Grounds and Moral Laws

Alexios Stamatiadis-Bréhier

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

School of Philosophy, Religion and History of Science

September 2021

To Yann Lauret (mon parrain). In loving memory.

Acknowledgements

I am deeply grateful to my supervisors, Pekka Väyrynen and Jack Woods, for their continuous and tireless assistance with this project. They are one of the nicest and smartest people I have met. Their professionalism is exemplary, and they have made me a better philosopher. This extends to my Viva examiners: Robbie Williams and Gideon Rosen.

Finally, I owe most of my metaethical background to Stelios Virvidakis who (among *many* other things) introduced me to Richard Boyd's seminal 1988 paper during the 2015 metaethics seminar at the History and Philosophy of Science Department at the University of Athens. His seminars and overall guidance were one of the main reasons I decided to attempt to pursue an academic career in philosophy. I also owe much of my background and interest in the metaphysics of science to Stathis Psillos.

I am grateful to the Onassis Foundation for supporting me for the duration of my PhD studies.

I am also lucky to have met an array of incredible people during my time in Leeds. These include: Ludovica Adamo, Thomas Brouwer, Andreas Bruns, Radu Bumbăcea, Arthur Carlyle, Adina Covaci, Paola Chapa, Alec Siantonas, George Stamets, Aleksander Domosławski, Zack Gudmunsen, Kevin Jones, Tadhg Ó Laoghaire, Sam Mason, Simon Newey, Nicholas Reimann, Jyl Schacher, Alec Siantonas, Alison Toop, Rodrigo Valencia.

Special shoutout to Pei-Lung Cheng and David Heering for numerous discussions on the nature of explanation and meta-philosophy. Shoutout to Dionysia Bouzioti, who has directed some of the best plays I have seen in my life. I am particularly grateful to Miriam Bowen and Olof Leffler for keeping me sane during the pandemic (and for being awesome all around).

Abstract

This dissertation consists in a series of papers on what could be called *moral grounding* (i.e. the application of metaphysical grounding to the moral domain). These papers motivate two distinct, but mutually complimentary, projects. The first concerns the nature of metaphysical grounding itself. In the first paper, I motivate and defend a functionalist theory of grounding (f-grounding). According to f-grounding, a relation R is a grounding relation iff R backs non-causal explanations. In the second paper, I elucidate one of the key components of f-grounding: the relation of *backing*. Specifically, I argue that backing is a form of truthmaking.

The second project concerns the motivation and defence of a neo-Humean theory of moral principles. First, I argue that moral principles have an important role in the definition of moral naturalness (ch. 3). Then, in chapter 4, I motivate NHM by showing that it provides dialectically compelling answers to questions concerning the explanation of pure moral principles as well as the explanation of moral supervenience. In that same paper, I defend the contingentist aspect of moral principles (roughly, the view that there are possible worlds with different pure moral principles). Finally, in the final paper, I defend the *explanatoriness* of moral principles qua moral generalizations. In doing so, I appeal to a pluralist theory of moral explanation.

Contents

Contents

Preface

Thesis summary

This dissertation consists of a series of papers on what could be called *moral grounding*. Specifically, I look into the interaction of metaphysical grounding with the question concerning the metaphysics of moral laws. The papers of this dissertation are organized in a modular way. In this sense, each paper is a self-contained piece. Still, put together, these papers provide the foundation for two distinct but mutually complimentary projects.

The first project concerns the nature of metaphysical grounding itself. In the first chapter, I propose a functionalist theory of grounding (what I call *f-grounding*) according to which grounding is multiply realizable by the relations that fulfil the non-causal explanatoriness role. I show that f-grounding has important virtues, and that it can successfully deal with a recent and important challenge against the existence of grounding. According to this challenge, explanatory unificationism puts pressure on the theoretical utility and existence of grounding. This is because non-causal explanations can be accommodated without appealing to non-causal determination relations. This makes grounding obsolete.

In the second chapter, I examine one of the key components of f-grounding. A relation is an f-grounding relation iff it backs non-causal explanations. But what does it mean to *back* an explanation? This question hasn't received the attention it deserves in the relevant literature. After considering and rejecting some alternatives, I propose that backing should be understood as an instance of truthmaking. I show that backing as truthmaking is the most plausible view given basic assumptions about the nature of explanation. It also avoids various challenges that its rivals face.

The second project concerns the application of metaphysical grounding to certain key metaethical themes. In chapter 3, I examine the notion of moral naturalness. It usually suggested that there is a deep metaphysical difference between moral naturalism and moral non-naturalism. But it is not clear what it really means for a property to be natural in this sense. A recent proposal accounts moral naturalness in terms of grounding. After examining two thin readings of that view, I conclude that such accounts face serious problems. To that end, I propose a novel account according to which a property is natural iff it figures in the nexus of natural laws. To avoid circularity, I appeal to a well-established feature of natural laws: a law is natural iff it is a physical law, or it is appropriately related to physical laws.

The details of how particular moral properties relate to their corresponding natural laws depends on one's background theory concerning the metaphysics of moral principles (which, according to the naturalist, is continuous to one's theory of natural laws). In chapters 4 and 5, I motivate and defend a neo-Humean theory of moral principles. Such a theory is (what I call) *non-substantivist* in nature: it construes moral principles as nothing over and above the instances that fall within their scope (to compare, rival accounts construe moral principles as something *separate* that, in a sense, *governs* particular moral instances). According to the neo-Humean theory, moral principles (in the *genetic,* non-propositional, sense) are *fusions* of particular moral instances which, in turn, supervene on the Humean mosaic (roughly, a mosaic of non-modal, freely recombinable, physical properties).

In chapter 4, I do two things. First, I motivate the neo-Humean (NHM) view by showing that it has at least two important virtues. First, it provides an answer to an important but neglected question: what *explains* pure moral principles? Secondly, in virtue of explaining that first question, it provides a dialectically powerful explanation of moral *supervenience.* Then, I focus and defend the *contingentist* aspect of NHM (i.e.

the view that moral principles are metaphysically contingent). I argue that NHM retains what is philosophically attractive about moral contingentism, while also avoiding the main challenges that its rivals face.

According to NHM, moral principles are mere generalizations of moral phenomena. In chapter 5, I defend the *explanatoriness* of moral generalizations against two important challenges. It has been suggested that moral generalizations are explanatorily idle in terms of the instances they subsume. Rather, moral generalizations are supposed to be explained in terms of their instances. Also, even if we somehow establish the thesis that moral generalizations explain their instances, there is a circularity problem to be faced. Given the *transitivity* of explanation, how could it be that moral generalizations explain the instances in virtue of which they are explained? In response, I propose of pluralist theory of moral explanation which accommodates, but also revises, the relevant notion of transitivity.

Future work

In this dissertation I present and motivate some key aspects of these two projects. In this sense, my goal is exploratory and doesn't aim to present a *complete* theory of either f-grounding or NHM. Still, if my arguments are successful then both projects should be taken seriously and explored further.

Consider how the adoption of f-grounding provides new ways of approaching traditional problems that figure in the grounding literature. For example, there is a question concerning the way grounding-facts are grounded. But if grounding is multiply realizable by different relations, then there will be a distinct "meta-grounding" problem for each individual relation. Also, it is plausible that certain *features* of some putative realizers of f-grounding provide straight-forward, or trivial, solutions to that

problem. If (say) truthmaking is a realizer of grounding, then the problem of what makes true the fact that P makes <P> true doesn't even arise (because the truthmaking-fact "P makes <P> true" is not a truth-apt entity).

Similar implications arise concerning the notion of *fundamentality*. Grounding relations are supposed to be indicators of relative fundamentality. It is usually suggested that if P grounds Q then P is more fundamental than Q. But given the multiplicity of the realizers of f-grounding, it turns out that some grounding relations can point to opposite directions. To take one example from chapter 5, it could be that a moral generalization is grounded in its instances but those very instances ground that same generalization. So it seems that our notion of fundamentality needs to be revised accordingly. This suggests the possibility of a pluralistic theory of fundamentality according to which there are many *distinct* notions of fundamentality relativized in terms of different f-grounding realizers.

Now consider NHM. In my dissertation I defend two key aspects of NHM: the explanatoriness of moral generalizations and the metaphysical contingency of moral principles. But there is more work to be done. Even though NHM is a metaphysical account of moral principles it would still be nice to have an *epistemology* of moral principles. Traditionally neo-Humean accounts of scientific laws are usually coupled with, so-called, *best-system analyses*: genuine laws of nature are supposed to be expressed by lawlike statements which figure in a deductive system governed by desiderata involving simplicity and explanatory strength. It is an open question that warrants further exploration whether NHM should mimic the epistemology of scientific laws or whether it should opt for more traditional accounts in moral epistemology. Finally, there is a technical worry about the way facts about moral determination should be understood in a neo-Humean framework. Determination-talk is usually accounted for

in terms of grounding. But how would facts about grounding be accounted for in terms of the Humean mosaic?

These are difficult and intriguing questions which I hope to address in future work. Still, I hope that in laying the foundations necessary to motivate NHM and f-grounding, I have provided enough reasons to take these projects seriously and pursue the new avenues of research that they create.

Part 1 | On Grounding and Explanation

Chapter 1 | Grounding Functionalism and Explanatory Unification

Abstract: I propose a functionalist theory of grounding: f-grounding. Specifically, I argue that grounding is a second-order phenomenon that is realized by relations that play the non-causal explanatoriness role. I also show that f-grounding can deal with a recent and powerful challenge. Kovacs (2017: sec. 5; 2020: 1) and Baron & Norton (2020: 3) appeal to explanatory unificationism to argue that successful non-causal explanations do not require backers in the form of grounding relations. Against this I argue that a systematization involving f-grounding is superior to its anti-relational counterpart.

1. Introduction

Metaphysicians often say that some things hold in virtue of other things. Chairs exist in virtue of certain arrangements of particles. Mental events exist in virtue of neural events. The normative characteristics of a given state of affairs hold (at least partly) in virtue of the non-normative features of that state. What is particularly striking in these examples is the 'in virtue of' claim. How is it supposed to be understood? Clearly, the relevant connection is non-causal: chairs are not *caused* by the appropriate arrangement of particles. To that end, many philosophers have proposed that 'in virtue

of' should be explicated by appealing to the relation of 'grounding'.[1] According to this picture, what these examples have in common is that they are grounding claims.

Metaphysical grounding is generally understood as a worldly, non-causal, determination relation that is tightly linked to explanation. Also, it is typically understood as metaphysically primitive and unanalyzable (Audi 2012).[2] Against this tradition I defend a reductive, functionalist, account of grounding (call this 'functionalist grounding' or 'f-grounding'). In this sense, a relation R is a grounding relation iff R plays the appropriate functional role. I will understand that functional role, minimally, as the function of backing non-causal explanations (to be elaborated in section 3.1.). The minimality of this function allows for, at least some, paradigmatic 'small-g' relations to figure as its realizers (Wilson 2014). 'Small-g' relations are non-causal determination relations that, according to Wilson, make general ("big-G") grounding relations obsolete (arguably, because many of the roles grounding is supposed to fulfil are already fulfilled by small-g relations). On the current proposal, small-g relations serve as *realizers* of grounding showing that the two kinds of relations are not in competition. Exemplars of such realizers include functional realization, composition, and other non-causal determination relations.[3]

This way of understanding grounding accommodates two important constraints. The first one is what I will call (Unity): different instances of grounding should be objectively unified. The second one is (Pluralism): a theory of grounding

---

[1] I adopt a relational understanding of grounding (Schaffer 2009). I will ignore the operator view as there are translation procedures from non-relational views to relational ones (assuming, of course, some ontological commitments along the way). (e.g. Saenz 2015: 4)

[2] For other conceptions of primitiveness see Dasgupta (2017: 81-2).

[3] For other potential candidates see Bennett (2017) (cf. fn 19). Non-standard small-g relations might include social construction (Ásta 2015), anchoring (Epstein 2015), machretic determination (Gillett 2016), or the 'governing' relation (Wilsch 2020).

should be as diverse as possible in terms of formal features. The resulting picture understands potential realizers of grounding as both significantly diverse and importantly unified.

A very important virtue of my proposal is that, in virtue of meeting those two constraints, it can meet a powerful challenge that every theory of grounding must face. According to this challenge, grounding relations do not exist since they are explanatorily obsolete. Under the unificationist framework, as championed by Kitcher (1981; 1989), explanations are understood holistically without appealing to non-causal determination relations like grounding. I respond to this challenge in section 4.

Still, I should note that the main objective of this paper is to present an initial *sketch* of the functionalist theory by showing how it deals with an important challenge. But there is more work to be done if a theory of f-grounding is to be fully defended. For example, it needs to be shown how f-grounding interacts with well-known issues in the literature, such as questions concerning the epistemology of grounding, or questions concerning grounding's connection to fundamentality (cf. Leuenberger 2020). In this sense, I take it that every theory of grounding (including f-grounding) needs to be developed in a piecemeal fashion. This paper presents an initial gloss and focuses on the unificationist challenge thus showing that f-grounding is a theory of grounding that should be taken seriously.

2. Constraints for a theory of Grounding

Grounding has instances. The singleton set involving Socrates is grounded by Socrates. A given mental state is grounded in its corresponding neural state. If these are genuine cases of grounding, then they must have something substantive in common. Orthodox grounding theorists typically argue that different instances of grounding are

unified in terms of a fixed set of formal features. The proponent of f-grounding has a different story. F-grounding is realized by paradigmatic small-g relations all of which realize the appropriate functional role. In this sense, even if it turns out that such realizers have no first-order features in common, it would still be the case that they are unified in terms of their functional role.

Both accounts have important challenges to face and require scrutiny. Still, both theories agree that for grounding to be a unified phenomenon, it must be shown in what way its different instances are objectively unified. This gives us our first constraint:

> **(Unity)** Different instances of grounding should be objectively similar to one another.

I will say more about the way in which f-grounding meets (Unity) in the next section. Still, why is (Unity) important?

First, one uncontroversial objective of realist metaphysics is to group together phenomena which appear qualitatively different but are objectively similar. In this sense, if the realizers of f-grounding are objectively similar then this is, on its own, a substantive result. Moreover, (Unity) has at least one interesting philosophical consequence. F-grounding is a reductive theory of grounding: it analyzes grounding in terms of something more familiar (the family of its realizers). This is epistemically fruitful. If P reduces to Q, then beliefs about Q can act as constraints for one's beliefs about P and vice versa. To illustrate, consider another reductive theory: ethical reductionism. If the ethical reduces to (say) the pleasurable, then we should expect for some features of the ethical to be revised accordingly (a candidate would, perhaps, be the putative intrinsic motivational force of moral properties).

Analogously, consider a classic problem that arises in the literature: the problem of what grounds grounding-facts (e.g. what grounds the fact that P grounds Q; call this

the meta-grounding problem). There are many proposed solutions (Wallner 2018). However, a reductive theory like f-grounding has important implications for the way the problem is framed. First, it would follow that the meta-grounding problem would be the problem of what f-grounds the fact that P f-grounds Q. If, for example, functional realization is understood as a grounding relation, then this allows for the meta-grounding problem to be restated: "What functionally realizes the fact that certain properties functionally realize higher-order states?".[4]

Arguably, an answer to this question would appeal to a variety of factors including, perhaps, facts about the nature of functions, and so on. This framing also naturally indicates a plausible *methodology* for approaching the meta-grounding problem. If the grounds of f-grounding facts are functional realizers of those facts, then tracing the grounds of grounding-facts would require the identification of the functional role of grounding-facts, and which facts actually fulfil that role. Finally, an f-grounding construal of the meta-grounding problem entails that there are *many* meta-grounding problems. Each realizer of f-grounding would involve a specific meta-grounding problem for *that* relation (e.g. one for constitution, one for social construction, etc.). This indicates that it is very likely that a "one-fits-all" solution is not possible.[5] Of course, all of this is controversial, but the point remains: if grounding

---

[4] Wilson (2014: 568 fn. 80; 2016: sec. 6) has argued that the meta-grounding problem is a 'spandrel' problem. Specifically, she argues that the meta-grounding problem doesn't even arise in the case of small-g relations. I think this move is premature (although not without some merit; see fn. 5). It is true that an analogous problem hasn't appeared in the 'small-g' literature, but this doesn't necessarily compromise the legitimacy of the question itself (especially if one considers all the different ways the meta-grounding problem can be formulated as per Kovacs [2019]).

[5] Some instances of the problem might have trivial solutions. Consider strong emergence as a realizer of f-grounding. Strong emergence-facts are typically brute so, in this case, there would be nothing f-grounding the meta-facts. Other instances of the problem are not even coherent. Truthmaking is a

reduces to something more familiar, then new ways of addressing old problems become available.[6]

Meeting (Unity) is a good thing, but one might worry that this will result in the exclusion of many non-standard cases of grounding. In this sense, (Unity) pulls in a direction that makes it difficult for an account to be pluralistic. Call this constraint, (Pluralism):

(Pluralism) Instances of grounding should be as diverse as possible.

I will focus on *formal* diversity.[7] For example, orthodox theories take every instance of grounding to have the same formal features. The most typical account takes grounding to impose a strict partial order on the entities it relates: every instance of grounding should be transitive, reflexive, and asymmetric.[8] Naturally, this way of understanding grounding excludes cases that do not exhibit such features.[9]

---

plausible realizer of f-grounding. But the question of what makes true a truthmaking-fact doesn't arise since truthmaking-facts are not truth-apt.

[6] Another application could concern the proper modal status of grounding-facts (cf. Skiles 2015).

[7] Since the most obvious feature by which relations can be characterized is their formal profile, my focus on formal diversity is warranted. Also, it seems to me that other kinds of diversity are downstream from formal diversity anyway.

[8] Naturally, there is disagreement even between orthodox grounding theorists about what these features are supposed to be (e.g. Rodriguez-Pereya 2015). But, regardless of their disagreement, what they agree upon is that these formal features stay the same across the board.

[9] Candidates for non-orthodox grounding are numerous and can be easily found in the literature (for an influential defense of this point see Wilson [2014: 569-70, 572]). Cases of reflexive grounding are advanced by Woods (2018) and Bliss (2018). Failures of transitivity are highlighted by Schaffer (2012). Finally, symmetric cases of grounding are accepted by Nolan (2018). Less straightforward rejections of asymmetricity involve what we might call *mixed* cases of grounding (Bennett 2017: 2.5.; Litland 2018) (i.e. cases where two different grounding relations go in opposite directions). The typical example

Building the impossibility of such cases *into* one's definition of grounding is problematic. Many paradigmatic small-g relations exhibit a diverse array of formal features. According to some, realization is non-reflexive (Shoemaker 2007) and non-transitive (Gillett 2016). Similarly, others allow for constitution to be symmetric (Pereboom 2002).[10]

Treating these cases as cases of grounding is perhaps controversial. Still, I will assume that they generate prima facie good reasons to have a relaxed view about the formal features of grounding. After all, recall that grounding is supposed to figure in the same contexts where many paradigmatic small-g relations appear. To my mind, a theory of grounding that can accommodate this plurality of formal features is better than one that cannot, all things considered.

---

involves Schaffer's priority monism (2010) where the parts of an entity hold in virtue of the whole, but the whole is composed by its parts (for a similar view with different relata see Bernstein [2020]). More sophisticated disagreement involves grounding's modal force. Traditionally, instances of grounding are taken to be metaphysically necessary (Rosen 2010), but others take them to be contingent (Skiles 2015). More extreme views highlight grounding cases that are indeterministic (Emery 2019) (cf. Bennett 2017: 3.3.1.).

[10] Similar cases of non-standard determination can be found in the literature on causation (e.g. backward causation) (Dowe 1996). One might worry that these are 'limiting cases'. Limiting cases of (say) realization are cases that are, in a sense, simply a technical by-product of the way the realization relation has been defined. For example, according to Shoemaker (2007), P realizes Q when Q has a subset of the causal powers of P. Contra Wilson (2011), he doesn't impose any constraints on parthood which means that, trivially, P can realize itself (since P can be a part of itself). That being said, I find the notion of a limiting case imprecise and obscure. Kovacs (2017: 5-6) attempts to characterize it but admits that it is difficult to spell out in more precise terms. Still, even if there *are* such limiting cases it doesn't follow that every case of non-standard small-g case is a limiting case. For example, Gillett's (2016) notion of realization is *typically* transitive and doesn't fall under any plausible characterization of a limiting case.

Naturally, (Unity) and (Pluralism) set the bar very high. Attempts to accommodate (Pluralism) make it increasingly more difficult to meet (Unity). The messier the phenomena the more difficult it is for them to be unified. Still, I will argue that my proposal delivers the right result.

3. Functionalist Grounding

I take grounding to be a second-order phenomenon. More specifically, some relations *realize* the grounding relation by having the right function. Typically, a function is specified via some job-description. But making the job-description too detailed risks disqualifying many potential realizers, thus violating (Pluralism).

To illustrate, consider functionalism about causation. Menzies (1996) identifies a list of platitudes about causation and then provides a definition that accommodates them. Some of these platitudes include causes preceding their effects and the relata of causation being events. However, as Psillos (2008: 6) notes, an account based on these platitudes is highly parochial as several conceptually coherent cases of causation are excluded. Similar worries apply to the grounding case. If, for example, the grounding job-description says that the relata of grounding are always facts, then that would exclude relations that are cross-categorical (like the truthmaking relation).[11]

There is a way to go forward. Minimizing the set of platitudes as much as possible yields a version of grounding functionalism that meets both constraints at once. A way to do this is to identify only a *single* platitude. I propose that a relation R is a grounding relation when R has the function to be explanatory in a distinct way. So, the question arises: in what kind of explanation are realizers of f-grounding involved in?

---

[11] Specifically, Rettler (2017: 13-4) falls victim to this sort of charge.

3.1. What kind of explanation?

I will understand explanatoriness in terms of 'backing' (Sjölin Wirling 2020). A relation is explanatory when it 'backs' (i.e. supports) explanations.[12] Grounding theorists typically cash out grounding explanations as *metaphysical* explanations. However, this is unhelpful at best as there is no consensus about what these explanations are supposed to be. Perhaps metaphysical explanations are essentially *synchronic*. But this is restrictive and controversial: there seems to be room for synchronic causation and diachronic grounding (Baron & Miller & Tallant 2020). Similarly, saying that metaphysical explanation is the "ultimate form of explanation" (Fine 2012) hardly seems illuminating.

A more distinctive variety of explanation is *non-causal* explanation. Note that to avoid circularity we cannot define non-causal explanation in terms of grounding (i.e. explanation that is backed by grounding). Rather, non-causal explanation can be defined negatively: take the set of every explanation, then identify the ones that are causal, and the remainder is the set of non-causal explanations. This way we have:

> **(Functionalist Grounding)** A relation R is a grounding relation iff R has the
> function to back non-causal explanations.

This is a promising way of understanding grounding explanations for at least two reasons. First, we know what those explanations are. They are defined negatively in terms of explanation simpliciter and causation. Secondly, (Functionalist Grounding) is specific but not too restrictive. On the face of it, no paradigmatic small-g relations are

---

[12] