Explanatory Depth*

Brad Weslake†‡

I defend an account of explanatory depth according to which explanations in the nonfundamental sciences can be deeper than explanations in fundamental physics.

1. Introduction. My aim in this article is to develop an account of explanatory depth that preserves the explanatory autonomy of the nonfundamental sciences, by which I mean all sciences apart from fundamental physics. By explanatory depth, I mean a measure in terms of which explanations can be assessed according to their explanatory value. In particular, I aim to defend the view that there are contexts in which it is possible for the nonfundamental sciences to provide deeper explanations than those provided by fundamental physics. Call this view *autonomy* and its negation *fundamentalism*. Autonomy does not entail that there is a single dimension of explanatory depth or that the nonfundamental sciences are capable of providing deeper explanations along all dimensions of explanatory depth. It also does not entail that there are phenomena that cannot be explained by fundamental physics but can be explained by nonfundamental sciences.[1] It simply asserts that there is at least one dimension of explanatory depth along which the nonfundamental sciences provide deeper explanations than those provided by fundamental physics

*Received July 2009; revised September 2009.

†To contact the author, please write to: Department of Philosophy, University of Rochester, Box 270078, Rochester, NY 14627; e-mail: bradley.weslake@rochester.edu.

‡Earlier versions of this article were presented at the Annual Conference of the Australasian Association of Philosophy, Melbourne, July 6–11, 2008, and Models and Simulations 3, University of Virginia, March 6–8, 2009, where I received especially useful comments from Marshall Abrams, Mark Bedau, and John Humphreys. I am also grateful to Alisa Bokulich for helpful correspondence on this topic and to Jeff Glick, Kevin McCain, Chris Pincock, Nandi Theunissen, Jonathan Vogel, and two anonymous referees for helpful comments on earlier versions.

1. A view defended by Putnam (1975) and, for very different reasons, Batterman (2002b).

Philosophy of Science, 77 (April 2010) pp. 273–294. 0031-8248/2010/7702-0001$10.00

and that there are contexts in which this dimension is salient.[2] Likewise, fundamentalism is not the view that nonfundamental explanations should be rejected in favor of explanations provided by fundamental physics or that nonfundamental explanations are not really explanations at all, although it is often used as a premise in arguments for such views.[3] It simply asserts that there is no dimension of explanatory depth along which the nonfundamental sciences can provide deeper explanations than those provided by fundamental physics.

I believe that autonomy is both an intrinsically attractive position and one that fits naturally with the practice of the sciences.[4] In section 2, I take for granted that autonomy is a desideratum of an account of explanatory depth and argue that two otherwise attractive accounts of explanatory depth fail to meet it. In section 3, I reject one way of constructing such an account and argue for my preferred view.

2. Fundamentalism.

2.1. The Deductive-Nomological Account. The contemporary debate over the nature of scientific explanation justifiably begins with the deductive-nomological (DN) view of explanation defended by Hempel and Oppenheim (1948), according to which to explain is to provide a sound deductive argument for some explanandum, with at least one essential premise describing a natural law. Somewhat surprisingly, while Hempel and Oppenheim devote a great deal of attention to the structure of scientific explanation, they never develop an account of explanatory depth. It might be thought that this neglect is justified by the DN view itself, on which with certain further assumptions, naturally attributed to Hempel and Oppenheim, it could be argued that all explanations of a particular event are more or less partial sketches of a single underlying explanation and hence only compete in terms of their relative completeness. This would leave out the question, however, of whether it is possible to rank explanations of different events in terms of their relative explanatory depth.[5] Moreover, this stance would gloss over a puzzle at the heart of the DN view itself.

According to the DN view, the link between explanation and understanding is secured by the notion of *expectability*—the reason that a sound

2. It is the view that Jackson and Pettit (1992) call "explanatory ecumenism."

3. This style of argument is what Block (1995, sec. 3.3) refers to as the "reductionist cruncher."

4. Jackson and Pettit (1992) agree, describing fundamentalism as "an extremely uncongenial doctrine" (171).

5. One may be skeptical that this is possible, a point to which I return later.

deductive argument for the occurrence of some event explains that event is precisely that the argument justifies, to the extent that the premises are justified, the expectation that the event will (or did) occur. One reason the natural law premise is required is to rule out purported explanations that involve deducing the occurrence of an event from itself, perhaps in conjunction with other inessential premises, perhaps including those involving a natural law. Another reason is that the invocation of a premise concerning a natural law furnishes the required deductive connection between the explanans and explanandum since—it may appear—no amount of further premises listing particular matters of non-nomological fact would be sufficient to secure that connection. On closer examination, however, it is clear that what is required for a sound deductive argument for some explanandum is not necessarily a natural law but simply any contingent premise that allows the explanandum to be soundly deduced from the explanans.

For example, consider a set of candidate DN explanations for a particular configuration of planets in the solar system at a time (E). Each explanation in the set contains a premise specifying the specific positions and momenta of the planets relative to the sun at some time:

(P) $P_1, P_2 \ldots P_n$.

Also, each explanation in the set contains a premise specifying a true generalization (L_i) that entails, together with this configuration, the subsequent configuration (i.e., L_i entails that whenever $P_1, P_2 \ldots P_n$, then E):

(L_i) L_i.

Each explanation therefore has the following form:

(P) $P_1, P_2 \ldots P_n$,
(L_i) L_i,
(E) $\therefore E$.

Consider now the following particular generalizations that may be substituted for L_i:

(L_1) Newton's laws of motion.
(L_2) Kepler's laws of motion.
(L_3) Whenever $P_1, P_2 \ldots P_n$, then E.

Intuitively, the explanation generated by substituting L_1 is deeper than that generated by L_2, and likewise for L_2 with respect to L_3. However, on the assumption that each specifies a natural law, the bare DN view does not provide the means to capture this fact about explanation.

