Interventionism, Explanatory Knowledge and Understanding

Abstract: *Despite widespread agreement among philosophers of science that understanding is constituted by explanatory knowledge, philosophical theories of explanation typically don't look like theories of understanding, or accounts of a species of knowledge, at all. As a results, such theories have rarely been evaluated explicitly with respect to how successfully they account for the role of understanding. Indeed, in the only sustained attempt to engage with this issue, Jaegwon Kim (1994) comes to the startling conclusion that extant accounts of explanation resolutely fail as analyses of understanding. Almost 20 years on, little effort appears to have been made to rectify this situation. In this paper, however, we argue that when evaluated explicitly with respect to how successfully they account for the connection between explanatory knowledge and understanding, it is the interventionist theory of explanation which emerges as a clear frontrunner. Crucially, interventionism can meet four challenges which face any attempted analysis of this connection. First, it provides a satisfying account of how explanatory knowledge differs from merely descriptive knowledge. Second, it allows us to accurately characterize the sense in which explanatoriness, and thus understanding, comes in degrees. Third, it is able to account for the role of both causal and noncausal dependence relations in promoting understanding. And finally, it is able to articulate the important connection between explanation, understanding and unification.*

Wordcount: *7815*

Keywords: *Understanding; Knowledge; Explanation; Unification; Interventionism*

"[T]his is the essence of scientific explanation – science increases our understanding of the world"

Michael Friedman (1974:15).

"Perhaps the most important fruit of modern science is the understanding it provides of the world in which we live… Such understanding results from our ability to fashion scientific explanations"

Wesley Salmon (1986:259).

"Scientific explanations are intended to provide objective understanding of nature. The task of characterizing the ideal notions of explanation… is thus one of brining into focus one of the basic aims of science"

Philip Kitcher (1989:54).

1. Introduction:

Among philosophers of science, there is a broad consensus that to *understand why p* occurs, one must possess an explanation of *p*; a proposition equivalent to "*p* because *q*".[1] As Alexander Bird argues, for example, '[t]o understanding why something occurred is to *know* what causes, processes, or laws brough it about.' (2007:84).[2] There is, however, a puzzle at the heart of this picture, one which has received surprisingly little attention within the relevant literature.

Despite the prevalence of the idea that understanding is constituted by explanatory knowledge, philosophical theories of explanation typically don't look much like theories of understanding, or accounts of a species of knowledge, *at all*. Consequently, such theories have rarely been evaluated explicitly with respect to how successfully they account for the role of understanding. Indeed, in the only sustained attempt to engage with this issue, Jaegwon Kim (1994) comes to the startling conclusion that extant accounts of explanation resolutely *fail* as analyses of understanding. Almost 20 years on, little effort appears to have been made to rectify this situation.

Engagement with Kim's (1994) work has tended to focus upon his ancillary claim that causation is merely one of a variety of dependence relations serving as the "objective correlates" of explanation. In light of recent interest in the notion of *grounding*, thought by many to be a noncausal dependence relation which provides the objective correlate for a

---

[1] A distinction is often drawn between "objectual understanding" and "understanding-why". In this paper, our focus shall be on this latter notion, typically taken to be a type of cognitive achievement expressed by sentences like "*S* understands why *E*", where *E* is an explanandum. For contemporary analyses of the former notion, not discussed here, see e.g., Lipton (2009); Kvanvig (2009); Gijsbers (2013); Strevens (2013); Kelp (2015, 2017); de Regt (2017); Khalifa (2017); Wilkenfeld (2017); Dellsén (2018).

[2] Also see e.g., Salmon (1984, 1989); Lewis (1986); Miller (1987); Hitchcock & Woodward (2003a); Woodward (2003); Lipton (2003); Strevens (2008).

distinctive form of metaphysical explanation, this focus is unsurprising.[3] However, acknowledging the contribution of noncausal dependence relations to explanatory knowledge is, according to Kim (1994), merely one small step upon the road to a viable characterization of understanding. Equally important is the need to explain *how* and *why* knowledge of these relations contributes to understanding.

In answering this question, Kim (1994) argues that a successful account of understanding ought to illuminate both: the unificatory role of dependence relations, in reducing the number of independent phenomena we need to recognize as fundamental; as well as the important sense in which explanatoriness and understanding come in degrees. Disappointingly, however, at this tantalizing juncture Kim throws in the towel. Bridging the gap between dependence, unity and understanding is, he admits 'the hard part; one scarcely knows where to start' (1994:69).

In this paper, we argue that *interventionism* represents the most viable option in terms of a theory of explanation capable of successfully characterizing the connection between explanation and understanding. In the first half (sections II-IV), we provide a detailed analysis of Kim's motivation in claiming that extant theories fail to adequately account for this connection. In the second, we build upon the work of James Woodward and Christopher Hitchcock (Hitchcock & Woodward 2003a, 2003b; Woodward 2003), in demonstrating how an interventionist interpretation of explanation can meet the demands of a theory of understanding highlighted by Kim (1994).

In section V, we show that interventionism provides a satisfying account of the contribution of explanatory knowledge to understanding, and how it differs from merely

[3] See e.g., Audi (2012); Trogdon (2013, 2018); Schaffer (2012, 2016, 2017); Wilson (2016, 2018); Maurin (2018); Trogdon & Skiles (2021).

descriptive knowledge. On the interventionist account, the distinctive value of explanatory knowledge lies in enabling us to 'distinguish those relationships that are potentially exploitable for the purposes of manipulation and control from those that are not' (Woodward, 2003:36).

In section VI, we argue that the interventionist notion of *explanatory depth* allows us to accurately characterize the sense in which explanatoriness and understanding come in degrees. The depth of an explanation is a matter of the range of interventions under which it is *invariant*. The greater this range, the deeper the explanation. We propose that the degree of understanding which can be attributed to an agent, *S*, is directly proportional to the depth of the explanation known to *S*.

Given the widespread assumption that interventions are a purely causal notion, Kim's position on the role of *non*causal dependence in facilitating understanding might seem at odds with our own interventionist commitments. However, in section VII, we argue that interventionism is best understood as an analysis of explanatory dependence *simpliciter*; with interventions functioning to provide knowledge of *causes* in the case of causal explanation and knowledge of *grounds* in the case of metaphysical explanation.

In section VIII, we turn to the role of unification in understanding. In a recent paper, David Kovacs (2020) suggests that appealing to dependence relations does little to elucidate the connection between explanation and understanding. Instead, Kovacs argues that if we are to take this connection seriously, in science or metaphysics, then we ought to adopt a *unificationist* analysis of explanation[4]. In contrast, we propose that an interventionist

[4] See e.g., Friedman (1974); Kitcher (1981, 1989).

methodology can better articulate the role of unification in explanation and understanding owing *precisely* to its focus upon the illumination of dependence relations.

