**This appeared in Kendler, Parnas and Zachar (eds.) *Levels of Analysis in Psychopathology: Cross Disciplinary Perspectives.* Cambridge University Press, 2020, pp. 424-449.**

This version may differ slightly from the published version. Please quote from latter.

**Levels: What are they and what are they good for?**

**Jim Woodward**

**HPS Pittsburgh**

**1**. **Introduction**

As many writers have observed, talk of “levels” (of description, organization, explanation etc.) is common in many areas of science. At the same time, many also complain that such talk is unclear and problematic in many ways. Part of the problem is that researchers operate with many different understandings of levels and how to distinguish them. Often the differences between these various notions are insufficiently appreciated with the consequence that researchers mistakenly attribute features associated with one notion of level to other notions where they do not apply. This may prompt the thought that we would better off avoiding level talk entirely (so-called “levels eliminativism”—see Eronen, 2015). I think this conclusion would be a mistake. Despite its multiple meanings (and in some respects, unclarity), level talk does useful work and, as I will illustrate below, there are many important methodological ideas that are naturally expressed in terms of levels.

 This paper begins (Section **2**) with a partial typology of some of the different ways in which the notion of level has been understood — these are not wholly independent of each other but they are far from exactly coinciding. Because of this variety, the search for a single general account that captures all legitimate uses of “level” is unlikely to be successful. However, awareness of the important differences among level notions can guard us against various confusions.

 A great deal of recent discussion of levels has focused on the implications of this notion (or notions) for causal reasoning. Can there be causal relations between factors at different levels (e.g., can there be “downward causation” from upper to lower levels) or must all causation be intralevel? I explore this issue in Sections **3-5**, focusing largely on influential recent discussions by Bechtel and Craver. In contrast to their views, I argue for the intelligibility and importance within science of interlevel causation and in particular downward causation.

 Sections **5-6** then attempt to put these ideas within a more general framework, linking issues about levels to issues about variable choice. It is a common and I believe entirely correct observation that different systems are best understood at different levels of analysis (or perhaps in terms of some combination of levels)—where this means that the systems are more appropriately analyzed or understood in terms of certain explanatory or causal variables (belonging to different levels) rather than others. I will describe one kind of consideration which has to do with what I call *conditional independence* that bears on this issue and which, I claim, helps to illuminate why “level talk” is often useful.

 A general feature of this account is that which level or levels are most appropriate turns out to be an empirical matter, dependent on the details of the behavior of the systems we are trying to understand. Sometimes “lower” or more fine-grained levels of analysis will be most appropriate and sometimes more “upper-level” variables will be most appropriate. In particular, it is a mistake to suppose that we always improve the quality of an explanation or causal analysis by invoking lower-level variables. In addition, as we shall see, the considerations that figure in variable choice will sometimes lead us to accounts which include (in a single model) variables are at different “levels” on some understandings of that notion—so-called multi-level or mixed level models. Thus another related theme is the contrast between such multi-level models and theories that appeal to variables that are at least for the most part at a single level.

**2. Some Different Notions of “Level”**

I begin, then, with a taxonomy of some of the different notions of level that have figured in recent discussion. But before getting started a brief aside on the language of “upper” and “lower” is in order. Although talk of upper and lower levels can seem very natural, it is, as Denis Noble remarks (2006), puzzling if one tries to take it literally. After all, it is not as though the “lower-level” atoms and molecules that are taken to compose “higher-level” cells are spatially below the latter or that the atoms are part of some edifice in which they function like a basement or ground floor and the cells as the upper stories. Exactly why we find the language of upper and lower in this context so natural is an interesting question – perhaps it is connected to expectations that explanations in terms of lower-level entities will be “deeper” or “more fundamental” than those that appeal to upper-level entities, with “deeper and more fundamental” being associated with what is lower. In any case, I am going to follow others in availing myself of this language, while at the same time recognizing that it involves a somewhat odd metaphor.

Let me also add (as should be obvious) that the list that follows is not meant to be exhaustive. Also the different notions of level that I discuss are not wholly distinct but rather partially overlap and interrelate in various complex ways.

**2.1. Levels as Compositional**. One of the most familiar notions is a compositional or mereological understanding of levels: objects (or entities) at a higher level are “composed of” or “made up of” or “constituted by” objects at lower levels in a way that generates at least a local hierarchy of levels. In a paradigmatic case, the lower-level objects are (or at least are thought of as) spatial parts of higher-level objects: Atoms are composed of protons, neutrons and electrons, molecules are composed of atoms, cells are composed of molecules, organs (like the brain or liver) are made up of different kinds of cells, multicellular organisms are composed of a range of different organs, social groups or societies are composed of organisms and so on. Each item on this list is regarded as at a “higher” level than the items of which it is composed. We find this idea (among others) in Putnam and Oppenheim’s classic paper (1958), and it often seems to be the preferred conception of levels both among metaphysically inclined philosophers and among some philosophers of biology, particularly those that stress the role of mechanisms in biological explanation. As noted below, some philosophers extend this compositional notion of level to include items such as processes or “activities” which they take to be related by temporal part/whole or “constitutive” relations, so that, e.g., long term potentiation is understood as one of the constituents of the formation of memories, with the latter being at a higher level than the former (Bechtel and Craver, 2007).

**2. 2. Levels as Individuated in Terms of Disciplinary Subject Matters**. Another common notion links levels to the present organization of disciplines or perhaps to differences among methods or research strategies that are presently characteristic of those disciplines. For example, one might take the “psychological level” to have to do with whatever psychologists study, the “biological level” with whatever biologists study and so on. Or one might take the psychological level to have to do with what can be discovered by research methods that investigate certain aspects of the behavior of whole organisms in contrast to research at the physiological or neurobiological level which instead involves measurements performed on brains, hormones and so on. Obviously these are not very principled or stable notions of level, but I think they clearly influence how researchers think about levels. One problem with this conception is that changes in research practices or interests (which can occur for any number of different reasons) will, on this conception, change what belongs to a given level—certain molecules will become entities at “the” biological level as biologists begin to study them, the brain will be thought of as at the level of economics to the extent that economists venture into neuro-economics and so on. Moreover, this notion of level is closely tied to how scientific disciplines are individuated in a way that threatens to make it very misleading unless those disciplines are individuated very finely-- “biology” may be thought of as a scientific discipline but talk of “the” biological level (or for that matter “the” neurobiological level) is obviously highly problematic. Some would subdivide even subdisciplines – such as the distinction between molecular and systems neuroscience. There is a tendency (see, e.g. Putnam and Oppenheim, 1958) to suppose that levels in this sense (individuated in terms of disciplinary subject matter) will coincide with levels characterized in other ways—e.g., in terms of composition -- so that the subject matter of chemistry has to do with “parts” of the objects biologists study but this also is highly problematic.

 **2. 3. Levels as Related to Abstractness and Coarse-Graining**. A common intuition about levels is that variables that are more fine-grained or specific are at a “lower level” than variables that are related to them by some sort of coarsening operation. “Coarsening” can take many different forms but often the idea is that the lower level more fine-grained variables are related to the more coarse-grained upper level variables via some many-to-one function—e.g. by averaging[[1]](#footnote-1). Consider a variable specifying the kinetic energy of each of a number of individual molecules composing a gas. This variable is typically understood to be at a lower level than a variable ascribing an average kinetic energy to the entire collection of molecules. Similarly for the relationship between total caloric intake and a more detailed breakdown of the individual foodstuffs making up those calories. As we shall see, coarsening operations of this sort are important in many areas of science—not because coarsening per se is valuable but because it is sometimes associated with other features of good explanations, including what I call conditional independence (Section **6**). Coarsening represents one form that “abstraction” can take in science with different degrees of abstraction being associated with different “levels”.

 **2. 4. Interactionist Conceptions of Level and Conditional Independence**. Here I will focus on one particular kind of case in which considerations having to do with strength and character of interactions affect assignments of levels. Suppose that some set of variables — *L* – influences some effect *E*, but that we can adequately summarize the impact of the *L* variables on *E* in by means of some much smaller number of variables or parameters (or variables with a small number of possible values) *U*, where the *U* variables correspond to a coarsening of the *L* variables. In other words, we find that *conditional* on the values of the *U* variables, further details about the *L* variables don’t matter for *E*, so that we can explain *E* just as well (or at least as well) by means of *U*, as we can by appealing to *L*. When this condition holds I will say that the *L* variables are *independent* of *E*, *conditional* on the *U*. (I will make this more precise in Section **6**) When this condition holds, we are likely to think of *U* as at a different (and “higher”) level than *L*. I call this conception of level interaction -based because it is based on the idea that the causal impact of the *L* variables on *E* can be fully represented or captured by the impact of the *U* variables on *E* (or at least that this is true to some very high level of approximation). For example, if we are interested in the relation between thermodynamic variables, we don’t need a detailed description of the behavior of individual molecules-- their aggregate impact as expressed in terms of variables like temperature or volume is all that is needed. Given the values of these thermodynamic variables, further variations in molecular details do not matter to the values of other thermodynamic variables. This is one consideration that leads us to think of the thermodynamic variables as at a different (and higher) level than the variables characterizing the behavior of individual molecules[[2]](#footnote-2).

 This interaction-based notion of level is often tied to ideas about the role of considerations having to do with “scales”—spatial, temporal and energetic – in constructing theories and models: sometimes when nature is kind we have “separation” or near separation of scales, so that what happens at one length or energy scale can be understood largely independently of what happens at other scales, and this in turn leads us to think of interactions at one scale as at a different level than interactions at other scales. Relatedly, in physical contexts, this notion is tied to the idea that “different physics” can be “dominant” at different scales, so that, with respect to many explananda (like rates of flow through a pipe), one can model the behavior of a fluid as though it is continuous, using the Navier-Stokes, equations, even though it is made up of discrete molecules, a fact that matters at other scales (Batterman, forthcoming a). Here the values of the continuous variables in the Navier-Stokes equations play the role of the *U* variables in the schema above, with variables characterizing the behavior of individual molecules playing the role of the *L* variables.

 It is important to understand that “scale” matters in this sort of framework because and to the extent that it bears on the kinds of interaction that occur—that is, size or the difference between longer and shorter times do not matter in themselves but only because (or to the extent that) they are thought to be related to the strength and nature of causal interactions and to what can be safely ignored (for purposes of understanding such interactions) conditionally on other information. So the idea is *not* that little things cannot interact (or rarely interact) with big things or vice-versa (after all viruses can destroy armies); it is rather that causal interactions often have systematic relations with spatial and temporal relationships, and that this is tied to what sort of causal details can be ignored or summarized by means of other variables.

 Put differently because this conception focuses on causal interactions, it is the character and relata of those interactions that figure in or are relevant to the characterization of levels. Note also that on most accounts of causation, including the interventionist account I will endorse below, *things* (or parts of things) per se don’t stand in causal relationships -- instead the causal relata are something more like properties or magnitudes or whatever is described by variables. By contrast, paradigmatic cases of compositional relations involve relations among things. This by itself marks a fundamental difference between composition- based and interaction-based conceptions of levels.

 As another illustration of this interactionist notion of level and its connection to the notion of scale, consider that for the purposes of understanding what is going on within the nucleus and phenomena such as radioactive decay, two of the four fundamental forces—the strong and weak nuclear forces-- are crucial. These forces are very strong at very short spatial scales. Gravity, another fundamental force, is effectively irrelevant for most purposes in modeling nuclear behavior. On the other hand, if we are interested in explaining/understanding chemical behavior—how atoms combine and form molecules and compounds—the strong and weak nuclear forces are effectively irrelevant in the sense that detailed features of these forces do not matter as long as they allow for the formation of stable nuclei. Instead, another force, the electromagnetic force plays a central role in chemical behavior. In many cases, this separation of levels or scales—the fact that nature permits us to construct theories that explain aspects of nuclear behavior that appeal to factors that are different from those that are required to explain chemical behavior, so that we can do nuclear physics without doing chemistry and vice versa – is crucial for successful science. As explained in more detail below, this separation enables us to avoid, at least to some considerable extent, problems of computational and other sorts of intractability that would likely otherwise hamstring scientific understanding. We avoid such problems by neglecting what is going on at other length, time and energy scales and we are justified in doing so because what goes on at these other scales either does not matter (much) for the phenomena we are trying to understand or else, to the extent that such goings on do matter, their influence can be summarized in a small number of manageable parameters or variables at the scale at which we are modeling, as in the thermodynamic case[[3]](#footnote-3).

 Both length and temporal scales are also important in biological contexts[[4]](#footnote-4). Some biological variables may change so slowly with respect to others and to the explananda of interest that the former can be effectively treated as constants – thus variations in them make effectively no difference for the problem at hand. For example, although large scale anatomical connections in the brain can change over time, they generally change much more slowly than large scale “functional” connections and thus the former can be treated as effectively unchanging for purposes of understanding the latter. Other variables may reach an equilibrium so quickly that they can also be treated as non-varying. Again, this can justify ignoring or greatly simplifying interactions involving those variables, as when we summarize the impact of a variable by a single constant.

