Forthcoming in BJPS; final version available via AdvancedAccess as of 12/18/2014

High-Level Explanation and the Interventionist's 'Variables Problem'[1]

L. R. Franklin-Hall

lrf217@nyu.edu

abstract

The interventionist account of causal explanation, in the version presented by Jim Woodward ([2003]), has been recently claimed capable of buttressing the widely felt—though poorly understood—hunch that high-level, relatively abstract explanations, of the sort provided by sciences like biology, psychology and economics, are in some cases explanatorily optimal. It is the aim of this paper to show that this is mistaken. Due to a lack of effective constraints on the causal variables at the heart of the interventionist causal-explanatory scheme, as presently formulated it is either unable to prefer high-level explanations to low, or systematically overshoots, recommending explanations at so high of a level as to be virtually vacuous.

1 Introduction

Though our world is a physical one, many of the best explanations for events within it mention exclusively 'high-level' features: the anger of the criminal, the action of natural selection, the rising of interest rates. The interventionist account of causal explanation, as articulated and developed by Jim Woodward ([2003], [2008a], [2008b], [2008d], [2010], [2011]), promises to make at least some sense of the special

[1] For helpful comments on this paper, many thanks to Christophe Malaterre, Tim Maudlin, Alex Reutlinger, Michael Strevens, Jim Woodward, an audience at Temple University, a reading group at University of Sydney, and my referees.

value of these high-level explanations, and its popularity among philosophers of the high-level sciences might indicate that it succeeded in doing so.[2]

On Woodward's view, to explain an event or outcome is, at minimum, to provide "information about the causes of that outcome"(Woodward [2010]: 291). Causes are then understood along interventionist lines: two features are causally related just when, given some background circumstance, there is a possible intervention—loosely, an ideal or surgical manipulation—on the state of one feature that changes that of the other. Since high-level features are just as able as fine-grained physical ones to satisfy this condition, interventionism straight-forwardly secures the causal and explanatory relevance of anything you might care about, from the biological, to the psychological, to the social scientific.

This account can sustain the judgment that high-level explanations are not completely mistaken. But can it go any further than this, showing that high-level explanations, of the sort provided by high-level scientists, are, at least sometimes, explanatorily superior? On early formulations of the interventionist explanatory theory, such as those provided by Woodward (2003) and Hitchcock and Woodward (2003), the answer to this question would appear to be *no*. According to the basic interventionist picture causal information is explanatory precisely because it can be used to answer *what-if-things-had-been-different* questions (w-questions). Though information about any interventionist cause will answer some w-questions, the best explanations, those deemed "deep and powerful," will answer the most. Judged by this standard of excellence, high-level explanations are uniformly impoverished; they explicitly represent fewer features of the world on which the explanandum depends than do lower-level "micro" explanations, limiting the range of w-questions they can answer.

Were this all that could be said, interventionism's gift to the high-level would be but thin; it is not alone among explanatory accounts in its ability to make sense of high-level explanations that are invariably *less explanatory* than micro-physical ones (e.g. Railton [1981]). Further, it would leave interventionism inferior—at least by high-level standards—to competing causal-explanatory accounts that are able to sustain the objective superiority of at least *some* high-level explanations (e.g., Strevens [2008a]).

Yet comments in a recent series of papers by Jim Woodward ([2008a], [2008b], [2008d], [2010], [2011]) suggest that the interventionist verdict on high-level explanation is neither so simple as this, nor so dire. Woodward claims that there are a least some circumstances in which "upper level causal claims provide better explanations than lower level claims"(Woodward [2008d]: 210; see also Woodward

[2]Recent work using interventionist high-level causal and explanatory accounts includes: Baedke [2012], Campbell [2007, 2008], Craver [2007], Glennan [2005], Malaterre [2011], Reisman and Forber [2005], Reutlinger and Koch [2008], Reutlinger [2011], Stegmann [2012], and Waters [2007].

[2010]). Dominating his discussions of high-level superiority is the requirement of causal *proportionality* (sometimes labeled *causal fit*). This requirement—which is strictly speaking a supplement to the version of explanatory interventionism formulated in Woodward (2003)—holds that good explanations cite causes that are 'just enough' for their effects; advocates claim that high-level factors, at least at times, uniquely satisfy this requirement.

Were this strategy effective, explanatory interventionism's popularity would be well deserved. After all, making sense of the explanatory—not merely the practical—superiority of high-level explanations in a physical world has been a kind of Holy Grail in the philosophy of science, long sought but never found.[3] In light of this, the aim of this paper is to scrutinize interventionism's high-level qualifications by exploring proportionality's prospects.

That exploration unfolds as follows. I start by describing relevant aspects of the interventionist causal-explanatory account (sections 2 and 3). Then, in three steps (sections 4 through 6), I argue that, at least as presently formulated, interventionism's high-level qualifications are poor. More specifically, section 4 takes seriously the *letter* of the proportionality standard—the precise definition given it—and shows it impotent to choose between levels and thus unable to prioritize high-level explanations. Next, section 5 offers an interpretation of the *spirit* of the proportionality standard, one that advocates may have in mind without stating outright, which adds to the letter of proportionality the requirement that explanatory causes exhaust the causal 'possibility space.' This suggestion proves not impotent, but rather *too* potent: it recommends explanations at so high of a level as to be nearly vacuous, and very unlike the explanations actually on offer in the high-level sciences. Then, section 6 explores a modest and interventionist-friendly fix: a requirement that explanations cite causal factors that best balance exhaustivity and the distinct explanatory virtue of stability, while also satisfying the more technical requirement of proportionality. This move proves likewise unable to recommend the intuitively satisfying high-level explanations of the kind scientists actually articulate.

A single argumentative strategy will be deployed throughout these three steps and is worth highlighting at the outset: to identify a *variable*—the causal relatum, on the interventionist's view—that can satisfy a proposed explanatory requirement while yielding an explanation that is not acceptable as judged by consistency with actual scientific-explanatory practice. The repeated success of this strategy reflects what I submit is the core source of interventionism's causal-explanatory shortcomings: its excessive ecumenism with respect to these causally related variables. If this is correct, interventionists must introduce more substantive constraints on good variables if

[3] Philosophers who have judged high-level explanations to be at least sometimes explanatorily superior include: Putnam [1979], Kitcher [1981, 1984], Garfinkel [1981], Batterman [2002], and Strevens [2008a].

they hope to 'save the explanatory phenomena.' Thus the paper concludes by distinguishing two species of constraint that interventionists might adopt, constraints appealed to by alternative causal-explanatory accounts. This reflects the fact that the 'variables problem' highlighted here is not actually one for interventionists alone; it is a broader challenge that, in various guises, every causal-explanatory theory is obliged to address.

2 The Interventionist Picture

2.1 Interventionist Type Causation

Though every account of causation will find some connection between causal relationships and those that are "potentially exploitable for the purposes of manipulation and control"(Woodward [2003], p. 25), causal interventionism is unique in taking such exploitability to be constitutive of causation, and not simply an indication or consequence of it. As Woodward puts it, interventionism "is intended as a characterization of what it *is* for X to cause Y. It is not claimed (and it is indeed false) that the only way to tell whether X causes Y is to experimentally intervene on X and see what happens to Y" ([2008a], p. 215). In particular, X causes Y, at the type level, just in case, in some background circumstance, *it is (in principle) possible to change Y by intervening on X*. For example, to say that tickling causes laughter is to say that—in some background circumstance, such as when candidate systems are neither extremely angry nor comatose—it is possible to change (the presence or intensity of) laughter by intervening on (the presence or intensity of) tickling.