II. Understanding as Nomic Expectability

Despite the close connection typically taken to hold between explanation and understanding, Kim notes that terms like "understanding", "intelligibility" and "explanatory knowledge" are often jettisoned when 'serious theory construction begins… [and] seldom make an appearance once the initial stage-setting is over' (1994:52). One reason for this, is the long-held suspicion that understanding is a purely "psychological", or "pragmatic" (Bunge 1973).[5]

This position can be traced back (as least) as far as Carl Hempel (1965) who argues that, in the context of understanding, explanation is 'a relative notion: something can be significantly said to constitute an explanation in this sense only for this or that individual' (1965:426). In contrast, by focusing upon *nomic expectability*, Hempel intends his own analysis to reflect only the "objective" aspect of explanation: 'the sense of being independent of idiosyncratic beliefs and attitudes on the part of the scientific investigators' (1983/2001:374).

On Hempel's deductive-nomological (DN) model, an explanation is an argument consisting of a premiss *G* (a statement representing the occurrence of an event *g*), a conclusion *E* (a statement representing the occurrence of an event *e*), and a law *L* (a statement specifying a universal generalization) allowing for the deduction of *E* from *G*. Crucially,

[5] Trout (2002), for example, has argued that understanding is nothing more than a feeling of confidence or satisfaction gained when one has seemingly answered a question adequately. And Humphreys similarly argues that a focus upon understanding will lead to a 'relativization of explanations to an individual' and that, as such, we ought to 'set aside that whole issue of what constitutes or promotes understanding' (1989:127).

Hempel argues that 'the argument shows that, given the particular circumstances and the laws in question the occurrence of the phenomenon *was to be expected*' (1965:337).

Given this focus, it ought to come as little surprise that Kim (1994) finds the DN model ill-place to illuminate the role of explanatory knowledge in promoting understanding. Kim argues that Hempel's view collapses the explanatory relation between *g* and *e* into a logical relation between their descriptions, *G* and *E*. Explanation becomes a matter of "logico-linguistic" connections between *descriptions* of events, and 'the job of formulating an explanation consists, it seems, in merely re-arranging appropriate items in the body of propositions that constitute our total knowledge at the time' (Kim, 1994:55-56). Understanding, on the DN model, takes place entirely *within* the "epistemic system" and, as a result, 'on the "subjective" side of the divide between knowledge and the reality known, or between representation and represented' (Kim, 1994:56).

What's more, such "explanatory *internalism*" fails to capture the important sense in which understanding and explanatoriness appear to come in degrees.[6] On the DN-model, explanation (and therefore, understanding) results from the explanandum statement being *derivable* from, or *entailed* by, the explanans statements. These logical notions are an "all-or-nothing" affair. *E* is either entailed by *L* and *G* or it is not; *G* is either a successful explanation of *E*, capable of promoting understanding, or it is not. There is no middle ground.

More promising, however, are the *unificationist* accounts of Michael Friedman (1974) and Philip Kitcher (1981), both of whom appear to be cognizant of the importance of understanding to an account of explanation. Friedman, for example, demands that 'a theory of scientific explanation tell us what it is about the explanation relation that produces

[6] As Hills notes: 'we do talk about understanding why (in at least some contexts) as if it comes in degrees and there is some suggestive linguistic evidence supporting this. "Understanding why p" is "gradable", that is, it is similar to verbs such as "regret" or adjectives such as "tall" (2015:665). Also see Khalifa (2013).

understanding' (1974:6); and Kitcher criticises the DN model for failing to explain 'why those derivations which employ laws advance our understanding' (1981:168).

Despite this apparent desire to carve out a central role for understanding, Friedman and Kitcher's analyses also tie explanation to law-based derivations; 'to other items within our epistemic system, not some objective facts about external events or phenomena' (Kim, 1994:64). As we shall see in the next section, while Kim agrees upon the importance of unification to our practices of explanation and understanding, he ultimately concludes that these unificationist accounts suffer from similar problems to the DN model.

III. Understanding as Unification

According to Kitcher, it is unification, rather than nomic expectation, which is doing the explanatory work on the DN model. Behind Hempel's "official" nomological view, we are invited to recognize an "unofficial" view which regards explanation as unification (Kitcher, 1981:167). As Hempel himself acknowledges, what explanation aims at:

> 'is not [an] intuitive and highly subjective kind of understanding, but an objective kind of insight that is achieved by a systematic unification, by exhibiting the phenomenon as manifestations of common, underlying structures and processes that conform to specific, testable, basic principles' (1966:83).[7]

In pursuing this "unofficial view", Kitcher argues that an argument or derivation is explanatory in virtue of its membership of a class which *best unifies* our system of beliefs.[8] Explanation provides understanding by 'showing us how to derive descriptions of many

[7] Also see Hempel (1965:345 & 444).
[8] Friedman (1974) provides a similar analysis which focuses on laws, rather than classes. As Kim highlights: '[f]or Friedman whether or not a given law explains another is determined crucially b the unifying power of the explaining law, and the concepts in terms of which the latter is explained are exclusively logical ones (equivalence, implication, etc.) and evidential ones ("independent acceptability")' (1994:64).

phenomena, using the same patterns of derivation again and again', and in doing so, 'teaches us how to reduce the number of types of facts we have to accept as ultimate (or brute)' (Kitcher, 1989:432).

Crucially for Kim, however, while Kitcher endorses Friedman's demand that a theory of explanation give an account of understanding, neither suggests that explanation involves discovering or imparting, *additional* knowledge. In both cases, explanation is once again an activity which consists in constructing derivations, the steps of which are logically related to the rest of our belief system. Unification, then, also appears to depend 'solely on factors internal to the epistemic system, such as the number of argument patters required to generate the given class of arguments, the "stringency of patterns, etc' (Kim, 1994:64).

Unificationist accounts of explanation suffer from similar problems to the DN model, in taking the epistemic virtues of understanding to concern 'our representations of the world, but not the world itself' (Kim, 1994:64). Explanation thus becomes a matter of the structure and organization of our entire belief system, rather than of the content of the propositions which constitute it; a "holistic" process, which makes it 'impossible to say anything precise and useful about the epistemic gain to be associated with individual explanations' (Kim, 1994:64).

As a result, unificationist treatments of explanation also face difficulty in accounting for the sense in which explanation and understanding come in degrees. On Kitcher's model, for example, explanatory power cannot be determined locally: 'explanations of individual events and actions seem to make no sense within such a picture' (Kim, 1994:65). Rather, a derivation is explanatory only if it belongs to the set which 'collectively provides the best systematization of our beliefs' (Kitcher, 1989:430). However, set membership is an all-or-nothing notion. A given derivation is therefore either: a member of the set of that best unifies

our belief system, in which case it can provide understanding; or it is not a member of this set, in which case it cannot provide understanding.