In my opinion, this interactionist notion is one of the most scientifically interesting and relevant notions of “level” although it can easily mislead us, as discussed below. I emphasize again that this basis for level talk (which interactions are important and which can be ignored) is conceptually distinct from the issues about composition or size that figure in the first notion of level distinguished above. Whether one object X is part of another object Y is obviously a distinct question from whether one can safely ignore features of X in explaining the behavior of Y[[5]](#footnote-5). Nuclei are “parts” of molecules but, as noted above, nuclear forces can be safely ignored in understanding the chemical behavior of molecules. Electrons are also parts of molecules but aspects of the behavior of electrons that have to do with electromagnetic forces are crucial to understanding chemical behavior. More generally, as noted above, little things can clearly interact with big things and conversely. Things that belong to the same kind or type as things that compose big things can interact with those big things. Proteins in the form of prions can affect brains and people even though proteins are among the constituents of brains and people. Thus strength of interaction considerations and ignorability considerations are only very imperfectly related to size differences or to compositional relationships. Although as noted above, compositional notions of level have been the subject of a lot of recent attention, especially among metaphysicians and philosophers of biology, in my view such notions have less bearing on interaction-based notions of level than many have supposed.

 Several other preliminary observations may be helpful at this point. Interactionist conceptions of level are, as I have said, based on the assignment of systems or variables to the same level to the extent that they interact strongly with one another and to the extent that various details about what is going on at other levels may be safely ignored or summarized by a much smaller number of variables. In some cases, this rationale for level assignment works smoothly and is very fruitful, as is illustrated by examples discussed above (and below). In these cases, we are able to identify a set of phenomena all of which can be explained in terms of some relatively small set of explanatory factors that interact primarily with one another and this encourages us to think of all these factors as “at the same level”. Put differently in such cases, the phenomena and explanatory factors to which we appeal are relatively self-contained: we don’t have to bring in other factors willy-nilly from some arbitrarily long and disparate list to account for what is going on. Thus general considerations based on scale and other factors help to provide principled motivations for restricting the list of possible explanatory factors and regarding them as separable from other explanatory factors. This sort of situation is sometimes described by physicists by means of talk of distinct “regimes” or “protectorates”—the idea being that there is a distinct, separable set of phenomena that can be adequately modeled by some restricted set of variables or explanatory factors. When we are in this sort of situation, level-based considerations can be very heuristically useful in guiding judgments about which causal factors can be safely ignored[[6]](#footnote-6).

 Although it is not uncommon to find explanatory factors and explananda that behave in the way just described – separable regimes and so on, this is by no means always the case. For this to be possible, nature has to cooperate and it does not always do so—an alternative possibility is that the list of possible explanatory factors for some set of explananda may be very open-ended and at different levels (by one or more of the criteria described above) with the kinds of heuristic arguments that in some cases allow us to exclude certain factors unavailable. Arguably this is the case for many explananda in psychiatry: many different causal factors—social or environmental factors, factors having to do with personality type, as well as genes and brain structure -- are relevant to paradigmatic mental illnesses such as depression and we do not have, as we do in the case of the nuclear and electromagnetic forces, strong general arguments that certain factors could not possibly be relevant (or at least any general arguments of this sort are much weaker than they are in the physics case.[[7]](#footnote-7))

 This fact can generate lots of misguided expectations and assumptions if we are not careful. As an illustration, consider the causal influence of environmental factors such as stressful life events on gene expression and mental illness. Assuming that such influences are real, what do they imply about levels? Suppose that we adopt a conception of levels according to which factors that interact are automatically at the same level (as one simple version of an interactionist conception might suggest). Then it follows that environmental factors and genes are at the same level. Of course this is not the way we usually think about environmental events like the death of a parent—intuitively we think of them as at a “higher level” than gene expression. Such intuitive judgments must reflect the influence of other conceptions of level besides the simple interaction-based conception described above: perhaps we are thinking that genes are parts of organisms, that environmental stressors involve whole organisms[[8]](#footnote-8), hence (via a composition-based notion) that the stressors must be a “higher level” than genes. Perhaps we are also thinking that environmental stress belongs to the subject matter of the social sciences and psychology and that these disciplines are at a “higher level” than genetics.

 Suppose we try to retain the idea that objects and systems at the same level interact preferentially or even exclusively with each other *and* combine this with a notion of level based on something other than interaction (e. g., a size or composition- based notion, as above). Now we have a recipe for confusion: on the one hand, the occurrence of environmental stressors is a higher-level event (based on size and composition considerations) than genes; on the other hand factors at different levels are not supposed to interact (much) with one another, so that it can seem problematic that environmental events can influence genes (so that some special story needs to be told about how to analyze the appearance of such influence—for example, by claiming that what is really doing the influencing is some set of much lower-level events that realize the environmental event). Accordingly, one finds, both in the philosophical literature and elsewhere, a number of arguments to the effect that objects and systems at different levels cannot interact with one another (or that such interaction is problematic and needs to reinterpreted in a way that makes it philosophically respectable, as in the views of Bechtel and Craver, discussed below). My opinion is that such views derive, at least in part, from illegitimately combining expectations that come from an interactionist view of levels with conceptions of level that are based on other sorts of considerations such as part/ whole relations. On my view, there is no problem in principle with the notion of inter-level causation, including the notion of downward causation from “upper” to “lower” levels. The temptation to think otherwise results, at least in part, from thinking of levels too exclusively in compositional terms, which helps to makes it seem plausible that downward causation should be modeled in terms of causal relations running from wholes to their parts.

**3. Causation and Levels**.

I turn now to some more specific remarks about causal relations between variables at different levels. This requires an account of causation and here I’m going to proceed within the interventionist framework that I have defended elsewhere (Woodward, 2003). According to this framework:

(**M**) *X* causes *Y* in background conditions B if and only if under some intervention that changes a value of *X* in B, the value of *Y* will change in a regular or stable way. (For our purposes we can think of “regular” here as meaning either that *Y* takes the same value for the intervention on the value of *X* or that *Y* exhibits the same stable expected value or at least that this is approximately true.)

Here *X* and *Y* are variables which can be either binary (corresponding to the occurrence or non-occurrence of some event) or can take many different values, as with real-valued variables. An intervention on *X* with respect to *Y* can be thought of as an idealized experimental manipulation of *X* which changes *X* in such a way that any change in *Y*, should it occur, occurs only through the manipulation of *X*. Put slightly differently, an intervention on *X* is an unconfounded manipulation of *X*. Interventions can be realized in, for example, randomized experiments among other possibilities.

