own – novel causal powers irreducible to the causal powers of their basal constituents" (Kim

¹³Consider for instance the set $\mathbf{V}=\{X,Y\}$ where X is the variable "John is smoking" and Y the variable "John's lungs develop cancer". Assuming medical studies have established the truth of " X causes Y " (i.e. **M** and **IV** obtain), one can see that **MC**, **D** and **MOD*** also obtain.

1999, 22)¹⁴. Adopting a manipulationist view of causation, Menzies (2008) and Raatikainen (2010) independently argue in favor of this causal efficacy. The idea is not so much that higher-level variables can be causally relevant, but that lower-level variables may, in some such cases, fail to be causal. As Raatikainen phrases it, “a mental state can truly be a cause of, e.g. a behavior, [whereas] more drastically [...] the underlying physical state may fail to be the cause” (Raatikainen 2010). Based on the definition of downward causation that I have proposed, I will rather argue that, even if downward causes can be made legitimate within a manipulationist account, they cannot exclude lower-level causes.

Consider the example discussed by Raatikainen: assume John is at home and thirsty; what makes him go to the kitchen? His belief (mental state) that there is beer in the fridge or his brain state B (that underlies his mental state)? Following Raatikainen, this situation can be represented by the set of causal variables $\mathbf{V} = \{X, Y, Z\}$, where X refers to John’s mental state, Y to his behavior and Z to his brain state, and where each variable can take two values (Figure 2).

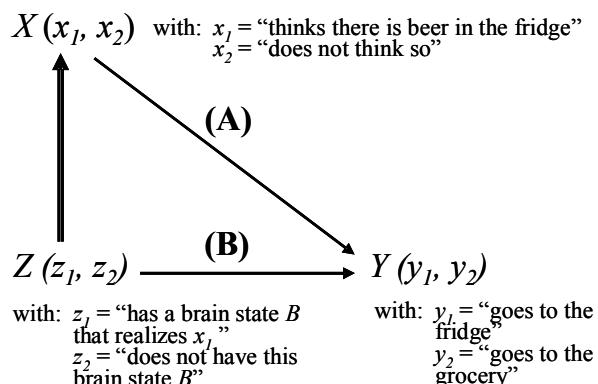


Fig. 2 – Illustration of the downward causation example presented by Raatikainen (2010). Causal relationships are represented by a single arrow, supervenience by a double arrow.

Let us consider the intervention I whereby Peter, John’s roommate, informs John that there is no more beer in the fridge: I changes John’s mental state from $X=x_1=\text{"thinks there is beer in the fridge"}$ to $X=x_2=\text{"does not think so"}$, which leads to a change in behavior from $Y=y_1=\text{"goes to the fridge"}$ to $Y=y_2=\text{"goes to the grocery"}$. According to Raatikainen, I qualifies as a proper intervention and X as a proper cause of Y . Can Z also qualify as a

¹⁴This argument can be understood as a counter-argument to the “causal exclusion principle”; see for instance Kim (1989).

cause of Y ? Raatikainen argues no. Granted the multiple realizability of mental states, there should be another brain state B' that also realizes the same mental state $X=x_1$ = “thinks there is beer in the fridge”. Consider the intervention I' that would change John’s brain state from B to B' ; such an intervention would change the variable $Z=z_1$ = “has a brain state B that realizes x_1 ” to $Z=z_2$ = “does not have this brain state B ”, yet this would not lead to a change in the behavior variable Y , since John still “thinks there is beer in the fridge”. Yet from the point of view of Z , I and I' cannot be distinguished: both interventions on Z switch from $Z=z_1$ to $Z=z_2$. And sometimes Y changes as a result of such intervention, sometimes not. The crucial point, according to Raatikainen, is that the counterfactual “If John’s brain state B were to be changed by an intervention to not having that state, he would have gone to the grocery (and not to the refrigerator)” is false. Thereby, Z fails to be a cause of Y . Hence the causal efficacy of higher-level variables (such as X) and the causal inefficacy of lower-level ones (such as Z). I think there are two problems with this argument.

The first problem relates to how one defines the set \mathbf{V} of variables. If one follows Raatikainen and includes both a supervening variable X and a variable Z that relates to its supervenience basis, then one cannot properly intervene on X while holding fixed Z at some given value; therefore I does not qualify as an intervention, and causal claims such as “ X causes Y ” are ill-defined (Baumgartner 2009). On the other hand, if one properly defines downward causation so as to avoid ill-defined claims (as I have proposed above), one cannot simultaneously include X and Z in \mathbf{V} , in order to satisfy the modularity condition **MOD***. One ought rather to consider $\mathbf{V}'=\{X, Y\}$ and $\mathbf{V}''=\{Z, Y\}$ (see Figure 3). And, relative to the set \mathbf{V}' , but not to \mathbf{V} , it can indeed be argued that X qualifies as a cause of Y and that the claim (A) “ X causes Y ” is true¹⁵.

The second problem relates to the choice of values that variables can take. This problem is associated with the question of choosing what Menzies and Raatikainen call the “causal contrast”. According to Raatikainen, the proper contrast for Z is whether John’s brain state is B or not, hence the choice of the two values z_1 = “has a brain state B that realizes x_1 ” and z_2 = “does not have this brain state B ”. However, as soon as one knows about B' that also realizes $X=x_1$ (and assuming, for the sake of simplicity, that no brain state other than B and B' realizes $X=x_1$) and about B'' that realizes $X=x_2$, one might as well argue that the proper values one ought to consider for Z are z_{11} = “has brain state B that realizes x_1 ”, z_{12} = “has brain state B' that realizes x_1 ”, z_2 = “has brain state B'' that realizes x_2 ”, z_3 = “does not have B nor B' , nor B'' ” (see Figure 3). Consider then the intervention I'' whereby one would change Z from $Z=z_{11}$ to $Z=z_2$. Obviously, such an intervention

¹⁵Strictly speaking, one would need to interpret Raatikainen’s example in such a way that Z relates to brain states (as does Baumgartner (2009)) and that **MC** and **D** obtain.

would be followed by a change in John's behavior, and, contra Raatikainen, Z would qualify as a cause of Y , relative to \mathbf{V}' , and the claim **(B)** "Z causes Y " would be true.

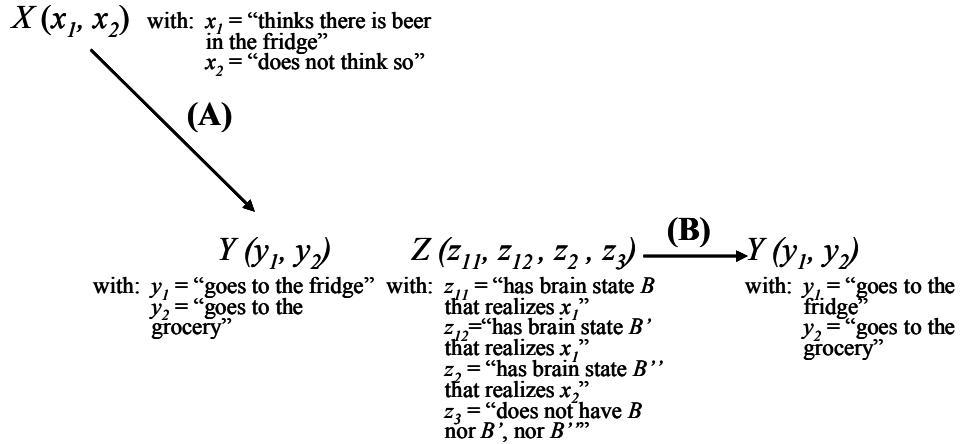


Fig. 3 – Two distinct causal claims restated

One might further argue that, anytime a downward causal claim is true, there exists a corresponding lower-level causal claim that is also true. Indeed, suppose the downward causal claim **(A)** relative to \mathbf{V}' is true (Figure 4).