Precisely what this view comes to depends on what it means *to intervene*. Interventions are causal manipulations—though ones we need not be able to actually carry out—that satisfy conditions that Woodward [2003] presents via a series of inter-linked definitions. Since these definitions are complicated in ways irrelevant to the coming critique, I make do with an intuitive sketch.[4] Woodward asks us to think about interventions as "idealized experimental manipulations" ([2003], p. 94), manipulations whose outcomes scientists aim to learn about through randomized experiments. Less practically, they can be pictured as events in which the 'hand of God' comes down and alters the value of X directly. Whichever image is preferred, what matters is that X causes Y just in case some non-confounded manipulation of X would change Y; a non-confounded manipulation is one in which Y is changed, if at all, via the contribution of X, and not via some other route.

As indicated already, most important for the argument of this essay will be the features—represented above with X and Y—that are causally related on the

[4] For more details, see Woodward's account of interventions in chapter 3 of his [2003], an account later discussed and revised in an exchange between Woodward [2008d] and Strevens [2007, 2008b]. Critical discussion of the notion of an intervention can be found in Baumgartner [2012] and Reutlinger [2012].

interventionist view. Some accounts of causation—either type or token—attempt to incorporate substantive commitments about these causal relata, insisting, for example, that they are (among other options) *events* (Davidson [1967]) or *situations* (Menzies [1989]), which might themselves need to satisfy some 'naturalness' condition (as in Lewis [1983]). Sharply contrasting with such constrained approaches, the interventionist account of causation—at least as it has been developed to date—remains maximally open about what sorts of things might be related as cause and effect, suggesting that the relata are *variables*. These variables are simply anything—events, features, properties, etc.—that can vary, and are neither required to satisfy a naturalness condition, nor anything else that is in any way "metaphysically portentous"(Woodward [2008b]: 231). The one restriction that variables must satisfy is that they be "capable of at least two different 'values"'(Woodward [2010]: 290; see also [2003]: 111), meaning that they have at least two different settings or states.

This account of causation is one that, along a number of dimensions, "allows a relationship to qualify as causal even if it lacks features thought by some to be characteristic of causal relationships"(Woodward [2010]: 290). First, in its openness to causation between all variables that can take at least two values, the set of features that might be causally related for interventionists is enormous, neither excluding high-level factors—such as mental states—nor those that are, from an intuitive point of view, gerrymandered. Second, the account is undemanding in virtue of its use of 'some,' rather than 'many' or 'all,' in its basic causal condition, which can be rephrased for clarity as follows: X is a type-level cause of Y just in case *some* possible intervention on X changes Y in *some* background circumstances. In line with this, X and Y may be causally related even though only one sort of intervention on X changes Y, and in only one particular background condition.

2.2 Interventionist Actual Causation and Event Explanation

On the interventionist view, to explain an event is—most centrally—to cite one or more of its *actual* or *token* causes. Thus, I offer the interventionist accounts of actual causation and event explanation simultaneously, labeling them in explanatory terms.

Woodward (2003) can be taken to offer two related discussions of event explanation. One of these is simpler than, and a special case of, the other. In what I call a *simple event explanation*, an explanation will cite, among other things, just one type-level causal claim (p. 203). In what I call a *complex event explanation*, an explanation includes information about a possibly extensive network of type-level causal relationships (pp. 74 - 86), and a procedure is offered to extract the actual cause of the target event from the network of causal relationships represented. The complex case is fascinating, and a discussion of it must form the heart of analyses of

actual causation problem cases, such as preemption scenarios.[5] Yet here I will almost exclusively consider simple event explanations, for two primary reasons: first, Woodward's own discussions of proper explanatory level—those that this paper aims to evaluate—themselves focus on the proper 'level' of alternative simple event explanations; second—and more importantly—this focus is apt given my present project: as I will explain in due course, it is only in application to *simple event explanations* that proposed explanatory standards, like proportionality, show any prospect of preferring high-level explanations to low.

The explanandum in an interventionist event explanation is a variable Y taking a particular value y_1, as instantiated in a particular system at a particular time. The simple event explanation of a case of *Y taking* y_1 has two parts—one general and one particular:

1) a true statement of an interventionist-approved type-level causal relationship (G) relating X and Y. G, minimally, must be such that an intervention on the cause variable, X, from one value (x_1) to another (x_2), changes Y from one value (y_1) to another (y_2).

2) a true statement that some cause variable, X, in the particular circumstance, took the value x_1 (Woodward 2003: 203).

An example can illustrate the interventionist approach to simple event explanation. As do many humans, Andrew sneezes when exposed to direct sunlight.[6] He just so sneezed, and this event is the target of my explanation. The explanandum, in interventionist terms, states that the sneeze variable took *sneeze*. Its explanans must include two elements, as noted above: 1) G, a type-level causal claim, true of Andrew, connecting sunlight exposure to sneezing,[7] and 2) the cause, that Andrew was *exposed to direct sunlight*. These combine to constitute a minimally adequate explanation for Andrew's sneeze.

3 Explanatory Comparisons

According to the interventionist explanatory account just sketched, any event will possess an exceedingly large number of minimally adequate simple event

[5] For helpful discussions of interventionist-style approaches to preemption scenarios, see Halpern and Hitchcock [2011] and Strevens [2007, §2].

[6] This is a consequence of the poorly understood photic sneeze reflex (also known as the 'sun reflex' and the 'ACHOO syndrome'), found in approximately one-fourth of human subjects. See Langer et al. [2010].

[7] For sunlight exposure and sneezing to be so connected on the interventionist analysis, it must be true that, in some background circumstance, some intervention on Andrew's exposure to direct sunlight would change his sneezing state.

explanations. These can vary along two primary dimensions: *horizontally* and *vertically*. The divide between these dimensions may be understood in terms of the 'distinctness' of the cause variables cited. Two explanations for a given event that differ horizontally cite *distinct* variables, while those that differ vertically cite *constrained* variables.[8] Let me explain these two notions.

Variables A and B are *distinct* just in case the values of the variables are capable of varying independently, meaning that there are no logical or metaphysical ties between their values. (Distinct variables may still be related causally or nomologically.) For the sake of expository simplicity, take both A and B to be binary, with A having two values, a_1 and a_2, and B two values, b_1 and b_2. In that case, the following four variable value combinations must be logically and metaphysically possible if these variables are to be distinct: $(a_1, b_1),(a_1, b_2),(a_2, b_1),(a_2, b_2)$.

In illustration of how distinct variables capture horizontally related causes of some event, consider two causal variables that may be cited in explanation of a car crash: Road State, which can take *wet* or *dry*; and Driver Emotional State, which can take *angry* or *serene*. One explanation cites the wet road; another cites the anger of the driver. Both of these, in concert with the appropriate type-level causal claims, form minimally adequate explanations, since—I stipulate—both the state of the road and the emotional state of the driver were interventionist causes of the crash: had the driver been serene, the crash would not have happened; similarly, had the road been dry, the driver would have, irrespective of any distraction-inducing emotional turmoil, cruised safely into the night. These causes are related horizontally because the cause variables are distinct; the driver might have been either angry or serene, while the road was either wet or dry.

Alternative causes, and causal explanations, of some event differ vertically—varying in 'level'—when they cite variables that are *constrained*, not distinct. Rather than representing independent difference-makers—such as emotional vs. road state—these variables track the same features of the world, but describe them differently. In the case of simple event explanations at different levels, the values of the high-level variable will be coarser than those of the low-level variable. Equivalently, each value of the low-level variable will be a determinate of some value of the high-level variable, its determinable. Such is the relationship between cause variables cited in the following two explanations of a crash: first, a low-level explanation that appealed to a binary variable, one of whose values was the exact speed of the car as it turned a bend, 50 mph, and another value representing some speed at which the car would not have crashed, such as 20 mph; second, a high-level

[8] Two explanations of a given event may also vary temporally, as when one explanation cites a factor that occurs earlier in the same causal chain leading to the target event than does a second explanation. Since I will characterize distinctness in terms of logical and metaphysical relations only, temporally related explanations will, along with horizontally related explanations, cite distinct variables.

explanation that appealed to another binary variable, with one value corresponding to the car's speed exceeding 30 mph, and the other value representing that it did not. The cause variables cited in these two explanations are constrained because their values cannot—in a very strong sense, and not simply because of nomic ties between their values—vary independently: the car could not have both traveled at 50 mph and not over 30 mph; neither could it have both traveled at 20 mph and, at the same time, over 30 mph.

Though there are closely connected questions about what makes alternative horizontal explanations for the same event—such as environmental or genetic causes of a phenotypic character—superior or inferior (see Franklin-Hall [forthcoming-b]; Waters [2007]), here I focus on vertical comparisons, as they more directly concern the problem of explanatory level.[9] Furthermore, the main discussion of explanatory level will itself be limited to evaluating relationships structurally similar to the car speed example just described, those relating multiple simple event explanations. These explanations vary exclusively in the coarseness or fineness of the cause variable cited, differing along what I will call the *coarse-fine* axis.

In considering only coarse-fine comparisons, I sidestep other kinds of 'leveled' explanatory comparison, such as those in which candidate explanations vary along what I will call the *macro-micro* axis. When two explanations vary in this way, the higher-level—or macro—explanation is a simple event explanation, but the lower-level—or micro—one is an instance of I earlier labeled a *complex event explanation*, possessing more content than one causal claim and a statement that the cause variable took a particular value. More specifically, the micro explanation will, roughly speaking, describe the mechanism underpinning the single causal claim featured in the paired macro explanation. For instance, consider two explanations of the acceleration of a car: 1) a micro explanation that represented both the depression of the gas petal, and the states of other individual car parts, along with causal connections between them, and 2) a macro explanation that black-boxed the engine's internal functioning, and exclusively described its behavior in terms of a system-wide causal input-output function, one connecting petal depression and acceleration.

There are two rationales for my focus on coarse-fine comparisons, to the exclusion of the macro-micro. First, in comparing macro and micro explanations for the same event, interventionism uncontroversially holds out no hope of prioritizing the macro-level. This is because well-formulated micro explanations—those that open up all black boxes and represent the workings of their innards—will always be more "deep and powerful," answering more w-questions, than do their macro cousins (for discussion, see Hitchcock and Woodward [2003]; Weslake [2010]). After

[9] As noted by Potochnik [2010], many actual alternative scientific explanations vary along both horizontal and vertical axes simultaneously. Thus, pure constraint and pure distinctness of variables should be seen as extreme cases on what is really a complex continuum.

all, in addition to answering all the questions that its paired macro explanation can, a micro explanation will be able to answer questions about what would have happened had the causal relationship cited in the macro explanation itself been abrogated, ceasing to hold. Second, the particular strategy that has been offered to show interventionism capable of prioritizing high-level explanations over low—proportionality—has been formulated in a way that only applies, at least in any straight-forward way, to explanations differing along the coarse-fine axis. Thus, an evaluation of interventionism's high-level prospects requires no wider gaze.

4 Proportionality to the Rescue?

Proportionality, or *causal fit*, is the central consideration that had been offered in defense of the superiority of high-level interventionist explanations.[10] Put informally, proportionality asserts that, other things equal, an explanation is superior when it cites a cause that is *just specific enough* to bring about its target effect. In being *just enough*—and no more than that—the hope is that proportional causes will be comparatively abstract and high-level, omitting details that don't make a difference to the effect to be explained. In this way, proportionality promises to make sense of the intuitive superiority of at least some of the explanations provided in the high-level sciences.

Woodward has adapted proportionality to the interventionist framework as a constraint on optimally explanatory causal variables, suggesting that, other things equal, "causal variables [should be] 'proportional to' their effects"(Woodward [2008a]: 239).[11] The idea is that while there are many causal variables that might *in principle* be cited in a minimally adequate interventionist explanation of some event, proportionality constrains which variables are chosen, ruling out some otherwise viable candidates that are at the wrong level, thus providing "some guidance on the choice of variable set or appropriate level"(ibid: 228).

In particular, the proportionality of a type-level causal claim G in terms of appropriate variables is characterized as follows:

> (P) There is a pattern of systematic counterfactual dependence (with the dependence understood along interventionist lines) between different possible states of the cause and the different possible states of the effect, where this pattern of dependence at least approximates to the following ideal: the dependence (and the associated characterization of the cause) should be

[10] The idea of proportionality is from Yablo [1992] and the interventionist application of it is most thoroughly explained in Woodward's [2008a], [2008b], and [2008d] and (especially) [2010]. It is also exploited in influential work on mechanistic explanation, such as Craver [2007].

[11] Malaterre [2011] also interprets proportionality as offering a constraint on explanatorily appropriate variables, though he is more optimistic about its efficacy.

> such that (a) it explicitly or implicitly conveys accurate information about the conditions under which alternative states of the effect will be realized and (b) it conveys only such information—that is, the cause is not characterized in such a way that alternative states of it fail to be associated with changes in the effect. (Woodward 2010: 298; see also Woodward 2008b: 234)

The first clause of (P) insists simply that G be a true interventionist type-level causal claim. (a) and (b) go beyond this by requiring that *only* and *all* changes in the value of the cause variable in G (via interventions) change the value of the effect variable. The idea is to rule out *cause variable-effect variable* pairings according to which the cause variable is able to take alternative values but in which interventions to those alternative values would not invariably change the value of the effect variable, as well as to make sure all alternative values of the effect variable are hooked up to alternative values of the cause variable. Put most succinctly, proportionality requires that cause and effect variables that constitute a type-level causal claim be such that there is a one-to-one relationship between alternative values of the cause variable and alternative values of the effect variable. It is just this one-to-one relationship that makes the cause and effect, intuitively speaking, *proportional*.

A simple example, originally from Yablo ([1992]) and adapted by Woodward ([2010]) and Craver ([2007]), can illustrate the task that proportionality must address and will help in the evaluation of the solution it offers. The example involves an explanandum event and a pair of minimally adequate explanations for it that differ in 'level,' related in a way analogous to the vertically varying explanations of the car crash considered earlier:

> A pigeon is trained to peck at a target when and only when presented with a stimulus of any shade of red. Suppose, on some particular occasion . . . the pigeon is presented with a particular shade of scarlet and pecks at the target. Consider the following two causal claims/causal explanations:
>
> (1): The *presentation of a scarlet target* caused the pigeon to *peck*.
> (2): The *presentation of a red target* caused the pigeon to *peck*.
>
> (Woodward 2010: 297)

Based on proportionality considerations, Woodward judges that (2), the "higher-level" option, "furnishes a better explanation"([2010]: 298) than does (1), and that proportionality "identifies the red rather than the scarlet color of the target as the appropriate level of description"(Woodward [2008a]: 234). Though this evaluation is made of a toy example, it points to a general characteristic of actual explanatory practice, and thus the task of making sense of and justifying the judgment is a

prudent project. Yet in advance of scrutinizing proportionality's ability to provide such a justification, the above explanations—(1) and (2)—must be clarified, something not explicitly done in most presentations of the standard. After all, the explanations sketched in this passage stand unfinished in two respects: 1) the other value or values that the effect variable can take is left unstated; 2) the other value or values that the cause variables can take is left unstated.