The basic challenge to the DN view is this: If to explain is to expect—

and if we have a set of sound deductive arguments for some event, each of which generate equally justified expectations—why should there be a form of explanatory virtue that discriminates between them? To put the point slightly differently, to the extent that we judge there to be a difference in explanatory virtue between the explanations generated by $L_1 - L_3$, it is a difference that requires going beyond the immediate link between explanation and expectability.[6]

As it turns out, the most natural way to incorporate an account of explanatory depth into the DN account is itself suggested by Hempel (1959, 302–3), who mentions in passing the predictive possibilities afforded by laws in situations other than the one under consideration. As I will understand it, the idea is this: while it is true that the explanations generated by $L_1 - L_3$ are identical with respect to the expectability they confer on the explanandum under consideration, $L_1 - L_3$ differ with respect to the range of phenomena they could be recruited to explain. In particular, the possible phenomena that could be explained by employing L_3 are a proper subset of the possible phenomena that could be explained by employing L_2, and likewise for L_2 with respect to L_1. If we now suppose that this relationship tracks a notion of explanatory depth, we have the means to discriminate between the candidate explanations in the manner intuition suggests: the deepest explanation is the one featuring the law employable in the widest range of possible explanations. In what follows, I will refer to this feature of an explanation, whether DN or otherwise, as *scope*.[7]

The DN view of explanation, supplemented with the scope-based account of explanatory depth, solves the puzzle with which I began this section. Unfortunately, it is inconsistent with autonomy. The reason is that fundamental physics aspires to provide the resources for an explanation for every physically possible event, while all other sciences aspire

6. This point is also made by Woodward and Hitchcock (2003b, 189–90). As Batterman (2000a, 230) has noted, Hempel is silent both on why greater expectability makes for better explanations and why expectability grounded in non-nomic generalizations does not provide the same degree of understanding as nomic expectability. On the first, Strevens (2000) provides part of the answer. On the second, I believe that the suggested amendment to the DN view I consider below provides an answer.

7. As Woodward and Hitchcock (2003b, 193) note, it would be a mistake to think that the sheer fact that one generalization is actually instantiated many times and another only once makes for a difference in explanatory depth between the two generalizations, so it is important to note that scope does not imply this implausible claim. In the terminology introduced by Weisberg (2004, 1076), scope tracks *p-generality*, not *a-generality* (see also Matthewson and Weisberg 2009).

to provide explanations for variously restricted subsets of these. It follows that DN + scope entails fundamentalism.[8]

2.2. The Interventionist Account. I turn now to the account of explanatory depth defended by Woodward and Hitchcock (2003b). The account is given in the context of a counterfactual theory of explanation defended in detail in Woodward (2003) and summarized in Woodward and Hitchcock (2003a). According to this theory, to explain is to exhibit patterns of counterfactual dependence relating explanans to explanandum, by describing generalizations that are invariant in the sense that they would continue to hold under various possible changes to the system in question. Woodward and Hitchcock give special emphasis to a particular class of possible changes they call interventions, so I will refer to the view as the *interventionist* account of explanation.

On the interventionist account, an explanation has a similar structure to the DN account, consisting of two components. First, there is a component, analogous to DN initial conditions, describing the actual values of variables in a causal model. Second, there is a component, analogous to DN natural laws, describing the invariant generalizations of the causal model. I will refer to the former as the *particular* component of an explanation and the latter as the *general* component of an explanation. Following Woodward and Hitchcock (2003b, 182) then, an interventionist explanation can be given in the following canonical form:

$$X_1 = x_1, \ldots, X_n = x_n,$$

$$Y = g(X_1, \ldots, X_n),$$

$$\therefore\ Y = y = g(x_1, \ldots, x_n),$$

where g is a functional generalization specifying how y is determined by $x_1, \ldots, x_n$.

The account of explanatory depth given by Woodward and Hitchcock ties depth together with the degree of *invariance* exhibited by the generalizations encoded within the causal model, with greater degrees of invariance making for greater degrees of explanatory depth. As described by Woodward and Hitchcock, there are a number of different ways in which a generalization may be more invariant than another:

8. More carefully, it follows if we make the (very plausible) empirical assumption that the generalizations employed by fundamental physics have wider scope than the generalizations employed in the nonfundamental sciences. Of course, some have denied this, most notably Cartwright (1983, 1999). However, here and throughout, I assume that it is true (see Hoefer 2003 and Sklar 2003 for discussion). I thank Paul Humphreys for drawing my attention to this point.

Accuracy. The generalization may be more accurate within a specific range.
Robustness. The generalization may be invariant under a wider range of interventions.
Continuity. The generalization may be invariant under a more continuous range of interventions.
Stability. The generalization may be invariant under a more continuous range of interventions, where the range includes the actual value of the explanans (this is a special case of continuity).
Insensitivity. The generalization may be invariant under a wider range of ways in which interventions may be performed.
Portability. The generalization may be invariant under a wider range of background conditions. This will typically be because the generalization has made explicit a dependence on factors left out of the original generalization.

In particular cases, these may compete with each other, as, for example, accuracy and portability. Moreover, they do not all naturally map onto corresponding notions of explanatory depth since, for example, we do not necessarily treat a more accurate generalization as ipso facto explanatorily deeper. In addition, there may well be pragmatic reasons for valuing one particular variety of invariance over another in one context and over another in another context. There may also be pragmatic or contextual reasons for valuing one range of interventions or background conditions over others in assessing the individual kinds of invariance. Nevertheless, what they all have in common is that they provide different ways in which a generalization can provide the resources to describe a greater range of true counterfactuals concerning possible changes to the system in question—that is, to answer more *w-questions*, in the sense of Woodward (2003, chap. 5). So as with DN + scope, the interventionist account of explanatory depth focuses on a particular kind of generality an explanatory generalization may possess. While the DN + scope view links depth with the range of systems to which the explanatory generalization applies, the interventionist view links depth with the range of w-questions the explanatory generalization answers for the system in question.

Unfortunately, the interventionist account of explanatory depth is also inconsistent with autonomy. The fundamental laws are those generalizations that are maximally accurate, robust, continuous, stabile, insensitive, and portable. So according to the interventionist view, whichever way the varieties of invariance are to be contextually set and weighed against each other, interventionist explanations employing the fundamental laws will always provide deeper explanations than those provided by the other sciences. Indeed, Woodward (2003, 17) proposes to understand laws of nature as themselves simply one species of invariant generalization, and it is natural on this view to understand the fundamental laws as those generalizations that are maximally invariant, in the sense that they are exactly those generalizations that would obtain under any physically possible transformation whatsoever.[9] It follows directly that interventionism entails fundamentalism.[10]

3. Autonomy.

3.1. The Informational Account. I turn now to consider how an account of depth that saves autonomy might be constructed. The most common strategy here is to make an appeal to the information provided by fundamental and nonfundamental explanations. In this section, I consider and reject two ways in which this strategy can be pursued, corresponding to two ways in which the information may be characterized.