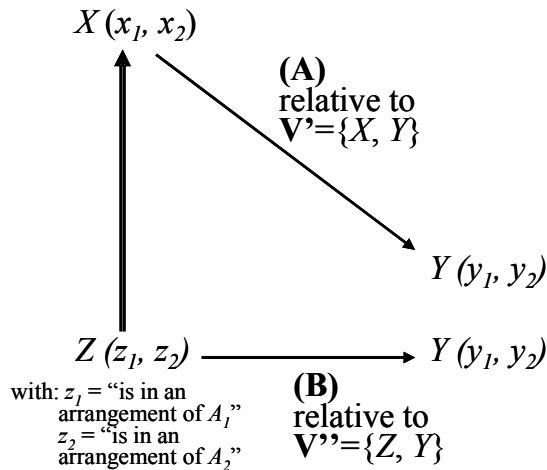


Fig. 4 – "Downward" (A) and "lower-level" (B) causal claims

Then \mathbf{V} satisfies **MC**, and there are lower level entities that compose the upper level entity to which X refers to. When taken together in a certain arrangement, some of these lower level entities make it possible for the

upper level entity that they realize to be such that $X=x_1$ (resp. x_2). Because of multi-realizability, there might be many different such arrangements; let A_1 be the set of all such arrangements (resp. A_2). Let $Z=z_i$ be the lower level property of being in an arrangement that belongs to A_i . Define $\mathbf{V}''=\{Z, Y\}$, and consider the intervention I'' that switches Z from z_1 to z_2 . Then I'' induces a change in Y , and the claim **(B)** “ Z causes Y ” is true.

Therefore it appears that, in agreement with Raatikainen, downward causation is relevant within a manipulationist account of causation, yet, contra Raatikainen, this downward causation does not exclude lower-level causation to also be relevant, and is therefore of no help to the “causal efficacy” argument.

6. Why is Downward Causation *Still* Interesting?

Despite this negative result, I think there is still room for enthusiasm over downward causation, and that this enthusiasm might come from the typological ideal of sorting out causal relations according to some of their intrinsic characteristics: indeed, in addition to having an account of causation, i.e. what it is for X to be a cause of Y , it is also interesting to see whether there might be different ways for X to be a cause of Y (e.g. Woodward 2010). In other words, once one has defined what a causal relation is, one might investigate whether there might be different types of causal relations and try to identify which characteristics capture the most salient dimensions of this diversity. It is in this respect that downward causation might be interesting, for, as I will argue, it might point to specific dynamic features of causal relations. More precisely, the idea is that some intrinsic characteristics of causal relationships might influence our choice of causal variables. In other words, if X and Y end up in the chosen set \mathbf{V} of variables, it is not just because X causes Y but also because that particular causal relationship has specific characteristics that alternatives (e.g. Z causes Y) lack.

As already mentioned above, manipulationist causation is relative to a set \mathbf{V} of causal variables. Yet, often, many variables are available to depict a given phenomenon. Take the case of cancer and the over-simplified causal relation according to which “smoking causes cancer”. Alternative variables to “smoking” could be “smoking cigarettes” or “smoking cigars”, or even “inhaling cigarette smoke” or “exposing one’s lungs to tar”; and similarly, one can imagine alternative variables to “cancer” such as “lung cancer”, “throat cancer” or even “malignant tumor growth in the lungs” etc. An obvious reason why some variables over others end up in \mathbf{V} is simply because there is indeed a causal relationship between them that can be assessed by an actual or a counterfactual intervention. In other words,

variables are chosen because they satisfy the “manipulationist” requirement **M** as well as the “intervention variable” requirement **IV**. And, in the case of downward causation, I have just argued that relevant variables ought also to meet the “mereological context” requirement **MC**, the “downward” requirement **D** and the “modularity” requirement **MOD***. All these requirements put some constraints over the choice of variables, yet alternatives remain possible, in particular when it comes to the granularity of the causal variables one chooses.