Guiding the clarifications I offer is the need to ensure that (1) and (2) satisfy the basic interventionist causal condition, as well as having some promise of satisfying the proportionality constraint. This is to proceed differently than is customary in explorations of proportionality—such as those found in Woodward ([2008d], [2010]) and Craver ([2007])—which implicitly complete competing explanations at different levels such that the 'lower-level' alternative falls immediately short. For instance, when the alternative cause variable value in (1) is taken to be 'not-scarlet,' (1) is deficient in light of the fact that some interventions on the cause variable—in particular, any intervention setting it to a non-scarlet shade of red—would not change the effect variable value to non-pecking. In this way, the explanation appears to fail to satisfy the basic interventionist causal condition by saying something *false* about what would happen were a non-scarlet shade of red shown to the pigeon.[12]

In focusing on such non-starters, one is prevented from evaluating proportionality's prospects for selecting an optimally leveled explanation from the many minimally adequate ones that are available. In light of this, I fill the lacunae in these candidate explanations as follows:

(1*)—The *presentation of a scarlet target* (other value: *presentation of a cyan target*) causes the pigeon to *peck* (other value: *not peck*).

(2*)—The *presentation of a red target* (other value: *presentation of a non-red target*) caused the pigeon to *peck* (other value: *not peck*).

As desired, this completion makes both explanatory alternatives—(1*) and (2*)—minimally adequate interventionist explanations. (1*) is adequate because an intervention on its cause variable, from scarlet presentation to cyan, would change the value of the effect variable from peck to not peck. (The pigeon has not been trained to peck when it sees cyan.) Similarly, a change in the value of the cause variable in (2*), from red presentation to non-red presentation, would also change the value of the effect variable from peck to not peck. (The pigeon has not been trained to peck when it sees non-red colors.) In consequence, both are good

[12] Alternatively, one might diagnose the failure of the scarlet/non-scarlet account by pointing to its failing to satisfy clause (a) of condition (P), as that clause requires that an explanation convey "accurate information about the conditions under which alternative states of the effect will be realized." Either way, the scarlet/non-scarlet variable is judged explanatorily sub-optimal.

explanatory causes of the explanandum event. Furthermore, on the coarse-fine approach to explanatory level presumed here, (1*) is at a lower-level than (2*); *scarlet* presentation is fine-grained or low-level in contrast to presentation of *red*; the same goes for the relationship between *cyan* and *non-red* presentation.[13]

With the explanatory options clarified, the central question: can proportionality distinguish between these candidate explanations for pigeon pecking, buttressing the intuitive judgment in favor of the higher-level (2*)? Perhaps surprisingly in light of the commentary surrounding it, it appears not: the letter of the proportionality standard, as defined by condition (P), is toothless with respect to this choice of explanatory level, possessing no capacity to prefer high-level explanations over low-level ones (or the reverse). After all, there are one-to-one relationships between values of both of the cause variables and the effect variable in the candidate explanations above, since all variables are binary (peck vs. not peck; scarlet vs. cyan; red vs. non-red). Further, the values of these variables line up just as required by the interventionist explanatory account. Interventions setting the cause variables to their respective values would change the effect variables as required.

This verdict in no way reflects an unfortunate choice of examples. Other illustrations of proportionality have been offered—some far more scientifically nuanced and rich with complexity—but they can all be accommodated in the same way. In general, no matter the event that is the target of the explanation, there will always be a low-level variable satisfying the letter of the proportionality standard. The actual cause value of this low-level variable will be the value that the variable took in the run-up to the explanandum event. Since any particular run-up is maximally concrete, this value can be as fine-grained as desired. The other value for the low-level variable may then be any other fine-grained value, such that, were an intervention to change the cause variable to this value, the effect variable would take its other value. There will always be such a value so long as the event to be explained depends on some feature of the world that is a possible target of intervention. And if it didn't so depend, that event wouldn't be a proper target for *any* interventionist explanation, no matter the 'level.'[14]

---

[13] One might question the wisdom of using an example that appeals to a property as conceptually vexed as color, even if it is the most discussed case in the literature on proportionality. Fortunately, the relevant features of the example do not depend on any of color's problematic features. To see this, consider an equivalent example that appealed to the more innocuous *mass*: a pigeon begins to peck when a 30 gram mass is placed on its back; what best explains this pecking event, 1) that a particular mass, 30 grams (rather than 10 grams), was placed on its back (in concert with the appropriate type-level causal claims linking mass placement to pecking)? or 2) that the mass placed on its back was 20 grams or above, the minimum at which it would peck (vs. under 20 grams)?

[14] See Shapiro and Sober [2012] for likeminded concerns about the power of proportionality. These authors, however, link the failure of proportionality to the contrastive aspect of counterfactual accounts of causation. I have side-stepped the thorny problem of whether the interventionist account is genuinely contrastive—both with respect to causes and effects—because the failure of proportionality can be evidenced even without that feature.

5 Exhaustivity to the Rescue?

If proportionality so straightforwardly fails to distinguish between explanatory level, why might anyone have thought otherwise? Perhaps fans of proportionality have excluded from the explanatory competition, either intentionally or otherwise, variables like the low-level, scarlet-cyan option considered above. With these out of the running, proportionality *is* able to prefer high-level explanations to low. This is because more orthodox low-level variables are many- or even continuum-valued. In the pecking example, for instance, such a variable's many values might each represent the presentation of a slightly differently colored swatch. Variables of this sort fail the letter of proportionality—the one-to-one requirement—in the context of any standard two-valued explanandum variable, one of whose values will represent the occurrence of the event to be explained and the other, its absence. This points towards a possible counter-move for the proponent of proportionality: to describe and defend a separate principle by which the scarlet-cyan variable, and others like it, might fall explanatorily short. When combined with the more technical requirement specified by condition (P), our high-level explanatory preferences might then be both clarified and rationalized.

A supplemental standard that appears to fit this bill is *exhaustivity*, and I will call the hybrid constraint constituted by both exhaustivity and condition (P) the *spirit of proportionality*, since it is what I suspect advocates of the proportionality standard have in mind, without stating it outright. Exhaustivity requires that the cause variable's values collectively exhaust the 'causal possibility space,' the range of circumstances by which the explanandum event—as well as its contrast—might be brought about. Though not mentioning exhaustivity as such in his discussions of explanatory level, Woodward gestures at the substance of the requirement in the course of motivating proportionality's putative rulings, explaining that an account of pigeon pecking should cite red, not scarlet, in order that it might tell us "about the full range of conditions under which the pigeon will peck or not peck"(Woodward 2008d: 161). Separately, he notes that the low-level explanation citing scarlet is inadequate because it "fails to convey the information" that "any shade of red would have caused the pigeon to peck"(Woodward 2010: 298).[15]

[15] As I've stated it, the exhaustivity standard calls upon a 'causal possibility space,' one that I take to include all alternative ways the target event 'might be brought about.' Minimally, these alternatives will be physical-nomological possibilities, that is, they must be instantiable in a world featuring the actual physical laws. Though a defender of the standard might well want to further constrain the space of possibilities—and doing this would be a way of responding to some of the challenges to come—it is not at all easy to do this in a principled manner. Thus, I will assume here only physical constraints on the causal possibility space, and leave it to any fan of the exhaustivity standard who aims to grapple with the concerns I raise to propose additional requirements.