This is not to say, however, that these unificationist accounts of explanation get *everything* wrong. Indeed, Kim is sympathetic to the idea that unification is a central function of our explanatory practices, although, he suggests that the unifying effect of explanations can be seen at work in *both*, our belief system, and in the world. In order to capture this worldly sense of unification, Kim argues that we require an account of explanation which not only connects propositions within our body knowledge, but also the "objective correlates" of such propositions, which lie *outside* of it (1994:56).

IV. Understanding as Knowledge of Causes

According to Salmon and Lewis, what warrants the use of a proposition *G*, in explaining an event *e*, is that the event which *G* represents, *g*, is a *cause* of *e*. Lewis (1986), for example, argues that *every* explanatory claim is a causal claim, and an explanation is informative precisely insofar as it is informative about the causal history of *e*. Similarly, Salmon suggests that 'causal processes, causal interactions, and causal laws provide the mechanisms by which the world works; to understand why certain things happen, we need to see how they are produced by these mechanisms' (1984:132).

One problem is that it is far from obvious just *how* understanding "springs" from our ability to 'discover and formulate causal judgements of the sort that Salmon [and Lewis] would consider explanatory' (Kim, 1994:61). Why, to use an example of Kim's, does knowledge that Socrates' death was caused by his drinking hemlock, promote understanding, when knowledge that this event took place in 399BCE, does not? What we need is an account

of why some knowledge has a distinctive explanatory character, purely in virtue of its content; an account which neither Salmon, nor Lewis, provide.

As I mentioned at the outset, however, recent interest in Kim's (1994) work has focused upon his ancillary claim that causation is just one of many explanation-grounding relations knowledge of which constitutes understanding. This presents explicitly causal accounts of explanation with another problem; how to account for *non*causal explanations. Causation, Kim argues, is 'one type of dependence, obviously one of central importance', however, he goes on to highlight *mereological* dependence as '[a]nother dependence relation, orthogonal to causal dependence and equally central to our scheme of things' (1994:67).

While Kim appears to take causal and mereological forms to be the most prolific types of dependence serving as "explanation-grounding" relations, he does consider others: Xantippe's widowhood appears to depend upon Socrates' death; mental states are widely regarded as depended upon the physical nature of the brain; and evaluative and normative properties seem to depend upon factual, or nonevaluative properties. In each case, Kim argues, the existence of a *non*causal dependence relation serves to 'generate explanations' (1994:68).

These days, such explanation-generating noncausal dependence relations are often thought to be united under the umbrella of "grounding". As Jonathan Schaffer summarizes the connection between grounding and explanation: '[g]rounding connects the more fundamental to the less fundamental, and thereby backs a certain form of [metaphysical] explanation' (2012:122).[9] Indeed, echoing Kim, it has recently been argued that the grounding relation is necessary in formulating physicalism (e.g., Schaffer 2009; Rosen 2010;

[9] Kim similarly argues that just as causal dependence 'rationalizes our attempt to look for the diachronic, temporally antecedent determinants of phenomena', mereological dependence rationalizes of attempt to 'search for their synchronic micro-determinants' (Kim, 1994:67). Also see: Kim (1984/1993:77).

Dasgupta 2014; Ney 2016), and may well be "indispensable" for normative theorizing more broadly (Berker 2018, 2019;). What's more, understanding is also often taken to be the epistemic aim of metaphysical explanation.[10]

So, while Salmon and Lewis are right to tie understanding to knowledge of the "objective correlates" of explanations, they are mistaken in thinking that causation exhausts the scope of such relations. In their defence, however, Kim (1994) argues that these externalist theories make better sense of the unificatory role of explanation than the internalist approaches of Friedman and Kitcher, despite also failing to account for the distinction between explanatory and descriptive knowledge.

We think of the world, according to Kim, 'as a system with structure, not a mere agglomeration of unconnected items, and much of the structure we seek comes from the pervasive presence of dependence relations' (1994:68). The "ontological contribution" of dependence relations like grounding and causation lies precisely in their role of reducing the number of independent events, states, facts and properties which we need to recognize as fundamental, or brute. In this sense, Kim maintains that '[u]nity and structure go hand in hand; dependence enhances unity by generating structure' (1994:68).

It is at this point, however, that Kim appears to give up: '[h]ow unification or simplification is to be connected to explanatory understanding is a difficult question, and I will have nothing useful to add to what has already been said by Friedman, Kitcher and others' (1994:67). When it comes to bridging the gap between the "objective correlates" of explanation, unity and understanding, Kim concedes that 'we lack an understanding of understanding that is rich and clear enough for useful theorizing' (Kim, 1994:69).

[10] See e.g., Thompson (2016); Dasgupta (2017); Maurin (2017); Schaffer (2017); Dellsén (2018).

So far then, we have seen Kim (1994) argue that philosophical *theories* of explanation ought to be centrally concerned with distinguishing explanatory knowledge from descriptive knowledge. Which is to say, a philosophical theory of explanation ought to account for the sense in which some knowledge plays a distinctive explanatory role, purely in virtue of its content. Despite this, a close examination of extant accounts of explanation reveals that this is not typically what we find.

While providing such an analysis of the connection between knowledge and understanding is a key task for a theory of explanation, it is not the *only* task according to Kim. In characterizing the distinction between explanatory and descriptive knowledge, a philosophical theory of explanation also needs to be able to account for three additional aspects of understanding.

These are: the sense in which explanatoriness, and thus understanding, come in degrees; the role of *non*causal dependence relations in generating explanations; and the unificatory element to explanation, facilitated by dependence relations. In the final three sections of this paper (sections VI, VII and VII) we show that interventionism can meet each of these requirements, and thus provides a successful account of the connection between explanatory knowledge and understanding. First, however, we argue that on an interventionist picture, the distinctive role of explanatory knowledge lies in its connection to *manipulability.*

V. Explanatory Knowledge as Knowledge of Manipulability

Like Kim, Woodward (2003) argues that by characterizing explanation in terms of the law-based derivations, both the DN and unificationist theories fail to adequately distinguish between explanatory and “merely descriptive” knowledge (Woodward, 2003:31; 2003:636). The issue here for Kim (1994) concerns the idea that explanation is entirely a matter of the

"logico-linguistic" relations between descriptions of events; factors internal to an epistemic system.

This methodology fails to adequately tie understanding to knowledge of the objective correlates of explanation. Such correlates are, according to Kim (1994), the dependence relations between the events themselves, which underpin the explanatory relation between their descriptions. Hitchcock & Woodward similarly argue that this difficulty in distinguishing between explanatory and descriptive knowledge arises as a result of prior theories failing to tell us precisely what the explanandum 'depends upon' (2003a:18).

Successful explanation, according to Woodward, ought to enable us to see 'what sort of difference it would have made for the explanandum if the factors cited in the explanans had been different in various possible ways' (2003:11). In order to do so, however, an explanation must be *invariant* under testing interventions, which is to say, it must 'describe a relationship which holds for certain *hypothetical* values of *X* and *Y possessed by the very object o*… where *X* is changed by an intervention' (Hitchcock & Woodward, 2003a:20).