In this respect, variable fine-graining may occur under three distinct types that I propose to refer to as “disjunctive”, “conjunctive” and “interpolative”. Disjunctive fine-graining consists in splitting a causal variable X into several causal variables X_i that play similar causal roles, such that X can somehow be understood as referring to either one of the X_i . Consider the variable X associated with the property “smoking”, with X taking different values depending on the intensity of smoking. Because “smoking” may mean “cigarette smoking” or “cigar smoking” or even “passive smoking”, X can be split accordingly into X_1 corresponding to the property of “cigarette smoking”, X_2 to that of “cigar smoking”, X_3 to that of “passive smoking” etc. In this case, each of the X_i plays a similar causal role to that of X with regard to the effect $Y = \text{“having cancer”}$, in the sense that one can state “ X_i causes Y ”. Applying a disjunctive fine-graining to X enables one to investigate more precisely the relative functional causal link between the different X_i and Y ; for instance, even though we know that “smoking causes cancer”, it is also interesting to know whether propositions like “cigar smoking causes more severe cancers than cigarette smoking” or “passive smoking does cause cancer, albeit to a lesser degree than cigarette smoking” are true or not.

On the other hand, “conjunctive fine-graining” consists rather in splitting a causal variable X into several causal variables X_j that can somehow be understood as jointly constituting X . Consider again the example of “smoking”: this activity consists in a set of sub-activities, such as “lighting a match”, “lighting a cigarette with a match”, “inhaling smoke”, “exhalting smoke”, “putting out the cigarette”, “disposing of the cigarette butt” etc. In this case, the X_j do not play similar causal roles with regard to the effect “having cancer”: for instance, “inhaling smoke” is more causally-relevant than “disposing of the cigarette butt”. Applying a conjunctive fine-graining to a given variable enables one to identify more precisely which sub-activities are indeed causally-relevant, and which ones are not.

Lastly, “interpolative fine-graining” consists in identifying variables Z_k causally in between the cause variable X and the effect variable Y . If one may say that “smoking causes cancer”, one might also wish to characterize more precisely the corresponding causal path, for instance by saying that “smoking” causes “lung exposure to tar” that causes “lung tissue injuries”

that causes “absorption of a carcinogen” that causes etc... that causes “cancer”. In this case, the different Z_k do not replace any of the previous variables X , Y , but rather find their way along the causal path from X to Y . “Interpolative fine-graining” enables one to make explicit different crucial causal steps in between a given cause and its effect, thereby providing grounds for more detailed explanations.

Obviously, these three types of causal variable fine-graining are not exclusive of one another but can happen simultaneously¹⁶. They capture different strategies when it comes to identifying causal variables with more precise roles. Each of these types of variable fine-graining may also come with varying degrees of precision, thereby aiming at ever more minute variables. In other words, each of these approaches to identifying ever more causally-relevant variables might be pursued more or less deeply to settle at a given degree of granularity as represented by that of the variables that end up in the set V . In this respect, it is interesting to ask why one settles at one particular degree of causal granularity rather than another one. In particular, relative to the notion of downward causation, one might ask whether there are reasons why one would rather stick to higher-level (and coarser-grained) variables than go with lower level (and finer-grained) ones.

The granularity of causal variables might first be chosen for a set of epistemic reasons. Obviously, the variables one selects must be somehow available. Yet this availability might be constrained by what we know today, that is to say by the current status of our best scientific theories and the variables that they manipulate, as well as by the instruments that we might use to attribute particular values to those variables. Take again the case of cancer. We now know that there are more than a hundred different types of cancer (e.g. Jemal et al. 2008); it is therefore possible today (but not yesterday, so to speak) to choose a variable such as Y_i =“having cancer of type T_i ” rather than the variable Y =“having cancer”. In addition, one might prefer some variables over some others for their meaningfulness relative to a particular cognitive context. Current molecular research on carcinogenesis does not articulate causal claims at the degree of granularity of variables such as X =“smoking” and Y =“having cancer”, which would be meaningless, but rather at a deeper degree of granularity where variables are phrased in terms of concentration of particular molecular compounds (e.g. carcinogens) or activation of particular genes (e.g. oncogenes or tumor suppressor genes).

In addition to such epistemic reasons, one also finds pragmatic reasons for choosing a particular variable graining rather than another one. Such pragmatic reasons might have to do with the usefulness of the chosen

¹⁶In addition to these three types of “variable graining”, one should also consider the degree of “value graining” that each variable might be subject to, and that corresponds to the number of different values that the variable may take. In this contribution, I focus only on “variable graining”.

variables with regards to given objectives. For instance in the health sciences, prevention and cure of diseases are of foremost importance. In this respect, the very coarse-grained causal claim “smoking causes cancer” can be regarded as formulated at the proper degree of granularity provided a public policy against smoking is considered as a lever for reducing cancer cases.

If such epistemic and pragmatic reasons appear legitimate, a most interesting question is whether there might also be some empirical reasons for choosing some type and degree of graining rather than other ones. One might argue, for instance, that the most elementary entities of physics as they appear in quantum theory constitute the most fundamental level of nature, and that, as a consequence, no finer causal graining might ever be achieved. The empirical non-existence of causal variables below that most fundamental level might therefore be considered as an empirical constraint to causal variable graining. In addition, one might ask whether some intrinsic characteristics of causal relations might also provide empirical grounds for preferring a given granularity over another one. In some recent work, Waters (2007) and Woodward (2010) propose that causal relations might be characterized by varying degrees of stability, proportionality or specificity¹⁷. And Woodward, for instance, proposes that such characteristics of causal relationships may “lead to the incorporation of more fine-grained detail in the specification of causes [...] or toward specifications that abstract away from such detail” (2010, 299). In other words, considerations of stability, proportionality and specificity may constrain the causal granularity one chooses and the variables that end up in \mathbf{V} in so far as one favors more stable, proportional and specific causal relationships as somehow “paradigmatic” cases of causation. Such considerations stem from an analysis of the functional relationship between a cause and an effect, i.e. a characterization of how the effect is a function of the cause. I would argue that a dynamic characterization of causal relationships might equally prove useful in mapping out the different dimensions along which to characterize causal relationships¹⁸. In what follows, I will more modestly defend the view according to which some downward causation claims might indeed point to some such interesting features of causal relationships and that such features might also lead to adopting certain degrees of causal granularity over others.

¹⁷Roughly speaking, stability has to do with whether a causal relationship is somehow stable under changes in the background conditions, proportionality with whether changes in the cause induce proportional changes in the effect, and specificity with whether a cause is indeed specific to a given effect (see Waters 2007 for more on specificity and Woodward 2010 for all three notions).

¹⁸ On the possibility of construing diseases, including cancer, as particular dynamical regimes in complex networks, see for instance Gross (this volume).

7. The Case of Tissue Disruption as a Cause of Cancer

As indicated above, some scientists in cancer research advocate the causal role of tissue dis-organization in carcinogenesis, and have voiced claims about downward causation (e.g. Soto and Sonnenschein 2005). Why such claims? One of the motives for choosing a tissue-level causal variable seems to be the underlying complexity of back-and-forth molecular interactions between cells: “cancer is construed as a tissue-level phenomenon within which the numerous cellular interactions that occur simultaneously to maintain the structure of a tissue make it practically impossible to sort out cause and effect into neat linear causal chains at the molecular level” (Soto and Sonnenschein 2005, 115). As we have seen, under normal conditions, neighboring cells inside tissues interact with one another in multiple ways (via adherens junctions, gap junctions, tight junctions or the extracellular matrix – see Bissell and Radisky 2001). These interactions are not only numerous at the molecular scale but run in multiple directions across cellular boundaries. For illustrative purposes, let us imagine the following causal situation based on a limited number of causal variables (see Figure 5).