3.1.1. Modal Information. It is often said that nonfundamental explanations provide modal information of a kind that is absent from fundamental explanations. Views of this kind have been defended in one way or another by Garfinkel (1981), Jackson and Pettit (1992), Wilson (1994), and passages in Woodward (2003). In order to evaluate this strategy, consider an example from Woodward (2003, 231–33):

> Suppose that a mole of gas at temperature T and pressure P is confined in a cylinder with a movable piston. The piston is then withdrawn and the gas is allowed to diffuse into the new volume V' while a heat source maintains its temperature at T.

First, there is the following *microscopic strategy* for explaining the new pressure P':

> One notes carefully the position and momentum of each of the 6×10^{23} molecules in the chamber immediately before the withdrawal of the piston (the initial microstate) and then explains the evolving energy and momentum of each molecule in terms of its initial state, the successive collisions it undergoes with other molecules, and the

9. Not all accounts of laws give them this feature. Most notably, the Lewis account of laws does not. See Lange (2008) for discussion.

10. I emphasize that I have here criticized the idea that the account developed by Woodward and Hitchcock (2003b) exhausts the dimensions of explanatory depth. Woodward and Hitchcock do not claim that it does, and Woodward (2003, 265) claims that it does not. I intend the view defended in this article to be fully compatible with the interventionist account of explanation.

> laws governing those collisions. The new pressure P' exerted by the gas is explained by aggregating the energy and momentum transferred by each molecule to the walls of the container. (231–33)

In short, the microscopic strategy consists in a DN or interventionist explanation employing fundamental laws. Second, there is the following *macroscopic strategy* for explaining the new pressure P':

> Given the laws governing molecular collisions, one can show that almost all (i.e., all except a set of measure 0) of the possible initial positions and momenta consistent with the initial macroscopic state of the gas (pressure P, temperature T, and volume V) will lead to a series of molecular trajectories such that the gas will evolve to the macroscopic outcome in which the gas diffuses to an equilibrium state of uniform density through the chamber at new pressure P'. Similarly, there is a large range of different microstates of the gas compatible with each of the other possible values for the temperature of the gas, and each of these states will lead to a different final pressure P_i'. (231–33)

The macroscopic strategy can be summarized in the form of the ideal gas law $PV = nRT$. In short, the macroscopic strategy consists in a DN or interventionist explanation employing nonfundamental laws. Regarding the connection between these two explanations, Woodward (2003, 232–33, 355–56) claims that the information provided by the macroscopic explanation is not "captured" or "represented" by the microscopic explanation, that the microscopic explanation is deficient in not providing information on the conditions under which P' would have been different, that an explanation employing the ideal gas law does provide this information and so provides a better explanation of P', and that while it is true that the microscopic explanation answers a wide range of w-questions, it does not answer the w-question demanded.

The basic doubt I will raise against these claims can be seen by noting that the macroscopic explanatory strategy is constructed precisely by appealing to the microscopic strategy over a range of possible initial conditions, which is to say that the modal information described by the macroscopic explanation has been exclusively procured from the microscopic explanation. This problem with Woodward's claims can be put more simply and generally as follows. The fundamental physical explanation for any event will employ laws of maximal generality in the sense of both scope and invariance. In particular, for any determinate w-question framed in terms of the variables employed by the fundamental physical explanation, the explanatory model will specify a determinate answer.[11] If we assume a reasonable form of physicalism, then there are no questions that can be formulated in terms of any other variables that do not correspond to one of these questions. So there are no physically possible counterfactuals on which the fundamental physical explanation is silent. The fundamental physical explanation provides the resources to answer any possible w-question. In short, a reasonable physicalism entails that there is no missing modal information of the kind claimed.[12]

A similar mistake is made by Jackson and Pettit (1992, 173–75), who likewise describe the difference between microscopic and macroscopic explanations in terms of the information the explanations provide. Jackson and Pettit present the idea in a slightly different way, focusing on information concerning the causes of a particular event and formulating the informational claim in terms of what it is possible for one to be ignorant about. In particular, Jackson and Pettit (1992, 177) argue that the information provided by a macrocausal explanation that is not provided by a microcausal explanation concerns counterfactuals of the form "if the actual history described by the microcausal explanation had not obtained, the explanandum would still have occurred."[13] Of course, not all grounds for believing a counterfactual of this kind provide explanations. For example, knowledge of an unrelated backup cause for any event will ground such a counterfactual while making no contribution to the explanation of that event. To respond to this problem, Jackson and Pettit appeal to the Lewis (1986) account of causal explanation, according to which an explanation must provide information about the actual causal history. This means that the difference between the acceptable macrocausal information and the unacceptable backup cause information must amount to a difference in the information provided about actual causal history, which Jackson and Pettit (1992, 178) say amounts to the question of whether "the counterfactual is grounded in the nature of the actual [causal] history." But this amounts to the admission that the information provided by the macrocausal explanation is in fact already implicit in the microcausal explanation, which by hypothesis is a microphysical description of

11. If the fundamental laws are probabilistic, this may take the form of a determinate probability distribution.

12. Of course, my response here implies that another way to secure autonomy is to reject this modal claim concerning fundamental physical explanations and to endorse the view currently under consideration. According to Loewer (2008, 2009), Fodor (1974, 1997) should be interpreted as taking this option. Thanks to Chris Pincock for prompting me to say more at this point.

13. Essentially the same view is defended by Garfinkel (1981), by Wilson (1994) under the banner "causal depth," and by Batterman (2000a) under the banner "structural stability."

the entire actual causal history. So again, there is no missing modal information of the kind claimed.