Woodward (2003) is also explicit that a philosophical theory of explanation ought to be able to account for the fact that some knowledge plays a special explanatory role purely in virtue of its content. He argues that any such theory ought to be able to explain how, and why, causal knowledge is of practical use: what it enables 'us (or other animals) to achieve with respect to such practical goals as survival and reproduction that other kinds of knowledge does not' (Woodward, 2003:30).

In answer to this question, Woodward proposes that the distinct benefit of explanatory knowledge has to do with our ability to manipulate and control the world around us. An intervention serves as something like 'a hypothetical or counterfactual experiment that shows us that and how manipulation of the factors mentioned in the explanation… would be a way

of manipulating or altering the phenomenon to be explained' (Woodward, 2003:11). It is, furthermore, 'only if a generalization is invariant under testing interventions that it conveys information about how one variable depends upon another' (Hitchcock & Woodward, 2003a:19).

In this sense, Hitchcock and Woodward follow Salmon and Lewis in taking the objective correlates of explanation, relations of dependence, to be '"out there" in nature' (Woodward, 2003:23). Unlike Salmon and Lewis, however, the interventionist can give a compelling account of why some knowledge has a distinctive explanatory character, purely in virtue of its content. For example, our knowledge that Socrates' death was caused by his drinking hemlock promotes understanding *precisely* because we would expect certain testing interventions upon this event to result in a different outcome. For example, had Crito intervened and knocked the hemlock to the floor, Socrates' life would have been saved.

On the other hand, knowledge that the event in question took place in 399BCE does not promote understanding, because the year in which Socrates consumed the hemlock has no bearing upon the outcome. Had Crito intervened to have Socrates granted a one-year stay of execution, this event would have taken place in 398BCE, yet the outcome remains the same. All else being equal, Socrates would still have consumed the hemlock and died. Given that this intervention upon the year in which Socrates' drank hemlock does not allow us to manipulate the outcome, knowledge that the event took place in 399BCE is not explanatory and thus does not contribute to our understanding of why Socrates died.

According to the traditional view, S *understands* why *e* occurs where S knows that *G* explains *E*. To know that *G* explains *E*, on an interventionist account, is to know that it is possible to intervene upon *g*, in such a way that changes the value of *e*. In other words, to understand *e*, is to possess a proposition equivalent to *"E because G",* which constitutes

knowledge that *e* depends upon *g*. In Kim's terminology, the 'objective relation connecting events, *g* and *e*, that grounds the explanatory relation between their descriptions, *G* and *E*' is one of dependence, and such dependence relations become known to us through the implementation of testing interventions (1994:56).

The interventionist account of explanation, then, looks a lot more like an account of a type of explanatory knowledge than those considered (and rejected) by Kim (1994). Indeed, unlike the accounts of accounts of Hempel (1965), Friedman (1974), Kitcher (1981), Salmon (1984) and Lewis (1986), interventionism is crystal clear on the distinction between explanatory knowledge and merely descriptive knowledge: explanatory knowledge enables us to 'distinguish those relationships that are potentially exploitable for the purposes of manipulation and control from those that are not' (Woodward, 2003:36). With this first challenge out of the way, in the next section, we examine the notion of explanatory depth available on an interventionist account and show how it can be used to characterise the sense in which explanation and understanding come in degrees.

VI. Degree of Understanding as Explanatory Depth

Once again echoing Kim, Hitchcock & Woodward (2003a, 200b) argue that both the DN and unificationist theories make explanatoriness an all-or-nothing affair and thus fail to account for the obvious sense in which understanding is a matter of degree. Traditionally, the laws needed to derive an explanation are taken to be *universal* or *exceptionless* generalizations. However, according to Hitchcock & Woodward, this distinction results in an "exhaustive dichotomy" of true generalizations: 'a true generalization is either a law, in which case it is explanatory, or it is accidental, in which case it is not explanatory. There are no other possibilities' (2003b:183).

As such, it becomes difficult to see how one generalization could be any more or less explanatory than another. On the DN model, given the explanans, every law makes the explanandum equally *expectable* (Hitchcock & Woodward, 2003a:18-21). On top of this issue, in taking only those derivations comprising the "best" systematization of our beliefs to be explanatory, unificationist accounts conclude that 'only our deepest, most unified theories are explanatory at all; everything else is non-explanatory' (Hitchcock & Woodward, 2003b:194).[11]

However, on an interventionist account, generalizations need not be exceptionless, or universal to be explanatory. Rather, by characterizing explanatory power in terms of invariance, Hitchcock and Woodward maintain that '[a]mong those generalizations that are invariant, some will be more invariant than others, and they will correspondingly provide *deeper* explanation' (2003b:183-184). For one explanation to be to be deeper than another, then, is for it to answer a greater range of *what-if-things-had-been-different-questions* concerning a given target object or system.

In the previous section, we suggested that Socrates' having consumed hemlock *explains* his death, in light of the possibility of using the former event to manipulate the latter: where Crito knocks the hemlock to the floor, Socrates survives. However, *hemlock* consumption is by no means the only explanation of this event available; nor is it the deepest. Consider, as another possible explanation, Socrates' having consumed *coniine*; the poisonous substance found in plants like hemlock, fool's parsley, and the yellow pitcher plant.

On Hitchcock & Woodward's account, an explanation of Socrates' death which references the fact that he consumed coniine will be *deeper* than an explanation of Socrates death which merely references the fact that he consumed hemlock. To see why imagine that,

[11] We shall have more to say about this in section VIII.

rather than knocking the hemlock from Socrates' hand, Crito now manages to intercept the drinking vassal while the concoction is being prepared and replaces the hemlock with a plant he found in the garden (which is, unbeknownst to him, fool's parsley).

Presumably feeling rather smug at having pulled off this feat of horticultural espionage, one can imagine Crito's confusion when Socrates' dies, nonetheless. Here, hemlock consumption clearly no longer explains Socrates' death; this explanation is not *invariant* under such an intervention. In contrast the fact that Socrates' ingested coniine, the poisonous chemical found in both hemlock and fool's parsley, explains Socrates' death in both scenarios.

What this means is that coniine consumption provides a *deeper* explanation of Socrates' death than hemlock consumption, because the former explanation remains invariant under a wider range of interventions than the latter. Consequently, someone who *knows* that Socrates consumed coniine will be able to answer a wider range of what-if-things-had-been-different question concerning this scenario, than someone who merely knows that he consumed hemlock.[12] While both parties will be able to answer questions like "what would have happened if Crito had knocked the hemlock to the floor?"; only the former will be able to answer questions like "what would have happened had Socrates consumed fool's parsley, or a yellow pitcher plant, instead of hemlock?".