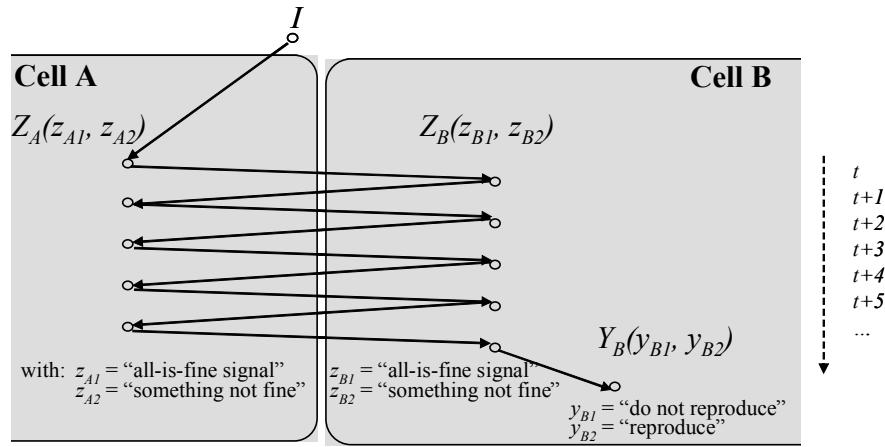


Fig. 5 – Illustration of back-and-forth causal relationships between two cells.

Suppose that one might intervene through I , for instance by adding a carcinogen nearby cell A, onto the signaling variable Z_A of cell A that indicates the type of (molecular) signal that cell A sends to cell B; take this signal to be either “all is fine” or “something’s not fine”. Imagine that Z_A has a causal influence on a similar signaling variable Z_B of cell B, that indicates the type of signal that cell B sends back to cell A. Let us further suppose that signals go back and forth between cell A and cell B a certain number of times before another variable Y_B of cell B, triggered by some sort of

cumulative effect, changes value from “do not reproduce” to “reproduce”, thereby engendering a tumor.

In this causal model, because Z_A , Z_B and Y_B are all cellular variables in the sense that they are attributes of cells, it appears legitimate to qualify the causal relationships between Z_A , Z_B and Y_B as cellular, such causal relationships being somehow instances of “same-level” causation. Yet, because of the back-and-forth interactions between cell A and cell B, the values of Z_A and Z_B at a given time depend on previous values of these same variables. And, strictly speaking, one has as many variables Z_A and Z_B as there are time-increments; one might label these variables $Z_{A,t}$, $Z_{A,t+1}$, $Z_{A,t+2}$ etc. (same for Z_B). The complete causal model at the cellular level in manipulationist terms is therefore based on the large set of variables $\mathbf{V} = \{Z_{A,t}, Z_{A,t+1}, Z_{A,t+2}, \dots, Z_{B,t}, Z_{B,t+1}, Z_{B,t+2}, \dots, Y_B\}$.

Yet, because the different $Z_{A,t}$, $Z_{A,t+1}$, $Z_{A,t+2}, \dots$ and $Z_{B,t}$, $Z_{B,t+1}$, $Z_{B,t+2}, \dots$ are variables whose values depend on one another, and belong to two different cells, one is inclined to group them into a single variable D that would refer to both cells A and B, and that would have a causal relationship to Y_B (see Figure 6).