The idea underlying **M** is that causal relations are relations that are potentially exploitable for purposes of manipulation and control; if *X* causes *Y* then if one wiggles *X* in the right way, *Y* should change. Conversely, if *X* changes under some intervention on *Y* then *X* causes *Y*. This does *not* mean that causal relationships can only be discovered through experimentation or by actually performing interventions -- I can certainly learn about causal relationships in non- experimental or “purely observational” contexts by employing various sorts of causal modeling procedures. However, according to (**M**), when I learn about causal relationships in such contexts what I am learning is what the results of a possible experiment involving manipulation of *X* would be, were I to perform the experiment. One consequence of this picture is that the kind of evidence that supports causal conclusions on the basis of non-experimental data should be evidence that would support conclusions about the outcomes of hypothetical experiments. I won’t try to defend this idea here but it is becoming increasingly accepted within statistics and econometrics —it underlies the use of instrumental variables and regression discontinuity designs (which involve looking for intervention-like processes in observational data), that are increasingly used in psychiatric epidemiology.

 Note that (**M**) by itself imposes no constraints connecting causal claims with the various non-interactionist notions of level. As far as (**M**) is concerned, a variable that is identified as “upper-level” according to some criterion like composition or abstractness—e.g., a variable like environmentally induced stress *S* or famine *F* -- can cause a lower-level variable having to do with e.g., a certain pattern of gene expression *G* as long as it true that under the right sort of wiggling of *S* or *F*, *G* would change in a regular way. That this is the case might be established either by experimental manipulation of the upper-level variable or from observational studies or a combination of these. For example, experimental manipulation of stressors imposed on laboratory animals can be shown to alter gene expression and observational evidence from a variety of sources (including instrumental variable type reasoning such as observations from the Dutch Hunger Winter) supports the conclusion that famine experienced by mothers can alter gene expression in their offspring. Such procedures can establish that upper-level variables have causal impacts on lower-level variables without having detailed causal information about exactly how the upper-level variables are realized by lower-level variables[[9]](#footnote-9). Of course if we say that variables are automatically at the same level as long as they causally interact, we will be led to conclude that *S* and *G* are at the same level but the point is that (**M)** imposes no further level-based constraints on what can cause what. For similar reasons, as far as **M** goes, variables that are “lower level” by some criterion (e.g., variables having to do with gene expression) can causally influence upper-level variables like whole organism psychology or behavior.

In my view this “level neutrality” is a virtue of (**M**)—it allows us to judge that the very common use of theories in the life sciences and the social and behavioral sciences in which there are causal relations holding among variables at different levels is legitimate and unproblematic and it fits naturally with the information on which we often rely to establish such conclusions[[10]](#footnote-10).

 **4. Levels and Downward Causation**

At this point, non-philosophers may ask, regarding the possibility of causation across levels: Why would anyone suppose otherwise? In fact several different reasons have been advanced (primarily but not exclusively by philosophers) for why causation “across” different levels, and in particular, “downward causation” from an upper to lower levels, is impossible. In this and the following section I explore and respond to several of these arguments, focusing first on these issues at a general level and then more specifically on an influential recent discussion by Bechtel and Craver (2007).

 A common complaint against downward causation is that this involves causation running from a whole to its parts and that this is always incoherent; typically this is claimed to be so because wholes and parts are not “suitably distinct” to stand in causal relationships and/or because whole/part relations are synchronic in a way that causal relationships are not (because, it is supposed causes must temporally precede their effects)[[11]](#footnote-11). A similar objection would apply to upward causation from parts to wholes. Thus Craver and Bechtel take as an example of a claim of downward causation the claim that “signal transduction” in the visual system causes changes in the “conformation of rhodopsin” and object that, because the latter is a temporal stage in the former, the relation between the two cannot be causal (2007, p. 552). For similar reasons they object to the claim that a mechanism considered as whole (that is, as a collection of parts or constituents standing in ordinary causal relations with each other) can exert downward causation on the parts or constituents of that mechanism. Heil (2017) takes as one of his paradigms of downward causation, the claim that the motion of a whole body of water causes the motion of its component molecules and objects that this involves a whole causing its parts, which is incoherent.

I agree that these (and other similar synchronic whole/part relations) should not be understood as causal relations. However, for the most part, these are *not* the sorts of examples that are described as cases of downward causation in the recent scientific literature. To begin with, as noted above, causal claims in most areas of science (and causal claims as these are understood within the interventionist framework) relate *variables* (or more pedantically what is described by variables) rather than things or thing-like entities such as processes or events. By contrast, whole/part relationships do relate thing-like entities—bodies of water and the molecules that compose such bodies, temporally extended processes and component sub-processes and so on. In my view, the latter are not appropriate candidates for the relata of causal relationships of any kind, whether interlevel or intralevel. In other words, to the extent that whole to part relations relate thing-like entities—and it is unclear what else the relata of such relations might be— this alone is a sufficient reason for disqualifying them as causal.

An even more important consideration is that when things stand in whole/part relationships, variables predicated of those wholes and parts need not (and usually do not) stand in whole/ part or constitutive relations of a sort that are inconsistent with a causal interpretation. As an illustration, consider that in the Hodgkin-Huxley model of the action potential, the potential difference *V* across the cell membrane is treated as a cause of the opening and closing and conductances of the ion channels in the cell membrane and of the various ionic currents through the membrane. The ion channels are literally part of the cell membrane and thus on a compositional conception one might think of the causal influence of the membrane potential on the ion channels as a matter of upper to lower or downward causation, which indeed is how it is often described (e.g. by Noble, 2006). However, although the ion channels are part of the membrane, it is dubious that it makes sense to describe the strength *Ii* of the ionic currents as “part” or a “constituent” of the membrane potential *V*. (I will say more about this later.) Moreover, whether or not this “parthood” language is appropriate, it seems clear that *V* and *Ii* are nonetheless “distinct” in a way that allows them to stand in causal relationships. In fact, the *V🡪 Ii* relation straightforwardly satisfies the interventionist criterion for causation; if one intervenes on the membrane potential the ionic currents will change. Indeed, Hodgkin and Huxley actually did this experiment with the then new device of a voltage clamp which allowed them to impose different potentials across the cell membrane and measure the resulting changes in the ionic currents.

Similarly, even if one thinks that genes are parts of people (or, if this makes sense, “parts” of environmental interactions involving people), and that the former are at a lower level than the latter, according to **M** this is no barrier to the truth of a causal claim according to which some high levels of stress cause changes in gene expression. Again an interventionist framework makes straightforward sense of such downward causal claims—one can, for example, manipulate social stress among laboratory animals and measure corresponding changes in gene expression.