Nevertheless, perhaps there is a way to make the distinction between microcausal and macrocausal explanations in a way that saves the idea that macrocausal explanations provide modal information that microcausal explanations do not. This would require microcausal explanations to take a different form than either DN or interventionist explanations since, as I have argued, these explanations implicitly contain all the modal information there is concerning the system in question. Since the idea here would be that the general modal information carried by laws in DN and interventionist explanations is absent from this variety of explanation, I will call these explanations *singular* microcausal and macrocausal explanations. The modal informational account of explanatory depth would then provide a sense in which singular macrocausal explanations are deeper than singular microcausal explanations in virtue of conveying modal information that singular microcausal explanations do not.[14] But this would be a Pyrrhic victory for the modal informational account of depth, for it would remain the case that fundamental DN and interventionist explanations are always deeper than their nonfundamental analogues, which is the result I have supposed an account of depth should license us to deny. The explanations provided by fundamental physics are simply not singular explanations in this sense.

3.1.2. Taxonomic Information. A second way in which informational accounts of explanatory depth have been formulated is in terms of what I will call *taxonomic* information, nonmodal information concerning the way in which explananda are described or individuated. For example, at one point Woodward (2003, 232–33) suggests that the information missing from the microscopic explanation of the gas pressure concerns the relationship between microscopic and macroscopic variables. This is puzzling given Woodward's description of the microscopic explanation since the last sentence of that description shows how the macroscopic variable P' is derived by aggregating the energy and momentum imparted by each molecule to the walls of the container. Nevertheless, suppose it were true that the microscopic explanation did not itself contain the information required to describe the explananda in macroscopic terms. In that case, the explanation should be rejected simply because it does not answer the original explanatory question. But this would again be a Pyrrhic victory

14. Although he is not concerned with explicating a notion of explanatory depth, this way of thinking of the relationship between fundamental and nonfundamental causal explanations is suggested by Campbell (1993, 263–64). It might also be the most charitable reading of Woodward (2008, 233–35).

for the informational account, for if we now suppose that the microscopic explanation were supplemented with the information required, we would have a fundamental explanation as deep as the nonfundamental explanation.

The basic problem here is that the taxonomic informational account of explanatory depth appeals to information that no reasonable account of explanation ought to count as explanatory. The debates concerning the relationship between the description of the world provided by fundamental physics and the descriptions given by the nonfundamental sciences are vexed and interesting, but they are orthogonal to the explanatory question under discussion. Indeed, the interventionist account of explanation itself connects explanation with the provision of modal information in the form of answers to specific w-questions, so this itself rules out nonmodal taxonomic information as beside the point from an explanatory point of view. Exactly the same point applies to the idea that nonfundamental sciences "capture patterns" that fundamental physics does not or provide explanations that are "simpler" or more "understandable."[15] In each case, the relevant informational notion, being nonmodal, is simply explanatorily irrelevant.

3.1.3. Conclusion. I conclude that the difference in depth between microscopic and macroscopic explanations cannot be elaborated in terms of the idea that there is information included in macroscopic explanations that microscopic explanations leave out. Either the supposed information is modal, in which case it is not missing, or it is taxonomic, in which case it is explanatorily irrelevant. The informational account of explanatory depth cannot deliver autonomy.

3.2. The Abstractive Account.

3.2.1. Three Varieties of Explanatory Generality. As I noted earlier, both the amended DN and the interventionist account of explanatory depth focus on a feature of explanations in virtue of which they can be more or less general. While the amended DN view defines depth in terms of the range of possible systems to which an explanatory generalization potentially applies (scope), the interventionist view defines depth in terms of the range of counterfactual questions an explanatory generalization answers (invariance). As I have shown, both views entail fundamentalism since the fundamental laws are exactly those generalizations that are maximally deep in both these senses.

In my view, these views are correct to make a connection between generality and depth but have overlooked the dimension of generality

15. On the former, I agree with Sober (1999, sec. 8).

required to secure autonomy. The key point to notice is that both views focus on varieties of generality possessed by the generalizations employed in explanations. The dimension of generality that has been overlooked, I suggest, is one possessed not by explanatory generalizations but by explanations per se. To see the distinction, consider again the microscopic and macroscopic explanations of the pressure in Woodward's example, focusing on a macroscopic explanation that appeals to the ideal gas law $PV = nRT$. The explanation shows how the new pressure P' of the gas is a function of the temperature T' and volume V' of the gas. Since the ideal gas law holds only under a restricted range of microscopic conditions, the microscopic explanation is more general than the macroscopic explanation in terms of both scope and invariance: the situations in which the ideal gas law applies form a subset of those in which the fundamental laws apply, and the w-questions addressed by the ideal gas law form a subset of those addressed by the fundamental laws. However, there is a sense in which the ideal gas law explanation as a whole is more general than the microscopic explanation since the ideal gas law explanation as a whole applies to a wider range of physically possible systems than does the microscopic explanation, which by hypothesis applies to a single type of physically possible system. To adopt the terminology of Garfinkel (1981), the microscopic explanation is in this sense *hyperconcrete*. Call the degree to which a whole explanation applies to a range of possible situations *abstraction*. My proposal, which I dub the *abstractive* account of explanatory depth, is that abstraction provides a theoretically important dimension of explanatory depth.

Note that I do not suggest that there exists a measure of abstraction. Since typically an explanation will be potentially satisfied by an infinite number of physical systems, the provision of a measure of the degree of abstraction of an explanation looks unforthcoming.[16] This in turn lends some credence to the intuition that the very idea of comparing explanatory depth across different explananda is incoherent.[17] However, a partial ordering of explanations in terms of abstraction is sufficient to secure autonomy. For notice that in Woodward's example, every case in which the

16. Here I agree with Matthewson and Weisberg (2009, sec. 4.3). See Woodward (2003, 261, 288–90, 365–66) and Jones (2005, 197–98) for the same verdict in different contexts. As an anonymous referee pointed out, the problem is not with defining but rather with justifying an appropriate measure. For a different technical problem afflicting the notion of abstraction, see Sober (1999, n. 9 549–50). A related question concerns whether abstraction should track the range of physically possible systems to which an explanation applies or, rather, the range of logically possible systems to which an explanation applies. I opt for logically possible systems, for reasons I explain below.

17. An intuition that, in conversation, Marshall Abrams and Chris Pincock reported some sympathy with (see also n. 5).

microscopic explanation applies is a case in which the macroscopic explanation applies but not vice versa. It follows that the macroscopic explanation is more abstract, whether or not we can compare the abstraction of the macroscopic explanation with the abstraction of other unrelated explanations.