This allows the interventionist to naturally cash out the sense in which explanatoriness, and thus understanding, come in degrees. *S* understands why *e* occurs where *S* knows that *G*

[12] In "Understanding as Knowledge of Causes", Grimm hints at something like this idea in noting that understanding will typically grant the 'ability to answer a variety of what [Woodward] calls "What if things were different?" questions' (2014:334). What's more, Grimm suggests the simple truism that understanding comes in degrees is illustrated by the fact that 'some of us will be able to answer many more of these "What if things had been different?" questions than others (2014:334). However, this idea is not pursued in detail. Also see Grimm (2012).

explains *E*. However, given that explanations vary in depth, or the range of interventions under which they remain invariant, *S*'s understanding of *e* can also vary in depth. The degree of understanding which can be attributed to *S* is thus, on an interventionist picture, directly proportional to the depth of the explanation known to *S*. Someone who knows that Socrates consumed coniine will thus have a better *understanding* of why Socrates died, than someone who merely knows that Socrates consumed hemlock.

In section VIII, we will provide another use for this interventionist notion of explanatory depth. Depth, we will argue, can also be used to characterize the role played by unification in connecting explanation to understanding and, what's more, that an interventionist interpretation of unification does a better job of accounting for this connection than archetypal accounts (e.g., Kitcher 1981). Before we do so, however, in the next section we shall argue that the notion of an intervention can be utilized in making sense of the role of dependence relations in providing the objective correlates for *noncausal* dependence.

VII. Understanding as Knowledge of Causes and Grounds

In section IV, we saw that Salmon (1984) and Lewis (1986) appear to a capture an important element of the connection between explanation and understanding, insofar as they suggest that understanding requires knowledge of the objective correlates of explanation. Despite this, Kim argues that these explicitly causal accounts of explanatory knowledge still face two hurdles. The first is that they fail to explain why some knowledge plays a distinctive explanatory role purely in virtue of its content. In section V, however, we argued that the interventionist analysis of explanation avoids this problem, by tying explanation to dependence through the notion of manipulation.

The second problem facing causal accounts of explanation, according to Kim (1994), is that not *all* explanations are causal in nature. While Kim agrees that understanding requires knowledge of the objective correlates of explanation, he argues that causation is merely one such relation. On the face of it, this appears to present a problem for our claim that interventionism can meet Kim's (1994) demands of a theory of understanding. This is because it is still widely assumed that where it is possible to intervene upon *X*, in such a way that changes the value of *Y*, *X causes Y*.[13]

As was also highlighted in section IV, however, the sort of explanation-generating noncausal dependence relations in which Kim (1994) is interested, are now often taken to be unified under the umbrella of "grounding". One notable trend within the recent grounding literature, has been to characterise the grounding relation explicitly in terms of Woodward's (2003) interventionist analysis of *causal* explanation.[14]

Schaffer (2016, 2017), for example, argues that metaphysical explanation is, like causal explanation, connected to manipulation. This manipulationist role, according to Schaffer, is elucidated through 'Woodward's (2003) guiding conception of explanations as serving to answer "what if things had been different" questions, and Skyrm's (1980:11) idea that wiggling the value of one variable wiggles the value of another' (2017:206). What's more, Schaffer sees this manipulationist aspect as connected to both the unifying feature of explanation, and its role in providing understanding:

[13] In a recent paper, Emmerson (*forthcoming*) labels this position "intervention puritanism" and lists, among its proponents, the likes of: Woodward (2003, 2015); Bokulich (2011); Leuridan (2012); Saatsi & Pexton (2013); Harinen (2014); Pexton (2014); Jansson (2015); Rice (2015); Romero (2015); Baumgartner & Gebharter (2016); Baumgartner & Casini (2017); French and Saatsi (2018); Jansson & Saatsi (2018); Khalifa *et al* (2018, 2020); Reutlinger (2018); Saatsi (2018); Lange (2019).

[14] See, for example, Schaffer (2016, 2017); Wilson (2016, 2018); Reutlinger (2017); Miller & Norton (2021, *forthcoming*).

> ‘Laws are the stable patterns which unify phenomena, provide recipes for manipulation and guide understanding. So I conclude that if there are metaphysical explanations, then there must be counterfactual supporting general principles in the metaphysical realm’ (2017:307).

In section V, we saw Hitchcock & Woodward (2003a) argue that we come by explanatory knowledge through the implementation of testing interventions, and that such knowledge is explanatory in virtue of conveying ‘information about how one variable depends upon another’ (Hitchcock & Woodward, 2003a:19). What Schaffer (2017) is suggesting, in appealing to interventionism in characterising grounding, is that the knowledge imparted as a result of a successful intervention is not always *causal* knowledge. While a successful intervention does indeed serve to highlight the existence of a dependence relation, this fact alone underdetermines whether said relation is one of causation, or one of grounding.

Schaffer is by no means to first to recognize the relation to manipulability shared by both causal and noncausal explanation. Indeed, in “Noncausal Connections” Kim (1974/1993) points to the “bringing about” relation as a unifying feature of both causal and noncausal explanation. Portending his later work, Kim argues that ‘there appear to be dependence relations between events that are not causal’ and that the thesis of universal determinism (“every event has a cause”) ought to be revised to include them ‘if we are to have a clear and complete picture of the ways in which events hang together in this world’ (1974/1993:22).

These observations motivate Kim to put forward the thesis that both causal and noncausal connections might be characterized in terms of a single unifying relation “*R*”: ‘a broad relation of event dependency that subsumes as special cases the causal relation and other dependency relations’ (1974/1993:27). While Kim (1974/1993) does not provide a substantive account of what this relation might be, it is clearly the very same relations “*R*”

which he later refers to as the "explanation-grounding" relation (Kim, 1994). What Kim *does* suggest, however, is that both the causal and noncausal connections, subsumed by "*R*", share a close connection to the "agency relation":

> 'the asymmetry of the agency relation in "We could bring about Xantippe's widowhood by bringing about Socrates' death" points to the asymmetry of the dependency relation between Xantippe's widowhood and Socrates' death. As in the causal case, the asymmetry of the former appears to be rooted in the asymmetry of the latter' (1974/1993:25).[15]

Nicholas Emmerson has recently argued that an interventionist methodology provides Kim's intuitive conception of "bringing about" with a more formal characterization, arguing that 'it is clear that *manipulation* is something like the notion which Kim is intending to highlight with his discussion of the connection between agency and dependence' (2021:3). Even Woodward's position on the topic of noncausal explanation appears to have softened of late.[16] He notes, for example, that 'there is an obvious sense in which it is true that by manipulating whether or not Socrates dies, one can alter whether Xantippe is a widow' (Woodward, 2018:121).

What this suggests, we believe, is that interventionism is best understood as a methodology for characterizing explanatory dependence, *simpliciter*; as opposed to a methodology for characterizing *causation*, or causal dependence, specifically. In the case of metaphysical explanations, successful interventions function to provide knowledge of

---

[15] Kim is perhaps most explicit about his manipulationist sympathies in his 1984 paper "Concepts of Supervenience", where he argues that 'the idea that "real connections" exist and the idea that the world is intelligible and controllable are arguable equivalent ideas', going on to suggest that it is in virtue of such connections that 'the world can be made intelligible; and by exploiting them we are able to intervene in the course of events and alter it to suit our wishes' (1984/1993:53).

[16] While Woodward appears to have been cognizant of the existence of noncausal explanations, he had previously argued that these cannot be given a manipulationist interpretation: '[w]hen a theory or derivation answers a what-if-things-had-been-different question but we cannot interpret this as an answer to a question about what would happen under an intervention, we may have a noncausal explanation of some sort' (2003:221). Also see Woodward (2015).

grounds; and in the case of causal explanation, successful interventions function to provide knowledge of causes.

On this interpretation, interventionism is able to overcome both of the problems identified by Kim with respect to the explicitly causal accounts of Salmon and Lewis. In the first instance, interventionism explains why some knowledge plays a distinctive explanatory role: such knowledge relays information about the possible manipulability of the phenomenon in question. And, in the second, this notion of manipulability can be utilized in characterising explanations resulting from the existence of both causal and grounding relations. Of course, one final hurdle remains.

VIII. Unification as Explanatory Depth:

In a recent paper, David Kovacs criticises grounding theorists for having become overly fixated upon the idea that metaphysical explanations 'track objective, worldly order', while ignoring their epistemological features; the sense in which they 'increase understanding, make the phenomena intelligible, satisfy our curiosity etc' (2020:1661). Indeed, Kovacs echoes Kim's charge against Salmon and Lewis, in arguing that 'it's not transparent what it is about backing relations that yields understanding, or, if you prefer, answers to why-questions' (2020:1672).

On the former point, we agree. This paper and Kovacs' (2020) share a core goal: to put scientific and metaphysical explanation back on an equal epistemological footing. Where we *dis*agree, however, is upon the best way to accomplish this task. In the second half of this paper, we have attempted to make clear the role played by "backing relation", or "objective correlates" of explanation in yielding understanding. According to Kovacs (2020), however,

rather than getting clear on exactly how such relations contribute to our understanding, we should simply abandon the idea that they have a role to play in promoting understanding altogether.

Instead, Kovacs argues that it is Kitcher's (1981, 1989) unificationist theory which provides the best hope of re-establishing 'the unduly neglected link between explanation and understanding in the metaphysical realm' (2020:1663). Following Kitcher, Kovacs (2020) claims that *explanatory* knowledge is knowledge belonging to the *most unified* set. Metaphysical explanation is thus a holistic affair, providing a "global" form of understanding which helps us to 'see how a large number of phenomena are the consequences of a small number of basic facts, from which they can be derived using relatively few and similar patterns of basic derivation' (Kovacs, 2020:1673).

However, Kovac's criticism of grounding theorist for having 'a lot to say about the worldly aspect of explanation but much less about its epistemological aspects' (2020:1962), loses its bite once one remembers that these seemingly disparate aspects of explanation are actually closely connected. As Kim (1994) argues, the ontological contribution of dependence relations like causation and grounding lies precisely in their unificatory role; reducing the number of phenomena which we need to recognize as fundamental. What's more we believe that an interventionist analysis of explanation is able to provide a more convincing account of the unificatory power of explanatory knowledge, *precisely* because of its focus upon the illumination of such worldly relations.

While both unificationism and interventionism share a commitment to the idea that successful explanation ought to apply to a number of different cases, they differ with respect to what they take the relevant cases to be. On Kitcher's (1981, 1989) view, this range of cases are understood in terms of *scope*: the set of objects or systems that fall under the antecedent

of a given law. As Hitchcock & Woodward describe this notion, for a generalization like "all *A*'s are *B*'s" to be explanatory, 'it must 'support' counterfactuals of the following form: if some object *o** that is different from *o* and does not possess property *A* were to be and *A*, then it would be a *B*' (2009a:19).

In this sense, the greater the range of *other objects* to which an explanation applies, the more unifying it is, and thus the greater the understanding which it provides. In contrast, interventionists are not concerned with the range of *other objects* for which a given explanation holds, but rather with the range of changes to the *actual object* being explained. Let's return to the example from section V to see this distinction play out. If you recall, we considered two possible explanations of Socrates' death: (a) Socrates' having ingested *hemlock*; and (b) Socrates' having ingested coniine.

While both the fact that Socrates consumed hemlock, and the fact that he consumed coniine, explain the *actual* event of his death, we argued that the latter provides a deeper explanation. As a result, we maintain that someone who knows that Socrates consumed coniine will have a greater understanding of why he died, than someone who merely knows that Socrates consumed hemlock. However, despite the obvious sense in which (a) provides less understanding than (b), on a unificationist account like that endorsed by Kovacs (2020), however, these explanations are mischaracterized as being equally unifying.

The problem here, lies with Kitcher's (1981) interpretation of unification in terms of scope. Consider, for example, a scenario in which it is Crito, rather than Socrates, who consumes the hemlock. In this scenario, we would expect Crito to meet the same unfortunate end as Socrates. It is in this sense, on a unificationist account, that (a) can be said to explain Socrates death: because were anyone else to consume hemlock, we would expect them to die as well. The problem is that the same is true of (b): had one of us, or you, or Crito, consumed

coniine, death would be the likely result. Indeed, all else being equal, there is no object with which Socrates can be replaced that will allow us to draw a meaningful distinction between (a) and (b) in terms of scope, and thus, unification.

For this reason (among others), Hitchcock & Woodward (2003b) argue, that we ought to adopt an interventionist account of explanation, where unification is understood in terms of explanatory depth. In the present case, that Socrates consumed coniine provides a deeper explanation of his death, because it remains invariant under a wider range of testing interventions. The appeal to coniine ingestion here thus allows us to unify *all* the scenarios in which (b) is invariant (including hemlock and fool's parsley consumption), under a single explanation. Of course, this notion of unification looks rather different to that advocated by Kovacs and Kitcher.

Indeed, an interventionist analysis vindicates many of the features which Kim deems central to an account of unity. For starters, it is *local*; a feature of the propositions and events involved in individual explanations, rather than a *holistic* or *global* feature of our whole belief system (Kim, 1994:68). We know that coniine consumption unifies Socrates' having ingested hemlock, and Socrates' having ingested fool's parsley, without needing to consult our whole belief system. Such wide-ranging consultation, as Kim points out, 'isn't something we do or need to do; it probably isn't something that any of us *can* do!' (1994:65).