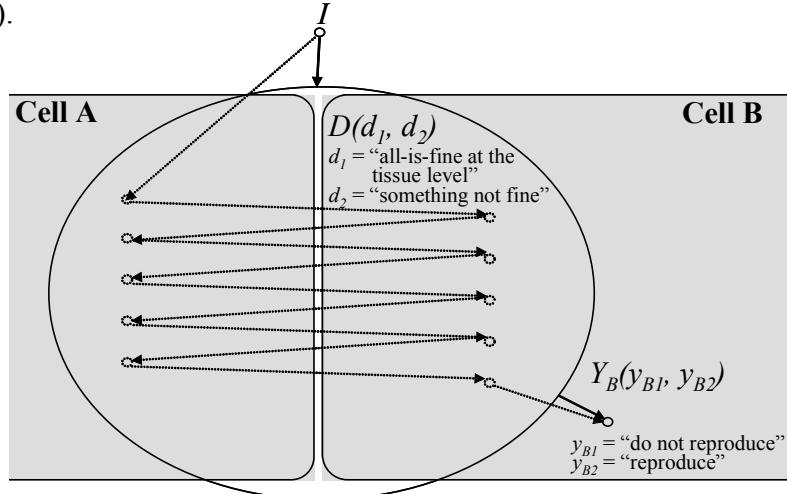


Fig. 6 – Illustration of a downward causal relationship induced by time-dependent variables causally related across distinct lower level entities. The higher level variable D is represented by the large circle; causal relationships relative to this new set of variables are represented with regular arrows; dashed arrows represent lower-level causal relationships.

In this respect, D would be a causal variable at the tissue-level that exerts a causal influence on a cell-level variable, Y_B . And one can show that such a situation would fulfill the formal requirements of downward causation. I would argue that such cases of causation, within which a “lower-level” causal model involves causal relationships that run back-and-forth between two variables that are attributes of two distinct “lower-level” entities and whose values depend on previous ones, naturally lead to an “upper-level” causal model that subsumes under an “upper-level” variable the numerous

lower-level variables just mentioned. Obviously, such a situation would legitimate a downward causal claim. Yet, most importantly, it is cases like this that might indicate how particular dynamic features of causal relationships are indeed salient in the sense of influencing the granularity of our causal models. Such considerations open up possibilities for identifying dynamic characteristics of causal relationships that might complement the functional ones already identified (e.g. Waters 2007, Woodward 2010). Whereas the later focus on a characterization of the functional relationship between a cause and an effect, the former aim at capturing dynamic or temporal features of causal relationships that are not accounted for in purely functional terms. What is at stake is to build a more complete picture of the different forms that causal relationships might take.

8. Conclusion

In this contribution, I have proposed a manipulationist account of downward causation and discussed some ways in which downward causation might be understood as useful or not. Starting from puzzling claims about downward causation and non-reductive physicalism in cancer research, my first objective was to make sense of the notion of downward causation within a manipulationist account of causation. I have proposed several conditions, including the “mereological context” **MC** and “downward” **D** conditions, to be added to manipulationist conditions such as the “manipulationism” **M**, the “intervention variable” **IV** and the “modularity” **MOD*** conditions. I have then argued that, so construed, downward causation makes full sense yet fails to support the non-reductive physicalist claims of the “causal efficacy of the emergents”. I have also argued that downward causation claims might give good hints about the existence of specific dynamic causal relationships that make higher-level variables preferable over lower-level ones, thereby also indicating particular characteristics of such causal relationships. Such considerations open up the possibility for identifying new and dynamic-related characteristics of causal relationships in addition to the functional ones already identified, and to map out, so to speak, a more complete “morphospace of causation”. This may, in turn, impact the usage of causal talk in biomedical contexts, and in cancer research in particular.

Acknowledgments

I am grateful to Michael Baumgartner, Isabelle Drouet, Michael Esfeld, Maria Kronfeldner, Staffan Müller-Wille, Federica Russo, and an

anonymous referee for very helpful comments and discussions. Special thanks also to the audience of the first European Advanced Seminar in the Philosophy of the Life Sciences, Geneva 2010, where an earlier draft of this paper was presented. Finally, I am indebted to the Fondation Brocher, to the Fondation Louis D. of the Institut de France and to the French National Research Agency ANR for financial support.