To explore the effectiveness of this tactic I will revisit the pecking explanations offered earlier. And for reasons that will soon become evident, in doing so a third and even higher-level explanatory alternative is added to the mix.

> A pigeon is trained to peck at a target when presented with a stimulus of any shade of red. Suppose, on some particular occasion, the pigeon is presented with a particular shade of scarlet and pecks at the target. Consider three explanations for her pecking:

(1*)—The *presentation of a scarlet target* (other value: *presentation of a cyan target*) causes the pigeon to *peck* (other value: *not-peck*).

(2*)—The *presentation of a red target* (other value: *presentation of a non-red target*), caused the pigeon to *peck* (other value: *not-peck*).

(3)—The *presentation of a red target or provision of food or tickling of the chin or electrical stimulation of the cerebellum* (other value: *none of the above*), caused the pigeon to *peck* (other value: *not-peck*)

All three of these options constitute minimally adequate interventionist explanations for pigeon pecking. For this to be the case, assume that the particular pigeon whose pecking is to be explained is one that, while exposed *in this case* to a scarlet target, would have pecked had it been presented with any red target, had it been provided food, had it been tickled, etc. In stipulating that the pigeon's pecking might have been brought about by a number of different causal pathways (as conventionally individuated)—a property we might call *diverse determination*—nothing tendentious has been assumed. In fact, it would be tendentious only to assume otherwise. First, any actual pigeon-pecking event will, in virtue of the complexity of the system underpinning it, display diverse determination, though the identity of the peck-inducing pathways may vary somewhat from pigeon to pigeon. Second, most if not all other phenomena that we aim to explain will likewise display diverse determination, being outcomes, in principle, of multiple causal pathways: the vase's breaking (by bat or dynamite), the neuron's firing (by electronic discharge or chemical stimulation), the person's vomiting (by whiskey or salmonella). In consequence, if we want our toy example to be relevant to the causal systems of our world—and thus to the explanation of the events they produce—we must grant it the same kind of structure.

If the spirit of proportionality is to select between explanations (1*), (2*), and (3), it must do so based on differences in their exhaustivity, as they each satisfy the one-to-one requirement of proportionality. And select it does. First, exhaustivity rules (2*) superior to (1*), as (2*) captures a number of causal possibilities that (1*) omits, such as what would have happened had the pigeon been exposed to cardinal, rouge,

green or purple. Explanation (1*), by contrast, deals exclusively with what the pigeon would have done had it been presented with a cyan or scarlet target. And on precisely the same grounds exhaustivity finds explanation (3) superior to (2*). After all, explanation (3) is even *more* exhaustive of the causal possibility space than (2*), providing information about what would have happened had the pigeon been given food, tickled, etc. It also captures what would have ensued had it not been exposed to *any* of these conditions. (2*), by contrast, says nothing of these possibilities, and is thus comparatively explanatorily inferior.[16]

Though the first part of this result will be congenial—making sense of our preference for the higher-level red account over the lower-level scarlet one—the second result should not be, at least if we are to judge accounts of explanation by their consistency with actual explanatory practice. Though scientists do sometimes appeal to high-level and even somewhat disjunctive causal factors in their explanations of both events and regularities, in preferring explanations like (3) the exhaustivity requirement—when constrained by condition (P)—goes well beyond any moderate high-level preferences with which you might be familiar: it recommends *maximally* disjunctive accounts, those citing causes that effectively lump together, into a single explanatory factor, *every single means* by which the effect might, in principle, have been brought about.[17] Such accounts are absent from the explanatory annals, presumably in part for their genuine explanatory inferiority; they are pitched at such great heights as to induce a kind of explanatory hypoxia,

---

[16] I have just pressed the priority of (3) over (2*) and (1*) on grounds of exhaustivity, as the case for high-level priority there is clear. Yet one might argue that (1*) and (2*) fall short of (3) based on the proportionality standard *narrowly construed*—that is, without the added requirement of exhaustivity—offering what I will call the *strong reading* of proportionality. Very briefly, according to the strong reading proportionality's requirement that a causal relationship "explicitly or implicitly conveys accurate information about the conditions under which alternative states of the effect will be realized" commits an explanatory causal relationship to conveying *complete* information about every circumstance that would eventuate in alternative states of the effect variable, such that only changes in the value of the cause variable could change the value of the effect variable. (1*) and (2*) would then fall short in virtue of being incomplete in this way; for instance, in the case of (1*) cyan presentation might remain constant, but the pigeon's pecking state could change in virtue of electrical stimulation of the cerebellum.

I have two comments for readers tempted by the strong reading. First, even if it is correct as a point of Woodward exegesis, the thrust of my argument is completely unaffected; it would simply mean that we arrived at one of the central problems—overshooting—one step sooner. Second, the wording of the proportionality standard does not suggest, to me at least, this interpretation. It seems that an explanation like (1*) can convey "accurate information about the conditions under which alternative states of the effect will be realized" without conveying *complete* information on such matters. Thus, a relationship may be proportional while still leaving out information about what would happen were features not represented in the explanation itself—e.g., the presence of absence of electrical stimulation—been themselves modified.

[17] Since disjunctiveness is language-relative, in saying that a variable is 'colloquially disjunctive,' or just plain 'disjunctive,' I mean—here and elsewhere—that it is syntactically disjunctive when its values are described in a language that individuates causal pathways as does our own.

specifying far too little about what actually brought about the explanandum event to be very explanatory of it.[18]

To highlight the inadequacy of these recommendations, consider how proportionality's spirit would treat an example of somewhat more fame than that of the pecking pigeon: a standard ecological explanation of the low rabbit population in some locale. The relevant rabbits, let us assume, are the prey of foxes, whose numbers rose the previous season. In such a case it would be customary—and intuitively acceptable—to explain the dearth of rabbits by citing that fox population

[18] Though the next section sketches my preferred response to the problem just described, a simpler strategy should be considered briefly: to formulate alternative explanations of the target event that are superior to, or co-equal with, (3) as judged by proportionality, yet are not as high-level as is (3). In particular, consider two proposals—both helpfully emphasized by a referee—which I'll call (2-complement) and (2-multi).

(2-complement) is an explanation just like (2*), but which appeals to a cause variable that takes the values (red presentation, no red presentation) rather than (red presentation, non-red presentation). Such an account may satisfy the 1:1 requirement of proportionality, while also being 'medium level,' between (2*) and (3). And because it is as exhaustive as (3), it appears to belie my claim that extremely high-level explanations are actually *required*—rather than just *allowed*—by the explanatory account under consideration. Yet this is not the case: (2-complement) fails explanatorily the same way as did that the explanation, discussed in section 4, citing (scarlet presentation, no scarlet presentation), and is thus not in real competition with either (2*) or (3). As in the other case, some may judge (2-complement) to fail the basic interventionist causal condition; others may see it in tension with (P)'s requirement that an explanatory causal relationship "conveys accurate information about the conditions under which alternative states of the effect will be realized." Either way, (2-complement) indicates that the non-presentation of red will change the pigeon's state from pecking to non-pecking. This is inaccurate because, under the umbrella of 'no red presentation' are circumstances, such as the electrical stimulation of the cerebellum, which would not change the pigeon's pecking state. (2*) and (3) do not have this unfortunate feature.

(2-multi) is related to (3) along what I call the 'single-multi axis.' It is a complex event explanation, one not exhausted by a single type-level causal claim. Instead, it cites a possibly complex network of causal claims in concert with initial conditions. In particular, let the core of (2-multi)—the part corresponding to the actual cause of the pecking event—be a causal claim connecting variable R, taking (red presentation/non-red-presentation), to pecking. Supplement this with a set of causal claim connecting other variables and the pecking variable, one for each disjunct in (3). Among others, add T (tickling-on-chin, non-tickling-on-chin) and C (stimulation-of-the-cerebellum, lack-of-stimulation). This explanation looks to be coequal with (3) in virtue of satisfying both proportionality and a modified version of the exhaustivity standard—one applying to *the entire causal network* rather than to *the values of the cause variable itself*.

Nevertheless, there are two reasons those like Woodward who aim to articulate an explanatory account friendly to the high-level should not take it seriously as an alternative. The first problem speaks against modifying the exhaustivity standard in the way just suggested. Once done, it is no longer possible for the combined standard to do the work Woodward originally proposed for proportionality: to respond to the explanatory fundamentalist by showing a 'red-level' explanation is superior to a 'scarlet-level' one. After all, in addition to the complex event explanation constructed above, there will be another account that is even lower-level, coequal with it, whose 'actual cause' is the fact that a variable taking either (scarlet presentation, cyan presentation) actually took 'scarlet.' Because of this, Woodward and others do well to restrict their application of the exhaustivity standard to the actual cause of an event. Secondarily, when complex event explanations are entered into competition the explanatory account will, as already noted, recommend the lowest-level explanation, and neither (3), (2), or (2-multi). As explained in section 3, when complex accounts are considered alongside simple event explanation, all high-level causal claims will be replaced by complex accounts that open up the black boxes that the higher-level accounts describe more abstractly.

boom, one that eventuated in a series of rabbit eatings and then the low rabbit census itself. As Garfinkel ([1981]), among others, has emphasized, this account is *somewhat* high-level in virtue of citing the fox *population*, rather than some more fine-grained property of the rabbits' lupine environment, such as the precise distribution or activities of the relevant predators.

Yet this explanation, offered as it is at the high-but-not-too-high level so characteristic of actual explanatory practice, is not one that the spirit of proportionality can recommend. It instead requires that we gather together, in one explanatory variable, all states of affairs that might have resulted in the small rabbit populace. Given that the rabbit's population, like the pigeon's pecking, displays diverse determination, just one disjunct of this will be the fox population boom, the prima facie explanatory factor. Among many others, the cause value of the explanatory variable will also contain as disjuncts: the occurrence of a local flood; the presence of human hunters; an outbreak of rabbit influenza; an earthquake that would collapse local warrens and suffocate baby bunnies nestled within. After all, just as tickling or the presentation of the pigeon with crimson might have brought about pecking in our earlier example, these are all means by which the rabbit population might have declined, making them elements of the explanandum event's causal possibility space. Proportionality's spirit thus demands that this factor—however curious or titanic it may appear—be cited in an optimal explanation.[19]

6 Stability to the Rescue?

Though the spirit of proportionality—principle (P) in concert with exhaustivity—systematically overshoots the appropriate explanatory level, proponents of proportionality have a response available: to add but a third standard on explanatory causal variables capable of dealing with this 'overshooting' problem. Such a standard might either completely exclude colloquially disjunctive variables from the explanatory competition, or it might identify some respect in which lower-level explanations were superior to high.[20] In the first case, so long as the exclusion could be explained and justified, the overshooting problem would be immediately solved. In the second case, the solution would go by way of demanding that optimal explanations trade off the satisfaction of the downward force of the proposed standard against the upward force of the spirit of proportionality. Assuming these trade-offs weren't uniformly strict—making possible, at least in principle, mid-level explanatory optima at which upward and downward considerations best balanced—

[19] See Weslake (ms.) for an allied exploration of this problem.

[20] Whether this suggestion would completely rule out all but the lowest-level explanations would depend on whether an account of non-disjunctive but higher-level variables was also provided.

interventionists would have a tool by which to recommend explanations of the sort scientists actually offer up.[21]

I cannot here explore all ways such a countervailing standard might be constructed. In particular, one otherwise-tempting strategy will be considered only briefly by way of conclusion: to solve the overshooting problem using a metaphysical principle—perhaps by demanding that explanatory variables track what Lewis ([1983]) called the "natural properties." This style of solution is considered just as a last resort because it is in strong tension with Woodward's explicitly non-metaphysical proclivities. As Woodward explained in his 2012 PSA Presidential Lecture, beyond a very modest realism his project is intended to require "no particular metaphysical commitments," an aim that cannot be lightly put aside: it is "precisely its unmetaphysical character"(Woodward 2008a: 194) that is, by Woodward's own lights, among interventionism's central attractions.

A strategy more in line with this down-to-earth aspiration is to identify an *explanatory* virtue by which lower-level explanations are superior to high. *Stability*, a property of causal relationships that Woodward ([2006], [2010]) has elsewhere emphasized, seems well-suited to this task.[22] The stability of a causal relationship concerns the number of background conditions that must be maintained for it to hold, where relationships with fewer such conditions are thereby more stable. Stable relationships will then be those that would continue to hold even were very many things different. Assuming background conditions are not themselves individuated in some gerrymandered way, stability promises to prefer lower-level explanations to high, as required for a trade-off-style solution to the overshooting problem. I will explain via an example: the superior stability of the causal relationship cited in the lower-level explanation (2*) in comparison with the higher-level (3).

If the stability of (2*) is to be greater than that of (3), the causal relationship cited in (2*) must require the maintenance of fewer background conditions. Because the causal-explanatory feature from (2*), red presentation, appears in (3) as but one of a number of disjuncts, we can be confident of this inequality. To see this, first recall that each disjunct in (3) reflects a different pathway by which the effect, pecking, might be brought about. For instance, the chain from red exposure to pecking would go by way of the occipital cortex, while that linking food to pecking might go by way of some intermediate olfactory way-station. Second, note that for any of these

[21] This telegraphic description of the trade-off strategy is all that space here permits. For a more thorough development of the trade-off-based approach to scientific explanation—though not one tied to the interventionist framework—see my [forthcoming-a and forthcoming-b].

[22] Note that though Woodward treats stability as an explanatory virtue in his most recent work, Hitchcock and Woodward [2003] did not do so. In contrast, they explain that two explanations G and G', where G' is more stable than G, the superiority of G' follows from the fact that it "makes explicit the dependence of the explanandum on variables treated as background conditions by G"(187). By making those relationships explicit, such explanations *will* be able to answer more w-questions, and will be better on those grounds.

pathways to effectively bridge cause and effect, certain background conditions must be maintained. In the red case, illumination must be present, and the nerve pathway from the eye to the occipital cortex must remain intact. There are equivalent requirements on the efficacy of any of the other of the diverse causes. Third, observe that these required background conditions will be at least somewhat different between any two pathways, neither one subsuming the other. For instance, illumination is required for the red presentation to work as a cause, but not for the food-induced pecking, which may instead require sufficient air circulation.