**5. Bechtel and Craver on Downward Causation and Mutual Manipulability**

 So far I have been objecting to the idea that claims of downward causation are to be understood in terms of part/whole relations (and hence are illegitimate). I want now to flesh out this portion of my discussion by turning to a more detailed look at one of the most developed and influential discussions of “levels” and downward causation in the philosophical literature, which is due to Bechtel and Craver (2007). These authors advance an account of levels according to which a necessary and sufficient condition for components to be at the “same level” is that they all be components or constituents of the same “mechanism”. That is, components of the same mechanism and only those components are at the same level. This basically relies on a compositional notion of level but the resulting notion is, as these authors stress, very “local” in the sense that components of different mechanisms will be non-comparable with respect to levels.

 Bechtel and Craver advocate the following condition (called **MM** for Mutual Manipulability) for whether something is a component of mechanism or more generally a constituent or part of something.

(**MM**) X and S are related as part and whole (X is a constituent of S) if and only if *F* is some behavior of X and *J* some behavior of S such that

(i) there is an intervention on X’s *F*-ing with respect to S’s *J*-ing that changes S’s *J*-ing;

(ii) there is an intervention on S’s *J*-ing with respect to X’s *F*-ing that changes X’s *F*-ing (Craver, 2007 , p. 153).

In other words, X is a part of S if and only if there is some way of intervening on X to change its behavior that changes S’s behavior and some way of intervening to change S’s behavior that changes X’s behavior. For example, rhodopsin conformation is a constituent of signal transduction because there is a way of intervening on the former that will change the latter and vice –versa. (Here rhodopsin conformation is the behavior corresponding to *F* in the above schema and signal transduction corresponds to *J*. Signal transduction is a behavior of the whole visual system -- S in the above schema-- and rhodopsin conformation is a behavior of a part X of S.) (**MM**) is satisfied because there is a way of intervening on the extent of rhodopsin confirmation that would alter overall signal transduction and, similarly, appropriate interventions on signal transduction would alter rhodopsin conformation. According to Bechtel and Craver, the presence of this sort of constitutive relationship between X and S precludes that relationship from being causal.

It follows from this picture that there is no such thing as inter-level causation literally speaking; what looks like inter-level causation is really a matter of the operation of intra-level causation and constitution relations, as captured by **MM**. When upper-level *U* appears to cause lower-level *L* which is really going on is that some set of lower-level constituents *C* of *U* ( where the *C*s at the same level as *L*) cause whatever happens with respect to *L*. Although in such cases we may observe *L* change after changes in *U*, strictly speaking the real causal action is all at the level of *L* (and C); *U* doesn’t do any causing. Apparent top-down causation is thus what Bechtel and Craver call a “hybrid” relation, that can be decomposed into a constitution relation between *C* and *U* and a causal relation between *C* and *L*, rather than a causal relation between *U* and *L*. To use another of their examples, the macro-level event *M* of Kane experiencing the loss of a sled leads to various lower-level neural events *N* as traces a memory of *M* are laid down—an apparent case of downward causation. However, this can be translated into a claim to the effect that *M* itself is constituted by various events *X* at the same level as *N*, which then cause *N*, so that all of the causation involved intralevel.

 My view is that making sense of downward causation (or more generally, interlevel causation) does not require this sort of hybrid picture. Indeed, this picture fails to make capture why we often find it so useful to employ interlevel causal claims. I will say more to motivate these claims in Section **6**.

Here I want to comment briefly on the **MM** criterion itself. One problem with this criterion is that it does not adequately distinguish part/whole or constitution relations from ordinary causal relations that are cyclic. Causal relations that involve cycles with *V1* causing *V2* which in turn causes *V1* are very common in biological, psychological and social scientific contexts, particularly when there is inter-level causation. This is because, when an upper-level variable causes a lower-level variable, often the lower-level variable will in turn feed back to affect the upper level variable (or vice-versa). For example, in the HH model, the membrane potential causally affects the channel conductances and the ionic currents and changes in these conductances/currents in turn affect the membrane potential over time—claims that as we have seen have a natural interpretation (and are readily testable) within an interventionist framework. Applying the **MM** criterion to this example, we seem forced to conclude, mistakenly, that the channels /ionic currents are constituents of the membrane potential and thus that these cannot stand in causal relationships. Similarly, as Kendler (2011), among others, notes, not only do environmental events influence gene action which in turn influences psychological states, those psychological states may in turn influence behaviors which help to create (or involve choosing to be in) environments which further influence expression of the same genes and so on. For example, environmental stressors may influence the action of genes contributing to depression which may in turn lead to behavior and choices that place the subject in environments which further accentuate gene action associated with depression. This is represented in the following diagram of Kendler’s which makes clear the cyclic nature of the process as well as the presence of interlevel causation:



[Kendler slide here: A Combined Disease Pathway: “Within the Skin” and “Outside the Skin”]

Because causal cycles are so common in the biological and social sciences, it seems to me to be a limitation in **MM** that it appears to classify cases in which cycles are present as cases involving constitution relations which are understood in a way which precludes causation, so that because the ionic channels are constituents of the cell membrane, the membrane potential cannot cause changes in those channels. A better criterion for when variables fail to be distinct in way that precludes their standing in causal relationships appeals instead to what I have elsewhere called independent fixability (**IF**): variables in set S are distinct in a way that permits their standing in causal relationships if and only it is “possible” to intervene on each variable independently, holding it fixed at each of its possible values while intervening to hold the other variables to each of their other possible values. Here “possible” includes settings of values of variables that are possible in terms of the assumed, logical, mathematical, or semantic relations among the variables as well as certain structural or space-state relationships[[12]](#footnote-12).

The variables in the HH model do meet the condition (**IF**). One can intervene to set the value of *V* independently of the values of the ionic currents *I* and one can intervene to set values of the latter via various pharmacological interventions that affect the behavior of the channels. Thus these variables can legitimately stand in causal relationships. Similarly for the other examples involving downward causation and cycles described above[[13]](#footnote-13).

 **6. Levels and Conditional Independence[[14]](#footnote-14)**.

In my remarks so far, I have objected to views about levels that tie this notion exclusively to compositional relations. Along with this, I have also attempted to respond to objections to the notion of inter-level causation that seem to be motivated by this compositional picture such as interpretations of downward causation according to which it involves a whole acting downwards on its parts. I am very aware, however, that some readers will think that I have failed to get to the heart of the matter. After all, they will say, in the case of, for example, of the HH model**,** the neuron itself is composed of atoms and molecules which interact locally, mainly through the electromagnetic force. The membrane potential difference, the channel conductances and so on must be the upshot or resultant of complex patterns of interaction among these atomic and molecular constituents. Thus variables like *V* and *I* do not represent anything “over and above” such constituents and their interactions. Why then do we need to make use of any notion of downward causation from upper-level variables? All that is “really going on” (it will be argued) involves causal interactions among lower-level variables. Talk of downward causation seems (at best) superfluous, if not positively misleading. Similarly for the other examples discussed above.