References

- Baumgartner M., 2009, "Interventionist causal exclusion and non-reductive physicalism", *International Studies in the Philosophy of Science* 23(2): 161-178.
- Bertolaso M., 2011, "Hierarchies and causal relationships in the interpretative models of the neoplastic process", *History and Philosophy of the Life Sciences* (this volume).
- Bissell M. J., Radisky D., 2001, "Putting tumors in context", *Nature Reviews Cancer* 1: 46-54.
- Broad C. D., 1925, *The Mind and its Place in Nature*, London: Kegan Paul & Co.
- Downward J., 2006, "Prelude to an anniversary to the RAS oncogene", *Science* 314: 43-44.
- Gross F., 2011, "What can systems biology tell us about diseases?", *History and Philosophy of the Life Sciences* (this volume).
- Hanahan D., Weinberg R. A., 2000, "The hallmarks of cancer", *Cell* 100: 57-70.
- Hausman D. M., Woodward J., 2004, "Modularity and the causal Markov condition: a restatement", *British Journal for the Philosophy of Science* 55: 147-161.
- Hempel C. G., 1969, "Reduction: Ontological and linguistic aspects". In Morgenbesser S., Suppes P., White M. (Eds) *Philosophy, Science, and Method: Essays in Honor of E. Nagel*, New York: St. Martins Press, 179-199.
- Jemal A., Siegel R., Ward E., Hao Y., Xu J., Murray T., Thun M. J., 2008, "Cancer statistics 2008", *CA Cancer Journal for Clinicians* 58 (2): 71–96.
- Kim J., 1989, "The myth of non-reductive materialism", *Proceedings and Addresses of the American Philosophical Association*, 63: 31-47.
- Kim J., 1999, "Making sense of emergence", *Philosophical Studies*, 95: 3-36.
- Kutschenko L., 2011, "Why classify diseases? Revisiting the epistemological function of medical classification systems", *History and Philosophy of the Life Sciences* (this volume).
- Malaterre C., 2007, "Organicism and reductionism in cancer research: Towards a systemic approach", *International Studies in Philosophy of Science* 21(1): 57-73.
- Menzies P., 2008, "Exclusion problem, the determination relation, and contrastive causation". In Hohwy J., Kallestrup J. (Eds.), *Being Reduced - New Essays on Reduction, Explanation and Causation*. Oxford: Oxford University Press, 196–217.
- Morange M., 2007, "The field of cancer research: an indicator of present transformations in biology", *Oncogene* 26: 7607-7610.
- Pearl J., 2000, *Causality. Models, Reasoning and Inference*. Cambridge: Cambridge University Press.
- Raatikainen P., 2010, "Causation, exclusion, and the special sciences", *Erkenntnis* Online DOI 10.1007/s10670-010-9236-0
- Russo F., 2009, *Causality and Causal Modelling in the Social Sciences*. Methodos

- Series: Springer.
- Russo F., Williamson J., 2011, "Epistemic causality and evidence-based medicine", *History and Philosophy of the Life Sciences* (this volume).
- Sonnenschein C., Soto A.M., 2008, "Theories of carcinogenesis: An emerging perspective", *Seminars in Cancer Research* 18(5): 372-377.
- Soto A.M., Sonnenschein C., Miquel P.A., 2008, "On physicalism and downward causation in developmental and cancer biology", *Acta Biotheoretica* 56: 257-274.
- Soto A.M., Sonnenschein C., 2005, "Emergentism as a default: cancer as a problem of tissue organization", *Journal of Biosciences* 30(1): 103-118.
- Soto A. M., Sonnenschein C. 2006, "Emergentism by default: a view from the bench", *Synthese* 151: 361-376.
- Spirites P., Glymour C., Scheines R., 1993, *Causation, Prediction and Search*, Cambridge: MIT Press.
- Waters K. C., 2007, "Causes that make a difference", *Journal of Philosophy* CIV:551–579.
- Weinberg R.A., 1998, *One Renegade Cell: How Cancer Begins*. New York: Basic Books.
- Woodward J., 2008, "Mental causation and neural mechanisms". In Hohwy J., Kallestrup J. (Eds) *Being reduced - New essays on reductive explanation and special science causation*, Oxford: Oxford University Press, 218–262.
- Woodward J., 2010, "Causation in biology: stability, specificity and the choice of levels of explanation", *Biology and Philosophy* 25: 287-318.
- Woodward J., 2003, *Making Things Happen: A Theory of Causal Explanation*, Oxford: Oxford University Press.