These features ensure that a more extensive range of background conditions must be maintained for the relationship cited in (3) than for (2*), making the stability of (2*) higher than that of (3). There are more background conditions required in (3) because its total condition will be the conjunction of those for each of its disjuncts, were these each to stand alone as causal-explanatory factors. And as I've just suggested, this combined condition will be more demanding than that for any one self-standing disjunct, since the conditions required for any two disjuncts will be somewhat different.

Stability thus looks to be just the downward force on explanation that was wanted. When balanced against the upward force already rehearsed, perhaps there would be some optimal explanatory level in the middle, the precise location depending on the empirical details, the individuation of background conditions assumed, and the weighing function used to trade-off explanatory virtues. Overshooting problem solved and high-but-not-too-high-level explanations vindicated? Unfortunately not. Though stability is an effective downward force *as between the options just considered,* a now-familiar difficulty lurks: there are other variables, perfectly kosher ones for the metaphysically-neutral interventionist, that might be appealed to in an explanation that maximally satisfied *all* explanatory standards under consideration: principle (P), exhaustivity, and stability. If all standards can be individually satisfied, there will be no trade-offs of any kind, and thereby no capacity for stability to solve the overshooting problem.

For instance, consider a modification of the variable cited in explanation (3), one tweaking it as follows: separately conjoin to each disjunct (red presentation, tickling, food provision, etc.) all the background conditions required for the effect to come about via the factor cited. So, if red presentation requires certain illumination conditions for pecking to ensue, conjoin these conditions to the 'red presentation' disjunct. Make equivalent additions to all other disjuncts. This will leave the cause value of the explanatory variable a disjunction of conjunctions (assuming we are using a language that individuates causal factors in the normal way), and the contrast value of the cause variable its negation. Such a variable will take the cause value just in case one of the many *complete circumstances* in which the pigeon would peck arose, and the other value in any of the total circumstances in which it would not.

An explanation offered in terms of this causal variable is optimally explanatory according to the development of the interventionist account presently under consideration. First, such an explanation is technically proportional; its cause and effect variables are both binary. Second, it is maximally exhaustive; for any way the world might have been, it tells whether, in that case, the pigeon would have pecked or not. Finally, it is maximally stable; there are no changes to background conditions that might disrupt it, and thus none must be maintained for it to hold. This is because what were background conditions in explanation (3) have been integrated into the value of the cause variable itself. This leaves the causal-explanatory relationship just as stable as can be, on a par even with the laws of physics.

This variable may sound far-fetched, even more so than those already considered. As before, it is not one that scientists—high-level or otherwise—would ever call upon. Yet, in their openness to all causal variables that can take at least two values, interventionists cannot rule out such factors. And if not ruled out, the interventionist is left recommending very peculiar, and in some respects very uninformative, explanations, those that don't go *any way* towards specifying the particular circumstance that brought about the effect to be explained. Ultimately, if Molière was right to mock the explanation of a man's drowsiness in terms of opium's 'virtus dormitiva,' then, according to the interventionist, it is partly because that explanation *said too much*.

7 Pointing Toward a Solution

If interventionists are to make sense of the special virtues of high-level explanations, as Woodward has claimed they might, they need to describe and rationalize principles that pick moderately high-level explanations from among the many minimally adequate ones, a collection vast in proportion to the undemandingness of the basic interventionist causal condition. Since deficient explanations may either be too low-level (e.g., fine-grained physical) or too high-level (e.g., disjunctive), to fix on those in between is to solve what Weatherson ([2012]), drawing on a discussion of Strevens ([2008a]), has aptly called 'the Goldilocks Problem,' that of characterizing what it is for an explanation to be, in level, 'just right.'

This paper has pressed how hard it is to solve this problem once we take seriously how many different variables are available to orthodox interventionists. More specifically, I've suggested that explanatory standards thus far proposed—proportionality, exhaustivity, and stability—do not suffice, individually or in combination, to pinpoint explanations that resemble the intuitively satisfying ones articulated by practicing scientists.

As usual, that something's gone wrong is clearer than just how to make things right. Still, at the most general level what's needed is this: an even further constraint

on explanatory variables that limits the features that a variable's values might track (and perhaps also the combination of features represented by a single variable's alternate values) to a sparser set of 'proper' features. Such a limit might be integrated into the above picture in a variety of ways, perhaps by constraining the size of the 'possibility space' over which exhaustivity is sought; or perhaps by acting as an initial variables sieve, filtering out the bizarre options from the start. However deployed, the limit would probably not eliminate the need for the virtues already explored (e.g., exhaustivity, stability); they would still come in to pick between the plurality of good variables remaining.

But whence such a well-tuned limit? The defeatist option is to insist that the best explanations exploit variables representing just the features to which *scientists* themselves appeal. As these include neither unusual fine-grained nor disjunctive ones, they are explanatorily taboo. Yet this suggestion is as unsatisfying as it is simple; in providing no analysis at all, it makes the philosopher's 'account' of explanatory level into a puffed-up science report. But is it possible to do any better? In conclusion, I briefly distinguish two strategies that might aspire to do so—one metaphysical and one pragmatic.

The first approach I have already anticipated: to add to the "thin [and] un-metaphysical" (Woodward [2011]) interventionist picture a requirement that explanatory variables be metaphysically 'natural.' This would be to take a cue from theorists of causation, like Lewis ([1983]), Menzies ([1996]), and Paul ([2000]), who maintain that only natural events or properties are causally related. If anything is different in the application of naturalness here it is that it comes in as an explanatory—rather than a basic causal—requirement. Among the 'unnatural' variables excluded would presumably be the most problematic problem-cases rehearsed above: the colloquially disjunctive one from section 5, and the disjunction of conjunctions from section 6.

How satisfying is this appeal to metaphysics? Presuming that the contours of the natural variables at each level are simply read off the science itself, it looks to be but a short walk from the science-mimicry already derided, and I am myself wary of it for that reason. Yet the minimal distance traversed does offer one small pay-off: it ensures a kind of philosophical honesty, as the philosopher endorsing the naturalness solution is at least admitting that *some* instrument—however mysterious it might ultimately be—is required to locate the explanatorily-apt variables in the bottomless interventionist pit.

A very different kind of approach is to place a *pragmatic* limit on explanatory variables, a limit that appeals in some way to human interests or practices—yet not to metaphysics. Though there are a wide range of pragmatic constraints that might be designed, the most developed suggestion to date, from Campbell ([2007], [2008]), is to prefer explanatory variables with values that we ourselves can presently

manipulate; these he calls the 'control variables.' In particular, Campbell suggests that our ability to manipulate a person's behavior psychologically—by simply talking with her, for instance—but not so easily neurophysiologically, might ground a preference for high-level psychological explanations of behavior over physical ones.

Aside from the question of just how it would apply to the pigeon pecking case, how successful is this proposal? Its advantage over the metaphysical suggestion above is that it places a real gap between the philosophical explanandum—which variables are explanatorily proper—and its explanans—available instrumentation. But another concern arises: that it leaves the selection of explanatory level too deeply anthropocentric. After all, the proper level would end up relative to the *particular technological tools* possessed at a particular time. Even ignoring explanations that seem tied to no manipulative instruments whatsoever, like those from astrophysics, it is far from clear that our explanatory practice has this exceedingly practical character, such that which manipulative tools we possess determines which accounts are explanatorily appropriate.

Apart from such concerns, should interventionists at least try to develop these proposals? Being wary of metaphysics, and yet also eager to distinguish what we can actually intervene on from how we should best explain, I suspect Woodward in particular will be inspired by neither. Instead, he and others will yearn for a solution to the variables problem that lies notionally between the deeply practical and the obscurely metaphysical. Whether this intermediate solution can be found remains to be seen.

References

Baedke, J. [2012]: 'Causal explanation beyond the gene: manipulation and causality in epigenetics', *THEORIA. An International Journal for Theory, History and Foundations of Science,* 27, pp. 153-74.

Batterman, R. W. [2002]: *The devil in the details: Asymptotic reasoning in explanation, reduction, and emergence*, New York: Oxford University Press, USA.

Baumgartner, M. [2012]: 'The Logical Form of Interventionism', *Philosophia,* 40, pp. 751-61.

Campbell, J. [2007]: 'An interventionist approach to causation in psychology', in A. Gopnik and L. Schulz (eds.), *Causal learning: Psychology, philosophy, and computation*, New York, NY: Oxford University Press, pp. 58–66.

--- [2008]: 'Causation in psychiatry', in K. Kendler and J. Parnas (eds.), *Philosophical Issues in Psychiatry: Explanation, Phenomenology, and Nosology*, Baltimore, Maryland: The Johns Hopkins University Press, pp. 196-215.

Campbell, J. (2010), 'II—Control Variables and Mental Causation', *Proceedings of the Aristotelian Society (Hardback)* (110: Wiley Online Library), 15-30.

Craver, C. F. [2007]: *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*, Oxford: Clarendon Press.

Davidson, D. [1967]: 'Causal relations', *The Journal of philosophy,* 64, pp. 691-703.

Franklin-Hall, L. R. [forthcoming-a]: 'The Causal Economy Account of Scientific Explanation', in C. K. Waters and J. Woodward (eds.), *Minnesota Studies in the Philosophy of Science*, Minneapolis, MN: University of Minnesota Press.

--- [forthcoming-b]: 'Explaining Causal Selection with Explanatory Causal Economy: Biology and Beyond', in C. Malaterre and P.-A. Braillard (eds.), *Explanation in Biology*, Basel: Springer.

Garfinkel, A. [1981]: *Forms of explanation: Rethinking the questions in social theory*, New Haven, CT: Yale University Press.

Glennan, S. [2005]: 'Modeling mechanisms', *Studies in History and Philosophy of Biological and Biomedical Sciences,* 36, pp. 443-64.

Halpern, J. Y. and Hitchcock, C. [2010]: 'Actual causation and the art of modeling', in R. Dechter, H. Geffner, and J. Y. Halpern (eds.), *Heuristics, Probability and Causality: A Tribute to Judea Pearl*: College Publications, pp. 383-406.

Hitchcock, C. and Woodward, J. [2003]: 'Explanatory generalizations, part II: Plumbing explanatory depth', *Noûs,* 37, pp. 181-99.

Kitcher, P. [1981]: 'Explanatory unification', *Philosophy of Science*, pp. 507-31.

--- [1984]: '1953 and all that. A Tale of Two Sciences', *The Philosophical Review,* 93, pp. 335-73.

Langer, N., Beeli, G., and Jäncke, L. [2010]: 'When the sun prickles your nose: an EEG study identifying neural bases of photic sneezing', *PloS one,* 5, pp. 1-6.

Lewis, D. [1983]: 'New work for a theory of universals', *Australasian Journal of Philosophy,* 61, pp. 343-77.

Malaterre, C. [2011]: 'Making Sense of Downward Causation in Manipulationism: Illustrations from Cancer Research', *History and Philosophy of the Life Sciences,* 33, pp. 537-62.

Menzies, P. [1989]: 'A Unified Account of Causal Relata', *Australasian Journal of Philosophy,* 67, pp. 59-83.

--- [1996]: 'Probabilistic causation and the pre-emption problem', *Mind*, pp. 85-117.

Paul, L. A. [2000]: 'Aspect causation', *The Journal of philosophy*, pp. 235-56.

Potochnik, A. [2010]: 'Levels of Explanation Reconceived', *Philosophy of Science,* 77, pp. 59-72.

Putnam, H. [1979]: *Mind, language, and reality: Philosophical Papers*, 2, Cambridge, UK: Cambridge Univ Press.

Railton, P. [1981]: 'Probability, Explanation, and Information', *Synthese,* 48, pp. 233-56.

Reisman, K. and Forber, P. [2005]: 'Manipulation and the Causes of Evolution', *Philosophy of Science,* 72, pp. 1113-23.

Reutlinger, A. [2011]: 'A Theory of Non-Universal Laws', *International Studies in the Philosophy of Science,* 25, pp. 97-117.

--- [2012]: 'Getting rid of interventions', *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences,* 43, pp. 781-95.

Reutlinger, A. and Koch, H. [2008]: 'Methodological Individualism and the Epistemic Value of Macro-Explanations', *Percipi,* 2, pp. 1-14.

Shapiro, L. and Sober, E. [2012]: 'Against proportionality', *Analysis,* 72, pp. 89-93.

Stegmann, U. [2012]: 'Causal control and genetic causation', *Noûs,* 48, pp. 450-65.

Strevens, M. [2007]: 'Review of Woodward, Making Things Happen*', *Philosophy and Phenomenological Research,* 74, pp. 233-49.

--- [2008a]: *Depth*, Cambridge, MA: Harvard University Press.

--- [2008b]: 'Comments on Woodward, making things happen', *Philosophy and Phenomenological Research,* 77, pp. 171-92.

Waters, C. K. [2007]: 'Causes that make a difference', *Journal of Philosophy,* 104, pp. 551-79.

Weatherson, B. [2012]: 'Explanation, Idealisation and the Goldilocks Problem', *Philosophy and Phenomenological Research,* 84, pp. 461-73.

Weslake, B. [2010]: 'Explanatory Depth*', *Philosophy of Science,* 77, pp. 273-94.

--- [ms.]: 'The Problem of Disjunctive Explanations'.

Woodward, J. [2003]: *Making Things Happen: A Theory of Causal Explanation*, Oxford: Oxford University Press.

--- [2006]: 'Sensitive and Insensitive Causation', *The Philosophical Review,* 11, pp. 1-50.

--- [2008a]: 'Comment: Levels of Explanation and Variable Choice', *Philosophical issues in psychiatry: explanation, phenomenology, and nosology*, pp. 216.

--- [2008b]: 'Mental causation and neural mechanisms', *Being Reduced: New Essays on Reduction, Explanation, and Causation*, pp. 218–62.

--- [2008c]: 'Response to Strevens', *Philosophy and Phenomenological Research,* 77, pp. 193-212.

--- [2008d]: 'Cause and Explanation in Psychiatry: An Interventionist Perspective', in K. Kendler and J. Parnas (eds.), *Philosophical Issues in Psychiatry*, Baltimore: The Johns Hopkins University Press, pp. 132-83.

--- [2010]: 'Causation in biology: stability, specificity, and the choice of levels of explanation', *Biology and Philosophy,* 25, pp. 287-318.

--- [2011]: 'Interventionism and Causal Exclusion', *PhilSci-Archive*, pp.

Yablo, S. [1992]: 'Mental causation', *The Philosophical Review,* 101, pp. 245-80.