As I see it, to adequately respond to these worries we need to explain more clearly what legitimate work is done by the notions of downward (and interlevel) causation and by notions of level that are not purely compositional[[15]](#footnote-15). Here the notion I called conditional independence in Section **2** will play a crucial role.

 I begin by filling out a kind idealized or limiting case sketched in Section 2. Suppose *L* is a fine-grained variable (or set of these, although I will ignore this possibility in what follows since it does not really change anything) which is causally relevant to some explanandum *E*  characterizing system S (where causal relevance is understood in terms of **M**). Here “fine –grained” means that *L* has many different possible values or states. For example, the values of *L* might represent all of the various possible combinations (all 36 x 1023 of them) of momentum and position of each of the individual molecules making up a mole of gas. Suppose also that there is another variable *U*which also characterizes S, and which corresponds to a “coarsening” of *L.* (“Coarsening” here means that *U* is a function of *L,*  in the mathematical sense of “function”[[16]](#footnote-16), but that *U* has many fewer possible values than *L*, so that the relation between *L* and *U* is that many different values of *L* are mapped into the same value of *U* as when, in the example above, the different possible combinations of momentum and pressure are mapped into a variable like temperature —this is connected to the abstractness notion of level in **2.3**)[[17]](#footnote-17). Assume that *U* is also causally relevant to the *E* (where again this is understood in terms of **M**) and that, furthermore, *conditional* on the values of *U*, the values of *L* are irrelevant to *E*. The notion of conditionality here is to be understood in terms of interventionist counterfactuals, rather than conditional probability: if we were to fix the value of *U* via an intervention, further changes in the value of *L* also produced by interventions that are consistent with the value of *U* make no further difference to *E*. For example, given that a gas has pressure *P=p*, then (as a matter of empirical fact) it is true or very nearly true that any variation in the positions and momenta of its component molecules which are consistent with *P=p* will have no further impact on other thermodynamic variables like the temperature and volume which we take to be our target explananda *E*. In such a case it is natural to think that all of the causal information about *E* that is in the fine-grained variable *L* is absorbed into the *U* -- in this sense *U* “screens off” *L* from *E.* Thus, within an interventionist framework one can just as well use *U* as *L* in explaining *E*. There is no loss of causal or explanatory information relevant to *E* in using *U* rather than *L* despite the fact *U* is a coarser variable[[18]](#footnote-18). When, in this sort of case, *U* is taken to be at a different level than *E* (where this judgment may reflect compositional or other sorts of considerations), we may legitimately think of *U* as an interlevel cause of *E*. Thus, on this view of the matter, claims of downward causation (and claims of interlevel causation more generally) can be thought of as claims about the irrelevance of certain kinds of information conditional on other sorts of information—we can legitimately make claims of interlevel causation when such conditional irrelevance relations are present.

Here are some additional illustrations of the basic idea. Returning to the HH model, think of this as embodying the claim that, given the potential *V* across the entire membrane, any further information about how that potential is realized in the electromagnetic forces associated with individual atoms and molecules does not matter for the impact of *V* on the variables measuring the ionic currents and the channel conductances. This does not mean, of course, that such forces do not exist or are not causally operative; rather the point is that they do not *matter* for the behavior of the ionic currents and channel conductances, *given* the value of *V* or at least they do not matter for the overall shape of the action potential. Put differently, this view does not deny that (as claimed in the objection envisioned above) there are causally relevant goings on at the level of individual molecules; rather the claim is that we do not *need* to advert to such details in explaining certain facts about the channel conductances and ionic currents, given the value of *V*[[19]](#footnote-19).

Consider another example: some environmental event (famine during pregnancy, death of a parent at an early age) is claimed to causally influence gene expression via some epigenetic process— in a putative case of downward causation. One might imagine an enthusiastic reductionist who resists this claim, insisting instead that what is “really” going on is that there is some molecular level instantiation of these environmental events with molecular level interactions involving these molecules being causally responsible via some complicated chain of molecular level intermediates for the changes in gene expression. According to the reductionist, all of the real causal action occurs at the molecular level. In responding it seems to me that the defender of downward causation should not deny (as some anti-reductionists do) the obvious point that causal (and explanatory) relations are present at the molecular level that influence gene expression. Rather the defender of downward causation should resist the claim that these are the *only* causal relations that are present (as the reductionist’s use of “really” seems to insinuate). Downward or interlevel causation will be present in such cases to the extent that is true that the appropriate conditional irrelevance relations hold. Suppose it is true (as it appears to be) that, given that a mother experiences a famine level reduction in total calories in the first trimester of pregnancy, then, independently of the details of the composition of those calories or their molecular realization, her offspring have changes in gene expression resulting in an increased tendency to obesity later in life. Then reduced caloric intake will be a legitimate downward cause, which indeed is how it is usually described.

The sort of case just described, in which there is complete conditional independence is, as I have said, a limiting case, although I believe that it is not as unusual as some philosophers suppose. One might relax this requirement in various ways: lower-level variable *L* might be conditionally independent of *E*, given *U,* for “almost all” even if not literally all values of *L*. Such conditional independence might hold for those values of *L* that are most likely (as measured in terms of relative frequencies) to occur or at least most likely to occur in environments of interest. Arguably it will be legitimate to continue to talk of upper-level causation in such cases, although there will be countervailing considerations[[20]](#footnote-20).

At this point the reader may wonder whether (and when and why) it is reasonable to expect even approximate conditional independence to hold. The answer to this question is going to depend on empirical facts about the domain under investigation. In physics, for example, there are a number of results that tell us when to expect such independence and even (in some cases) why it occurs. Renormalization group type arguments show why various phenomena having to do with phase transitions are independent of facts about the details of the material composition of the systems undergoing such transitions (Batterman, Forthcoming b). Decoupling theorems in high-energy physics show that the correct physics at lower energy scales is effectively independent of the correct physics at higher energy scales. Turning to macroscopic organisms like ourselves, there are a number of general reasons for expecting that conditional independence relations (or some approximation to them) in which relatively coarse grained environmental variables screen off more fine-grained variables will sometimes hold. Roughly, this is because it is often the values of such coarse-grained variables rather than variables that make further fine-grained distinctions that are biologically meaningful for us and that we have been shaped, by natural selection and various learning processes, to be sensitive to. For example, it makes biological sense that insofar as the effect of interest is weight gain in adulthood, that this would be sensitive to an upper-level variable like maternal caloric intake rather than tracking further fine-grained variations in how this variable is realized. Of course, with respect to other effects more fine-grained dietary details may matter[[21]](#footnote-21). Similarly, although different environmental stressors may have different effects on gene expression, it is plausible that these stressors will fall into certain broad groups (e.g., those that involve loss versus threat) having similar effects based on the biological significance of the stressor and the psychological and neuroendocrine systems it engages. In other words, given that the stressor involves threat, further variation in the details of the stressor may not matter to its genetic effect, in which case the experience of threat is the downward cause. More generally, there are obvious ecological reasons why sensory systems are likely to be sensitive to the values of relatively coarse-grained environmental variables rather than fine-grained realizations of these (the presence of a tiger rather than small variations in the molecular realization of the tiger) and thus that it will be these “upper level” macroscopic variables that will drive lower-level neural, hormonal and genetic responses. Whenever this is the case, one has a good approximation to conditional independence[[22]](#footnote-22).

So far my argument has been that causes framed in terms of upper-level variables will sometimes do at least as well, for purposes of causal explanation, as causes framed in terms of lower-level variables—this shows that such claims are not superfluous. There are, moreover, several additional points that help to explain why it is often useful to operate with notions of interlevel causation.

The first is that, in many cases, explanations in terms of lower-level variables are simply not accessible or constructable, both for computational and epistemic reasons. Consider the project of replacing an explanation in terms of an upper-level variable like death of parent with a molecular level characterization of this cause. Obviously we are in no position to actually exhibit or construct such an explanation—we don’t know the molecular details of the realization of the upper-level variable on any particular occasion and even if we did, these details would be different for other episodes of parental death. Furthermore even if we did have such information, a bottom-up calculation from these details showing how they result in, say, depression, would be (to state the obvious) completely intractable. Similarly for many of the other examples of causal claims involving upper-level variables discussed above. It is thus fortunate indeed that when the appropriate conditional independence relations hold, we don’t need to appeal to such lower-level causes. We can establish the truth of upper-level causal claims and that the appropriate conditional independence relations hold by means of the usual methods of experimentation and causal modeling involving such variables without making use of “underlying” lower-level information. We thus have two-part rationale for appealing to upper-level causes and downward causation: (i) we may lose nothing by doing so in terms of relevant difference-making information and (ii) it may be impossible to actually construct or exhibit explanations in terms of lower level causes. Note that even if (ii) is regarded as a “merely pragmatic” consideration, (i) is not. Whether conditional independence or something close to it holds depends on what nature is like, not on our calculational and epistemic limitations or other pragmatic considerations. These are among the considerations that show why appeals to upper-level variables as causes (including causes of lower level variables) is not “superfluous” or a product of some sort of confusion.

Let us now return, in the light of these observations, to Bechtel and Craver’s proposal that interlevel casual claims should be understood as hybrid claims that invoke both interlevel constitution relations and intralevel causal relations. As noted above, we often lack information about the compositional and interlevel relations that according to Bechtel and Craver underlie interlevel causal claims. Focusing for the moment on claims of downward causation, we often don’t know the necessary details concerning the lower-level constituents of the upper-level causes, or what the laws or causal relations governing the lower-level variables are. Thus their account seems to hide the very reason that we appeal to downward casual claims in the first place: we do so to a substantial extent because we don’t have the very information on which their account rests. Put differently, appeals to downward causation are to a considerable extent part of a strategy of avoiding modeling unknown lower details – the very details to which Bechtel and Craver appeal.

Finally, let me return to the question posed in Section **1** (and in the title of this paper) concerning the function of level talk—what legitimate work does it do? In addition to the role played by compositional considerations, we can think of claims about levels as encoding information about what factors it is permissible to ignore in modeling and causal analysis – either because these factors are unconditionally irrelevant to some effect of interest or (more commonly) because, they are irrelevant, conditional on information represented by other variables.

**References**

Batterman, R. (Forthcoming a) “Multiscale Modeling in Inactive and Active Materials”

Batterman, R. (Forthcoming b) “Universality and RG Explanations,” *Perspectives on Science.*

Bechtel, W and Craver, C (2007) “Top-down Causation Without Top-Down Causes” *Biology and Philosophy* 22:547–563.

Chalupka, K. Eberhardt, F., and Perona, P. (2017) “Causal Feature Learning: an Overview” *Behaviormetrika* 44:137–164.

Craver, C. (2007) *Explaining the Brain*. Oxford: Oxford University Press.

Eronen, M. I. (2015). “Levels of Organization: A Deflationary Account” *Biology and Philosophy* 30 (1):39-58.

Glymour, M. , Veling, W. and Susser, E. (2011) “Integrating Knowledge of Genetic and Environmental Pathways to Complete the Developmental Map” in Kendler, K., Jaffee, S. and Romer, D. *The Dynamic Genome and Mental Health*. New York: Oxford University Press.

Kendler, K. and Campbell, J. (2009) “Interventionist causal models in psychiatry: repositioning the mind–body problem” *Psychological Medicine* 39(6):881-7.

Kendler, K. (2011) “A Conceptual Overview of Gene-Environment Interaction and Correlation in a Developmental Context”. In Kendler, K., Jaffee, S. and Romer, D. , *The Dynamic Genome and Mental Health*. New York: Oxford University Press.

Heil, J. (2017) “ Downward Causation” In Paoletti, M. and Orilia , F (eds.) *Philosophical and Scientific Perspectives on Downward Causation*. New York: Routledge, 42-53.

Lewis D. (2000). “Causation as influence”. Reprinted in Collins J., Hall N. and Paul L. (eds), *Causation and Counterfactuals*. Cambridge: MIT Press.

 Noble, D. (2006) *The Music of Life.* Oxford: Oxford University Press.

 Oppenheim, P., and Putnam, H. (1958). “The Unity of Science as a Working Hypothesis” . In Feigl, H., Scriven, M. and Maxwell, G. (Eds.). *Concepts, theories, and the Mind-Body Problem.*  Minneapolis: University of Minnesota Press, 3-36.

Wimsatt, W. (1994) “The Ontology of Complex Systems: Levels of Organization, Perspectives, and Causal Thickets”, in *Biology and Society: Reflections on Methodology*, M. Matthen and R. Ware, *Canadian Journal of Philosophy*, supplementary volume 20: 207–274.

Wilson, M. (2018) *Physics Avoidance: Essays in Conceptual Strategy*. Oxford: Oxford University Press.

Woodward, J. (2015) “Interventionism and Causal Exclusion” *Philosophy and Phenomenological Research* 91: 303- 347.