As noted earlier, the abstraction of an explanation is an explanatory virtue that must be traded off against the virtues of scope and invariance. I take it that it is a contingent matter that the world is structured so as to permit a gain in abstraction without significant cost in invariance. The evidence for the world being this way comes from the very existence of nonfundamental sciences and from the particular varieties of explanation I consider in what follows. At least part of the explanation for the world being this way is provided by fundamental physics itself, for example, by way of renormalization group methods (Batterman 2000b).[18] However, for my purposes it is important only that this trade-off exists.

3.2.2. Placing Abstraction. It is surprising that the abstractive account of explanatory depth has never been stated in a form free of further theoretical commitments. The basic idea behind the abstractive account of explanatory depth originates with Hilary Putnam (1967, 1973, 1975). However Putnam claimed not merely that abstraction is a dimension of explanatory depth but that it dominates other dimensions in all contexts, for some explananda. I see no reason to think this is the case.[19]

The idea also bears some resemblance to the unificationist account of explanation defended by Philip Kitcher (1981, 1989, 1999), although the motivation for the requirement is different. In contrast with the unificationist criterion for explanatory depth, abstraction does not necessarily result in the minimization of "the number of types of facts that we have to accept as . . . brute" (Kitcher 1989, 432). Moreover the account is not subject to some important problems for Kitcher's account raised by Woodward and Hitchcock (2003b).[20] The first problem turns on the example of the local temperature of microwave background radiation, which can be repeatedly derived from the assumption that it is homogenous throughout the universe. The problem is that the derivation does not provide answers to any w-questions and so fails to be explanatory. This is not an issue for the abstractive account of depth, which is perfectly compatible with the interventionist theory of explanation. The second problem turns

18. The possibility of explanations of this kind had earlier been noted by Fodor (1974, 107).

19. Here I agree with Sober (1999, 550–51; 2010, sec. 5), Jaworski (2002), and Potochnik (2007).

20. The examples also appear in Woodward (2003, 366–69).

on an example involving two neural mechanisms, N_1 and N_2, the first of which is actually pervasive and the second of which is rare, but both of which can be used to answer the same range of w-questions. The problem for Kitcher's account is that the unification afforded by the explanation of N_1 is greater than that afforded by N_2, which is in tension with the intuition that the two explanations are equally good. What this highlights is that, as with scope (see n. 7), abstraction should be taken to measure the number of possible situations to which an explanation applies, not the number of actual situations to which it applies. Setting aside worries concerning the comparability of explanations for different explananda, presumably N_1 and N_2 are equally abstract in this sense. So again, this is no problem for the abstractive account. The third problem is that it looks as if Kitcher's way of handling explanatory asymmetries rules out the possibility of saying that there can be two explanations of the same phenomena, one more unified and one less unified. This means that degree of unification cannot provide an account of explanatory depth capable of handling cases such as the gas example from Woodward. Again, it is clear that the abstractive account does not have this problem.

The account of explanatory depth most similar to the abstractive account I have outlined is provided by the *kairetic* theory of explanation due to Strevens (2004, 2009). Strevens rightly emphasizes the generality of whole explanations, rightly criticizes the unificationist emphasis on actual generality rather than possible generality, and rightly notes some of the explanatory benefits of abstraction. However Strevens, following Putnam, develops his account by drawing a tight connection between abstraction and the notions of explanatory relevance and difference making. I have shown that the abstractive account of depth can be motivated without making this commitment.

3.2.3. Problems for Abstraction. I turn now to a problem concerning disjunctive predicates that arises in one way or another for any account of explanatory depth that ties depth with generality. The problem is that it seems possible to gain generality cheaply by employing disjunctive predicates in either the particular or the general components of an explanation. For instance, suppose we aim to explain why some given configuration of planets obtains. Consider the explanation with a particular component consisting in the disjunction of the initial configuration of the plants and the initial macroscopic configuration of a particular chamber of gas and a general component consisting in the disjunction of Newton's laws and the ideal gas law. Call this gerrymandered law L_4. This explanation is more general than an explanation employing any of L_1–L_3 since it can explain not only the configuration of the planets but the macroscopic properties of gases. But no account of explanatory depth should say that it provides a deeper explanation in either context.

The solution to this problem is straightforward. The problem with the explanation employing L_4 is that it contains redundant, nonexplanatory information in any particular context in which it is employed. This is in effect the position taken both by Strevens (2004, 170–72), who describes the redundancy in terms of a loss of cohesion (which demands that a causal model contain as many causally active elements as possible), and by Woodward and Hitchcock (2003b, 190), who describe the redundancy in terms of a lack of increase in the interventionist w-questions addressed by the causal model.[21] Either way, the problem is that generality has been bought at the cost of explanatory redundancy. This in turn shows that abstraction is to be understood as not simply requiring maximal explanatory generality but requiring maximal explanatory generality consistent with explanatory nonredundancy.

A second problem may be thought to arise in connection with disjunctive predicates introduced into fundamental explanations.[22] Notice that the microscopic explanation for the gas pressure involves a particular component consisting in an exact description of the initial positions of all gas molecules and so is hyperconcrete and minimally abstract. However, the corresponding explanation involving a particular component consisting in the infinite disjunction of all physically possible initial positions consistent with the gas pressure is, it may be alleged, exactly as abstract as the macroscopic explanation employing the ideal gas law. In reply, I deny that the disjunctive microscopic explanation is as abstract as the macroscopic explanation. To see this, notice that the ideal gas law is independent of whether the underlying mechanics is Newtonian or quantum mechanical. This means that there are physically impossible systems to which the macroscopic explanation applies but to which the microscopic explanation does not. So the macroscopic explanation is more abstract than the disjunctive microscopic explanation. I do not deny that there exist macroscopic explanations that do not have this feature, in which case they will be explanatorily equivalent to disjunctive microphysical explanations. And of course, in this case we still prefer the macroscopic explanation to the gerrymandered microscopic explanation—but

21. See also Woodward (2003, 261). Kitcher's (1981) notion of "stringency" plays a similar role with respect to a similar problem.

22. I am grateful to an anonymous referee for catching an important error in an earlier version of this paragraph.

as I argued earlier, this preference is not grounded in a difference in explanatory virtue.[23]

3.2.4. The Virtues of Abstraction. Why believe that abstraction provides a genuine dimension of explanatory depth? My central claim, of course, is that abstraction provides the best explanation for the truth of autonomy. So far I have defended this claim by way of a single example, in which I showed that abstraction grounds the claim that an ideal gas law explanation is in one respect deeper than the corresponding fundamental explanation for the volume of a gas. It is important to see, however, that the example generalizes. The conditions under which fundamental and nonfundamental explanations will be related in this way are as follows:

- Every possible situation in which the fundamental explanation applies is a case in which the nonfundamental explanation applies.
- There are possible situations in which the nonfundamental explanation applies in which the fundamental explanation does not apply.

This will be the case whenever the nonfundamental explanation supervenes on and is multiply realizable by the fundamental explanation.[24] So it is reasonable to believe that nonfundamental explanations are often autonomous to the extent that it is reasonable to believe that the variables employed in explanations provided by the nonfundamental sciences often supervene on, and are multiply realizable by, the variables employed in fundamental physics. In my view, this makes the autonomy of nonfundamental explanations with respect to fundamental explanations pervasive. Indeed, there is a continuum here corresponding to the traditional hierarchy of the sciences, with physics at the bottom and the social sciences at the top. From the perspective of the abstractive account of explanatory depth, the explanatory strategies of these sciences can be seen to involve different norms concerning the way invariance and abstraction are to be

23. This is not to say that the preference is wholly pragmatic. It may, e.g., be the case that the macroscopic explanation has confirmatory virtues that the microscopic explanation lacks (Pincock 2007). More important, note that even if many or most macroscopic explanations are equivalent to disjunctive microscopic explanations, abstraction still provides an account of what it is about these explanations that renders them superior to nondisjunctive microscopic explanations. That is, abstraction entails that some fundamental physical explanations (those equivalent to nonfundamental explanations) are deeper than others (those not equivalent to nonfundamental explanations) because they are more abstract. In what follows, I offer more examples where equivalence fails.

24. This holds, so long as multiple realization implies the existence of physically impossible potential realizers. This qualification is assumed in what follows. I argue that this is the case for a variety of different explanations below.

weighed against each other. An explanation in fundamental physics has the advantage of maximizing invariance, while explanations in the nonfundamental sciences have the advantage of variably increasing abstraction.[25]

In making this claim, I should emphasize that I am engaging in a certain idealization. I do not claim that any of the explanations currently provided by the fundamental and nonfundamental sciences are related in this way. Nor do I claim that at the limit of inquiry explanations will be related in this way. Rather, I am making a claim concerning what we might think of as *ideally complete explanations* in terms of the variables distinctive of a particular science.[26] In physics, these ideally complete explanations will take the form described by Woodward in his example, and my claim is that abstraction provides a dimension according to which these explanations are inferior to certain nonfundamental explanations. Potochnik (2010) has argued convincingly that many, if not most, of the explanations actually provided by more fundamental sciences are not ideally complete in this sense and do not serve as subvenience bases for less fundamental explanations. In her example, the variables composing genetic explanations for the evolution of a trait typically do not provide subvenience bases for the variables composing corresponding phenotypic explanations. Instead of it being the case that one explanation is more abstract than the other, they are abstract in different ways—one abstracts away from phenotype, and one abstracts away from genotype. Both explanations, meanwhile, abstract away from a vast array of microphysical detail. My claim, then, is that abstraction provides the best explanation for the autonomy of nonfundamental explanations with respect to ideally complete fundamental explanations.

It would be nice to have some additional reasons to believe the abstractive account of explanatory depth, and indeed there are many to be found. As it turns out, a large array of explanatory practices and preferences manifested in the sciences can be explained and unified by the abstractive account of explanatory depth.

First, the abstractive account of depth tracks the attitudes of scientists. For instance, as documented by Weisberg (2004), Hoffmann (1995, 1998) has claimed that more general, less accurate "qualitative models" in chemistry aid understanding in a way not captured by more precise computational models. The abstractive account provides an objective notion of

25. Woodward (2003, 262–65) argues that different sciences also involve different norms concerning the variables that generalizations are valued for being invariant with respect to.

26. Here I deliberately echo the notion of an "ideal explanatory text" introduced by Railton (1980, 1981).

explanatory virtue that justifies this claim, without requiring recourse to subjective notions of simplicity and understandability.[27]

Second, the abstractive account explains why indeterministic explanations can be preferable to deterministic explanations since a gain in abstraction can often be made at the price of predictive accuracy. As Glymour (1980, 37) writes, it "is a striking feature of scientific reasoning that we are willing to sacrifice a bit of empirical accuracy for a gain in explanatory unification."[28]

Third, the abstractive account explains why it is that idealizing models possess a distinctive kind of explanatory value since a gain in abstraction can often be made by omitting representational details, or by sacrificing representational accuracy, while only negligibly sacrificing invariance.[29] Fourth, the abstractive account justifies the explanatory intuitions behind some classic problem cases for theories of causal explanation, in which deep explanations are provided precisely by abstracting away from causal details. I have in mind, in particular, equilibrium explanations and geometric explanations.

In equilibrium explanations, the final state of a system is explained by showing that it is an endpoint of many different initial configurations. Equilibrium explanation was initially brought to the attention of philosophers by Sober (1983), who took as an example Fisher's famous explanation for equilibrium sex ratios in biological populations (see Charnov 1982 for a survey). Other examples abound. The "asymptotics of the first kind" discussed in Batterman (2000a) are equilibrium explanations. The explanatory advantage provided by the inflationary universe hypothesis over the standard big bang model in cosmology is the advantage of an equilibrium explanation over a nonequilibrium (Maudlin 2007, sec. 7). And part of the role of the concept of equilibrium in evolutionary game theory is to provide equilibrium explanations in this sense (Skyrms 1992, 2000).

In geometric explanations, geometric facts combine with empirical facts to explain. Examples from folk physics include the inability to untie a

27. Indeed, while abstraction sometimes lines up with simplicity, it does not always, as will be obvious from the range of explanations covered in what follows. For this reason it is a virtue of the abstractive account that the explanatory value of abstraction does not come from simplicity but from generality.