What's more, the interventionist analysis takes unity to be a feature 'of events and facts of the world as well as of our beliefs and propositions' (Kim, 1994:68). As Kim puts it, the world is no 'mere agglomeration of unconnected items', rather, we typically think of it as having structure, structure which 'comes from the pervasive presence of dependence relations' (1994:68). This is a central benefit of the interventionist analysis of explanation; as

was argued in section VI, interventions themselves serve to provide us with knowledge of such relations.

An agent can be said to understanding why *e* occurs, where they possess an explanation of *e* equivalent to "*E because G*". On the interventionist picture, to possess such an explanation, is to know (roughly) that it is possible to intervene upon *g*, in such a way that changes the value of *e*. In other words, the content of explanatory knowledge concerns a dependence relation which holds between *e* and *g*; the worldly events (or properties, states, facts etc) represented by *E* and *G*.

Concluding remarks:

At the outset of this paper, we identified a curious puzzle arising from a common assumption among philosophers of science. Despite the prevalence of the view that understanding is merely explanatory knowledge, philosophical theories of explanation themselves typically fail to provide a substantive account of how the latter gives rise to the former. It has been our aim to rectify this situation. We have argued that, when evaluated explicitly with respect to how successfully they account for the connection between explanatory knowledge and understanding, it is the interventionist analysis of explanation which emerges as a clear frontrunner.

First and foremost, interventionism provides a satisfying account of the contribution of explanatory knowledge to understanding; how it differs from merely descriptive knowledge. Explanatory knowledge is distinctively valuable insofar as it enables us to distinguish those relationships which are (at least potentially) exploitable for the purposes of manipulation, from those that are not. What's more, interventionism allows to us accurately characterize the

sense in which explanatoriness and understanding come in degree, through the notion of *explanatory depth*. The degree of understanding which can be attributed to *S*, is directly proportional to the depth of the explanation known to *S*.

We have also argued that interventionism is able to account for the role of noncausal dependence relations in promoting understanding. On our preferred interpretation, interventions function to provide knowledge of cause, in the case of causal explanation, and knowledge of grounds, in the case of metaphysical explanation. Last, but by no means least, we have shown how an interventionist theory is able to articulate the role of unification in explanation and understanding. The resulting notion is both *local* (applying to individual explanations), as well as being a feature of events and facts in the world (not just our beliefs about it).

References

Audi, P. (2012). A clarification and defences of the notion of grounding. In Fabrice Correia & Benjamin Schneider (eds) *Metaphysical Grounding: Understanding the Structure of Reality*. Cambridge: Cambridge University Press. 101-121.

Baumgartner, M. & Gebharter, A. (2016). Constitutive Relevance, Mutual Manipulability, and Fat-Handedness. *British Journal for the Philosophy of Science* 67 (3):731-756.

Baumgartner, M. & Casini, L. (2017). An Abductive Theory of Constitution. *Philosophy of Science* 84 (2):214-233.

Berker, S. (2018). The Unity of Grounding. *Mind* 127 (507):729-777.

Bird, A. (2007). What is scientific progress? *Noûs 41 (1):64-89.*

Bokulich, A. (2011). How scientific models can explain. *Synthese* 180 (1):33-45.

Bunge, M. (1973). Philosophy of Physics. *Journal for General Philosophy of Science/ Zeitschrift für Allgemeine Wissenschaftstheorie* 4 (2):407-409.

Dasgupta, S. (2014). The Possibility of Physicalism. *Journal of Philosophy* 111 (9-10):557-592.

Dasgupta, S. (2017). Constitutive Explanation. *Philosophical Issues* 27 (1):74-97.

Dellsén, F. (2018). Beyond Explanation: Understanding as Dependency Modelling. *British Journal for the Philosophy of Science* (4):1261-1286.

Emmerson, N. (*forthcoming*). A Defence of Manipulationist Noncausal Explanation: The Case for Intervention Liberalism. *Erkenntnis* https://doi.org/10.1007/s10670-021-00497-4.

French, S. & Saatsi, J. (2018). Symmetries and Explanatory Dependencies in Physics. In Alexander Reutlinger & Juha Saatsi (eds) *Explanation Beyond Causation: Philosophical Perspectives on Non-Causal Explanations*. Oxford: Oxford University Press. 185-205.

Friedman, M. (1974). Explanation and Scientific Understanding. *Journal of Philosophy* 71 (1):5-19.

Gijsbers, V. (2013). Understanding, explanation, and unification. *Studies in History and Philosophy of Science Part A* 44 (3):516-522.

Grimm, S., R. (2014). Understanding as Knowledge of Causes. In Fairweather, A. (ed.), *Virtue Epistemology Naturalized*.

Grimm S., R. (2014) Understanding as Knowledge of Causes. In Fairweather A. (eds) *Virtue Epistemology Naturalized: Bridges Between Virtue Epistemology and Philosophy of Science.* Synthese Library (Studies in Epistemology, Logic, Methodology, and Philosophy of Science). Berlin: Springer. (366):329-345.

Harinen, T. (2014). Mutual manipulability and causal inbetweenness. *Synthese* 195 (1):35-54 (2018).

Hempel, C. (1965). *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science.* New York: The Free Press.

Hempel, C. (1966). *Philosophy of Natural Science*. Prentice-Hall Foundations of Philosophy Series. New Jersey: Prentice-Hall.

Hempel, C. (1983/2001). Valuation and objectivity in Science. In Fetzer, J., H. (ed.), *The Philosophy of Carl G. Hempel*. New York: Oxford University Press. 372-396.

Hitchcock, C., & Woodward, J. (2003a). Explanatory Generalizations, Part I: A Counterfactual Account. Noûs 37 (1):1-24.

Hitchcock, C., & Woodward, J. (2003b). Explanatory Generalizations, Part II: Plumbing Explanatory Depth. Noûs 37 (2):181–199.

Humphreys, P. (1989). *The Chance of Explanation: Explanation in the Social, Medical, and Physical Sciences.* Princeton: Princeton University Press.

Jansson, L. (2015). Explanatory Asymmetries: Laws of Nature Rehabilitated. *Journal of Philosophy* 112 (11):577-599.

Jansson, L., & Saatsi, J. (2018). Explanatory Abstractions. *British Journal for the Philosophy of Science* 70 (3):817-844.

Kelp, C. (2015). Understanding phenomena. *Synthese* 192 (12):3799-3816.

Kelp, C. (2017). Knowledge First Virtue Epistemology. In Carter, A., Gordon, E., & Jarvis, B. (eds.), *Knowledge First: Approaches in Epistemology of Mind*. Oxford: Oxford University Press. 233-245.

Khalifa, K. (2017). Understanding, Explanation, and Scientific Knowledge. Cambridge: Cambridge University Press.

Khalifa, K., Millson, J., & Risjord, M. (2018). Inference, explanation, and asymmetry. *Synthese*: https://doi.org/10.1007/s11229-018-1791-y.

Khalifa, K., Doble, G., & Millson, J. (2020). Counterfactuals and Explanatory Pluralism. *British Journal for the Philosophy of Science* 71 (4):1439-1460.

Kim, J. (1974/1993). Noncausal Connections. In Kim, J. (ed.), *Supervenience and Mind.* Cambridge: Cambridge University Press. 22-32.

Kim, J. (1984/1993). Concepts of Supervenience. In Kim, J. (ed.), *Supervenience and Mind.* Cambridge: Cambridge University Press. 53-78.

Kim, J. (1994). Explanatory knowledge and metaphysical dependence. *Philosophical Issues* 5:51-69.

Kitcher, P. (1981). Explanatory Unification. *Philosophy of Science* 48 (4):507-531.

Kitcher, P. (1989). Explanatory Unification and the Causal Structure of the World. In Philip Kitcher & Wesley Salmon (eds.), *Scientific Explanation*. Minneapolis: University of Minnesota Press. 410-505.

Kovacs, D. (2020). Metaphysical Explanatory Unification. *Philosophical Studies* 177 (6):1659-1683.

Kvanvig, (2009). The value of understanding. In Pritchard, D., Haddock, A., &Millar, A. (eds.), *The Value of Understanding*. Oxford: Oxford University Press. 95-112.

Lange, M. (2019). Asymmetry as a challenge to counterfactual accounts of non-causal explanation. *Synthese* 198 (4):3893-3918.

Lewis, D. (1986). Causal Explanation. In Lewis, D. (ed.), *Philosophical Papers II*. Oxford: Oxford University Press. 214-240.

Leuridan, B. (2012). Three Problems for the Mutual Manipulability Account of Constitutive Relevance Mechanics. *British Journal for the Philosophy of Science* 63 (2):399-427.

Lipton, P. (2003). Inference to the Best Explanation. New York: Routledge.

Lipton, P. (2009). Understanding without explanation. In de Regt, H., Leonelli, S., & Eigner, K. (eds.), *Scientific Understanding: Philosophical Perspectives.* Pittsburgh: University of Pittsburgh Press.

Maurin, (2018). Grounding and metaphysical explanation: it's complicated. *Philosophical Studies* 176 (6):1573-1594.

Miller, (1987). *Fact and Method: Explanation, Confirmation and Reality in the Natural and the Social Sciences.*

Miller, K. & Norton (2021). *Everyday Metaphysical Explanation*. Oxford: Oxford University Press.

Miller, K. & Norton, J. (*forthcoming*). Non-cognitivism about Metaphysical Explanation. *Analytic Philosophy*.

Ney, A. (2016). Grounding in the Philosophy of Mind: A Defense. In Aizawa, K., & Gillett, C. (eds.), *Scientific Composition and Metaphysical Grounding*. London: Palgrave-Macmillan. 271-300.

de Regt, H., W. (2017). *Understanding Scientific Understanding*. Oxford: Oxford University Press.

Reutlinger, A. (2017). Does the Counterfactual Theory of Explanation Apply to Non-Causal Explanation in Metaphysics? *European Journal for Philosophy of Science* 1-18.

Reutlinger, A. (2018). Extending the Counterfactual Theory of Explanation. In Alexander Reutlinger & Juha Saatsi (eds) *Explanation Beyond Causation: Philosophical Perspectives on Non-Causal Explanations*. Oxford: Oxford University Press. 74-95.

Rice, C. (2015). Moving Beyond Causes: Optimality Models and Scientific Explanation. *Noûs* 49 (3):589-615.

Romer, F. (2015). Why there isn't inter-level causation in mechanisms. *Synthese* 192 (11):3731-3755.

Rosen, G. (2010). Metaphysical Dependence: Grounding and Reduction. In Hale, B. & Hoffmann, A. (eds.), *Modality: Metaphysics, Logic, and Epistemology*. Oxford: Oxford University Press. 109-135.

Saatsi, J. & Pexton, M. (2013). Reassessing Woodward's Account of Explanation: Regularities, Counterfactuals, and Noncausal Explanations. *Philosophy of Science* 80 (5):613-623.

Salmon, W. (1984). *Scientific Explanation and the Causal Structure of the World.* Princeton: Princeton University Press.

Salmon, W., C. (1989). 4 decades of scientific explanation. *Minnesota Studies in the Philosophy of Science* 13:3-219.

Skyrm, (1980). Causal Necessity: A Pragmatic Investigation of the Necessity of Laws. New Haven: Yale University Press.

Schaffer, J. (2009). On what grounds what. In Manley, D., Chalmers, D., & Wasserman, R. (eds.), *Metametaphysics: New Essays on the Foundations of Ontology*. Oxford: Oxford University Press. 347-383.

Schaffer, J. (2012). Grounding, transitivity, and contrastivity. In Correia, F., & Schneider, B. (eds.), *Metaphysical Grounding: Understanding the Structure of Reality*. Cambridge: Cambridge University Press. 122-138.

Schaffer, J. (2016). Grounding in the image of causation. *Philosophical Studies* 173 (1):49-100.

Schaffer, J. (2017). Laws for Metaphysical Explanation. *Philosophical Studies* 27 (1):303-321.

Strevens, (2008). *Depth: An Account of Scientific Explanation.* Cambridge, MA: Harvard University Press.

Strevens, (2013). No understanding without explanation. *Studies in History and Philosophy of Science Part A.* 44 (3):510-515.

Thompson, N. (2016). Metaphysical Interdependence. In Jago, M. (ed.), *Reality Making*. Oxford: Oxford University Press.

Trogdon, K. (2013). An Introduction to Grounding. In Hoeltje, M., Schnieder, B., & Steinberg, A. (eds.), *Varieties of Dependence.* Munich: Philosophia Verlag. 97-122.

Trogdon, K. (2018). Grounding Mechanical Explanation. *Philosophical Studies* 175 (6):1289-1309.

Trogdon, K., & Skiles (2021). Should explanation be a guide to ground? *Philosophical Studies* 178 (12):4083-4098.

Trout, J., D. (2002). Scientific explanation and the sense of understanding. *Philosophy of Science* 69 (2):212-233.

Wilkenfeld, (2017). MUDdy understanding. *Synthese* 194 (4):1273-1293.

Wilson, A. (2016). Grounding Entails Counterpossible Non-Triviality. *Philosophy and Phenomenological Research* 92 (3):716-728.

Wilson, A. (2018). Metaphysical causation. *Noûs* 50 (4):1-29.

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

Woodward, J. (2015). Interventionism and Causal Exclusion. *Philosophy and Phenomenological Research* 91 (2):303-347.

Woodward, J. (2018). Some Varieties of Non-causal Explanation. In Alexander Reutlinger & Juha Saatsi (eds) *Explanation Beyond Causation: Philosophical Perspectives on Non-Causal Explanations*. Oxford: Oxford University Press. 117-141.