Woodward, J. (2016) “The Problem of Variable Choice” *Synthese* 193:1047-1072

Woodward, J. (Forthcoming) "Explanatory Autonomy: The Role of Proportionality, Stability, and Conditional Irrelevance". *Synthese*

1. Other, more complex forms of coarsening are discussed in Batterman, forthcoming b. [↑](#footnote-ref-1)
2. In this case, the interactionist and compositional notions of level lead to the same judgments since it is also the case that the gas is composed of molecules, hence that variables characterizing the gas a whole seem (on a composition-based notion) at a higher level than individual molecules. But as noted below, level assignments based on interactionist and compositional ideas do not always coincide. [↑](#footnote-ref-2)
3. This general theme of looking for ways of modeling certain features of a system which allow the modeler to avoid having to model other features in any detail is explored in great depth in Wilson (2018). [↑](#footnote-ref-3)
4. Alon (2007) provides illustrations of important differences in temporal scale in biological context: inputs change the activities of gene transcription factors on a sub-second scale, in contrast, “binding of the active transcription factor to its DNA site reaches equilibrium in seconds. Transcription and translation of the target gene takes minutes and the accumulation of the protein product can take many minutes to hours” (p.10-11). [↑](#footnote-ref-4)
5. Of course, as noted above, sometimes composition-based and interactionist notions of level will coincide or track one another. My point is that this is by no means always the case. [↑](#footnote-ref-5)
6. An illustration: My former colleague David Goodstein, a solid state physicist at Caltech, recounts the reaction of physicists to the claims of the chemists Pons and Fleischmann to have discovered cold fusion. Goodstein observes that one reason why most physicists disbelieved those claims (even in the absence of detailed calculations or failed attempts at replication) was based on a heuristic application of separation of scale/ level considerations. The energy supplied in the apparatus these researchers were using was many orders of magnitude too small to produce fusion and the distance relations between atoms that could be achieved by the electrolytic reactions the chemists were employing were many orders of magnitude greater than the distances required to achieve fusion. The processes involved in fusion occur on distance scales and on time scales that are effectively completely separate from anything that might be involved in the kind of experiment the chemists conducted. [↑](#footnote-ref-6)
7. Wimsatt (1994) speculates, plausibly, that what he calls “causal leakage” (that is causal influence across levels and causal explanations employing causes at different levels) becomes more common as one moves to phenomena at “higher levels”—it is more common in, say, psychology than in high-energy physics. [↑](#footnote-ref-7)
8. Note that an environmental stressor like “death of a parent at an early age” seems to be well-defined and measurable only at the level of whole organisms, rather than at, say, the level of individual genes, even though the environmental event can influence gene expression. [↑](#footnote-ref-8)
9. This is not to deny that it is of great interest to discover the mediating variables by which environmental events affect gene expression: they will often do so via the organism’s sensory system which will in turn affect neural processing, hormone and neurotransmitter levels and so on. These mediating processes will be “lower-level” but this is compatible with the initial triggering environmental events often being best conceptualized as at a higher level—e.g. as “stress” or perhaps a particular kind of stressor. [↑](#footnote-ref-9)
10. For a defense of this claim within psychiatric contexts see Kendler and Campbell, 2009. [↑](#footnote-ref-10)
11. The idea that a causal relationship requires that the relata of that relationship be “distinct” in some appropriate way is a common place of the philosophical literature—see, for example, Lewis, 2000. One of his examples is that saying “hello” loudly cannot cause one to say, “hello”, because saying “hello” loudly logically entails saying “hello” and this sort of entailment implies that the events in question are not distinct. I agree that causation requires that the causal relata satisfy some appropriate distinctness requirement but, for reasons discussed below, do not think that standard claims of downward causation in the scientific literature violate a reasonable version of such a requirement. [↑](#footnote-ref-11)
12. For further discussion see Woodward, 2015. It is worth explicitly noting the difference between (IF) and (MM). (MM) has to do, roughly, with whether it is true both that Y changes under some intervention on X and that X changes under some intervention on Y (where X and Y correspond to the behaviors F and J in (MM). By contrast, IF says distinctness fails when it is impossible to intervene to set X and to set Y to some combination of values. [↑](#footnote-ref-12)
13. There is a tendency in philosophical discussion to suppose that causal representations with cycles are always illegitimate—indeed another part of Bechtel and Craver’s objection to downward causation is that this often involves causal cycles which they take to be objectionable. To the extent that this is a worry, cyclic representations can sometimes be replaced with acyclic representations involving time-indexed variables: X at time t causes Y at time t+1 which causes X at time t+2. Although I lack the space for detailed discussion, I believe, however, that there are legitimate cases involving cycles -- e.g., models which describe certain kinds of equilibrium relations) for which this sort of interpretive move is not available. Such models nonetheless can have a straightforward interventionist interpretation. [↑](#footnote-ref-13)
14. The ideas in this section regarding conditional independence are in some respects similar to and have been influenced by ideas developed much more formally by Frederick Ebehardt and colleagues in a machine learning context—see, e.g., Chalupka et al. (2017). I am indebted to Eberhardt for helpful discussion. [↑](#footnote-ref-14)
15. I see this as part of a general program of elucidating important notions in science by spelling out what their legitimate function is or what we aim to accomplish in using them. [↑](#footnote-ref-15)
16. That is, each value of L is always mapped into just one value of U. [↑](#footnote-ref-16)
17. In other words, although F is a function, it is not 1-1 or injective. [↑](#footnote-ref-17)
18. In other words, although L contains more information than U, U contains all the information that is relevant to E in L. [↑](#footnote-ref-18)
19. A number of defenses of upper-level causation (or explanation) in the philosophical literature contend that, when true, the upper-level claims rule out the truth of lower-level causal claims, so that there is “downward exclusion”. I reject such views. [↑](#footnote-ref-19)
20. The countervailing considerations are that as we relax the conditional independent requirement in the ways described, we admit downward causes that have more and more heterogenous effects. At some point this heterogeneity becomes sufficiently great that we don’t have well-defined effects or well defined responses to interventions on the candidate cause. [↑](#footnote-ref-20)
21. Glymour et al. (2011) argue that folate deficiency in the diets of mothers during the Dutch hunger winter is partially responsible for increased the increased incidence of neural tube defects and schizophrenia in their offspring. If this is correct, then, even conditional on total caloric intake, dietary folate levels, are relevant to NTDs so that the downward cause is folate deficiency rather than calorie deficiency. [↑](#footnote-ref-21)
22. Put differently many biological systems operate by neglecting or abstracting away from certain kinds of detail or variation—think of systems that have a binary response to continuous input variables etc. [↑](#footnote-ref-22)