28. I thank an anonymous referee for reminding me that an argument for this claim, based on the explanatory role of fitness, is spelled out persuasively by Sober (1984/1993, sec. 4.3).

29. Here I disagree with Strevens (2007, 2009), on whose account a gain in explanatory power can never be achieved by omitting details that are causally relevant to the explanandum. See Matthewson and Weisberg (2009) for a detailed defense of the claim that abstraction trades off against precision.

knot explained by topological features of the knot (Kitcher 1989, 426) and the geometrical explanation for why a square peg cannot fit into a round hole (Putnam 1973, 295–97). A more sophisticated example along the same lines is Euler's explanation for the impossibility of walking certain kinds of paths over the bridges of Königsberg (Pincock 2007, 257–60). Examples from real physics include dimensional explanations (Lange 2009) and explanations invoking space-time geometries (Nerlich 1979). A particularly lovely example from biology is the explanation for allometric scaling laws due to West, Brown, and Enquist (1997), which makes central use of the fractal properties of vascular systems.

What is the distinctive explanatory virtue of all of these explanations? In each case, the equilibrium or geometric explanation supervenes on and is multiply realizable by more fine-grained explanations and so is more abstract. According to the abstractive account of explanatory depth, it is the abstraction of equilibrium and geometric explanations that makes for their explanatory virtue—and the mistake of traditional causal theories of explanation such as Lewis's (1986) is to suppose that explanations can never be (nonpragmatically) improved by omitting rather than incorporating causal details.

Note that many of the explanations I have described provide more examples of explanations that would remain explanatory even if the fundamental laws of nature were different, within a certain set of constraints, from what they actually are. To take just one example, just as the ideal gas law is independent of whether the underlying mechanics is Newtonian or quantum mechanical, the square peg explanation works not just in the nomologically possible worlds but in all possible worlds consistent with basic laws concerning rigid bodies.[30] It is for this reason that I propose that abstraction is to be understood in terms of the set of logically possible systems compatible with a given explanation. An explanation is deeper in this sense, then, if it applies not merely to a wide range of other nomologically possible situations but to a wide range of logically possible situations.

Finally, we are often interested in explanatory questions that demand an abstracted answer. For instance, to take an example from Batterman (2002a, 26), suppose we observe a stiff ribbon of steel being gradually loaded up until it buckles. We may be interested in the precise details of how this particular ribbon buckled, but more typically we are interested in general properties of steel with the same macroscopic characteristics and so seek an explanation that would have applied not just to this piece but to all pieces sharing those characteristics.

30. Lange (2000, chap. 8; 2002; 2005) also identifies this feature of higher-level generalizations as contributing to their explanatory power.

4. Conclusion. It is a widely held intuition that generality is an important dimension of explanatory depth and that this has something to do with the autonomy of nonfundamental explanations. However, attempts to articulate this intuition have either identified the wrong form of generality or tied the right form together with more contentious theoretical commitments. I have shown that the abstractive account of depth both identifies the right form of generality and possesses a theoretical attraction independent of the details of any particular theory of explanation. Autonomy is true because abstraction is a genuine dimension of explanatory depth.

REFERENCES

Batterman, R. W. 2000a. "A 'Modern' (= Victorian?) Attitude towards Scientific Understanding." *Monist* 83 (2): 228–57.

———. 2000b. "Multiple Realizability and Universality." *British Journal for the Philosophy of Science* 51 (1): 115–45.

———. 2002a. "Asymptotics and the Role of Minimal Models." *British Journal for the Philosophy of Science* 53 (1): 21–38.

———. 2002b. *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence*. Oxford: Oxford University Press.

Block, N. 1995. "The Mind as the Software of the Brain." In *An Invitation to Cognitive Science*, vol. 3, *Thinking*, ed. E. E. Smith and D. N. Osherson, 377–426. Cambridge, MA: MIT Press.

Campbell, J. 1993. "A Simple View of Colour." In *Reality, Representation and Projection*, ed. J. Haldane and C. Wright, 257–68. Cambridge: Cambridge University Press.

Cartwright, N. 1983. *How the Laws of Physics Lie*. Oxford: Oxford University Press.

———. 1999. *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.

Charnov, E. L. 1982. *The Theory of Sex Allocation*. Princeton, NJ: Princeton University Press.

Fodor, J. A. 1974. "Special Sciences; or, The Disunity of Science as a Working Hypothesis." *Synthese* 28 (2): 97–115.

———. 1997. "Special Sciences: Still Autonomous after All These Years." *Noûs* 31:149–63.

Garfinkel, A. 1981. *Forms of Explanation: Rethinking the Questions in Social Theory*. New Haven, CT: Yale University Press.

Glymour, C. 1980. "Explanations, Tests, Unity and Necessity." *Noûs* 14 (1): 31–50.

Hempel, C. G. 1959. "The Logic of Functional Analysis." In *Symposium on Sociological Theory*, ed. L. Gross, 271–87. New York: Harper & Row. Repr. with revisions. 1965. In *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, 297–330. New York: Free Press.

Hempel, C. G., and P. Oppenheim. 1948. "Studies in the Logic of Explanation." *Philosophy of Science* 15 (2): 135–75.

Hoefer, C. 2003. "For Fundamentalism." *Philosophy of Science* 70 (5): 1401–12.

Hoffmann, R. 1995. *The Same and Not the Same*. New York: Columbia University Press.

———. 1998. "Qualitative Thinking in the Age of Modern Computational Chemistry; or, What Lionel Salem Knows." *Journal of Molecular Structure* 424 (1): 1–6.

Jackson, F., and P. Pettit. 1992. "In Defense of Explanatory Ecumenism." *Economics and Philosophy* 8 (1): 1–21.

Jaworski, W. 2002. "Multiple-Realizability, Explanation, and the Disjunctive Move." *Philosophical Studies* 108 (3): 298–308.

Jones, M. R. 2005. "Idealization and Abstraction: A Framework." In *Idealization XII: Correcting the Model; Idealization and Abstraction in the Sciences*, ed. M. R. Jones and

N. Cartwright, 173–218. Poznán Studies in the Philosophy of the Sciences and the Humanities 86. Amsterdam: Rodopi.
Kitcher, P. 1981. "Explanatory Unification." *Philosophy of Science* 48 (4): 507–31.
———. 1989. "Explanatory Unification and the Causal Structure of the World." In *Scientific Explanation*, ed. P. Kitcher and W. Salmon, 410–505. Minnesota Studies in the Philosophy of Science 13. Minneapolis: University of Minnesota Press.
———. 1999. "Unification as a Regulative Ideal." *Perspectives on Science* 7 (3): 337–48.
Lange, M. 2000. *Natural Laws in Scientific Practice*. Oxford: Oxford University Press.
———. 2002. "Who's Afraid of Ceteris-Paribus Laws? or, How I Learned to Stop Worrying and Love Them." *Erkenntnis* 57 (3): 407–23.
———. 2005. "Ecological Laws: What Would They Be and Why Would They Matter?" *Oikos* 110 (2): 394–403.
———. 2008. "Why Contingent Facts Cannot Necessities Make." *Analysis* 68 (298): 120–28.
———. 2009. "Dimensional Explanations." *Noûs* 43 (4): 742–75.
Lewis, D. 1986. "Causal Explanation." In *Philosophical Papers*, vol. 2, 214–40. Oxford: Oxford University Press.
Loewer, B. 2008. "Why There *Is* Anything Except Physics." In *Being Reduced: New Essays on Reduction, Explanation, and Causation*, ed. J. Kallestrup and J. Hohwy, 149–63. Oxford: Oxford University Press.
———. 2009. "Why Is There Anything Except Physics?" *Synthese* 170 (2): 217–33.
Matthewson, J., and M. Weisberg. 2009. "The Structure of Trade-Offs in Model Building." *Synthese* 170 (1): 169–90.
Maudlin, T. 2007. "A Modest Proposal Concerning Laws, Counterfactuals, and Explanations." In *The Metaphysics within Physics*, 5–49. Oxford: Oxford University Press.
Nerlich, G. 1979. "What Can Geometry Explain?" *British Journal for the Philosophy of Science* 30 (1): 69–83.
Pincock, C. 2007. "A Role for Mathematics in the Physical Sciences." *Noûs* 41 (2): 253–75.
Potochnik, A. 2007. "Optimality Modeling and Explanatory Generality." *Philosophy of Science* 74 (5): 680–91.
———. 2010. "Levels of Explanation Reconceived." *Philosophy of Science*. 77 (1): 59–72.
Putnam, H. 1967. "Psychological Predicates." In *Art, Mind, and Religion*, ed. W. H. Capitan and D. D. Merrill, 37–48. Pittsburgh: University of Pittsburgh Press. Repr. 1975. "The Nature of Mental States." In *Philosophical Papers*, vol. 2, *Mind, Language and Reality*, 429–40. Cambridge: Cambridge University Press.
———. 1973. "Reductionism and the Nature of Psychology." *Cognition* 2 (1): 131–46.
———. 1975. "Philosophy and Our Mental Life." In *Philosophical Papers*, vol. 2, *Mind, Language and Reality*, 291–303. Cambridge: Cambridge University Press.
Railton, P. 1980. "Explaining Explanation: A Realist Account of Scientific Explanation and Understanding." PhD diss., Princeton University.
———. 1981. "Probability, Explanation, and Information." *Synthese* 48 (2): 233–56.
Sklar, L. 2003. "Dappled Theories in a Uniform World." *Philosophy of Science* 70 (2): 424–41.
Skyrms, B. 1992. "Chaos and the Explanatory Significance of Equilibrium: Strange Attractors in Evolutionary Game Dynamics." In *PSA 1992: Proceedings of the 1992 Biennial Meeting of the Philosophy of Science Association*, vol. 2, ed. D. Hull, 374–94. East Lansing, MI: Philosophy of Science Association.
———. 2000. "Stability and Explanatory Significance of Some Simple Evolutionary Models." *Philosophy of Science* 67 (1): 94–113.
Sober, E. 1983. "Equilibrium Explanation." *Philosophical Studies* 43 (2): 201–10.
———. 1984/1993. *The Nature of Selection: Evolutionary Theory in Philosophical Focus*. Repr. Chicago: University of Chicago Press.
———. 1999. "The Multiple Realizability Argument against Reductionism." *Philosophy of Science* 66 (4): 542–64.
———. 2010. "Evolutionary Theory and the Reality of Macro Probabilities." In *The Place*

of Probability in Science, ed. E. Eells and J. Fetzer. Boston Studies in the Philosophy of Science 284. Dordrecht: Springer.
Strevens, M. 2000. "Do Large Probabilities Explain Better?" *Philosophy of Science* 67 (3): 366–90.
———. 2004. "The Causal and Unification Approaches to Explanation Unified—Causally." *Noûs* 38 (1): 154–76.
———. 2007. "Why Explanations Lie: Idealization in Explanation." Unpublished manuscript, Department of Philosophy, New York University.
———. 2009. *Depth: An Account of Scientific Explanation*. Cambridge, MA: Harvard University Press.
Weisberg, M. 2004. "Qualitative Theory and Chemical Explanation." *Philosophy of Science* 71 (5): 1071–81.
West, G. B., J. H. Brown, and B. J. Enquist. 1997. "A General Model for the Origin of Allometric Scaling Laws in Biology." *Science* 276 (5309): 122–26.
Wilson, R. A. 1994. "Causal Depth, Theoretical Appropriateness, and Individualism in Psychology." *Philosophy of Science* 61 (1): 55–75.
Woodward, J. 2003. *Making Things Happen: A Theory of Causal Explanation*. New York: Oxford University Press.
———. 2008. "Mental Causation and Neural Mechanisms." In *Being Reduced: New Essays on Reduction, Explanation, and Causation*, ed. J. Kallestrup and J. Hohwy, 218–62. Oxford: Oxford University Press.
Woodward, J., and C. Hitchcock. 2003a. "Explanatory Generalizations," pt. 1, "A Counterfactual Account." *Noûs* 37 (1): 1–24.
———. 2003b. "Explanatory Generalizations," pt. 2, "Plumbing Explanatory Depth." *Noûs* 37 (2): 181–99.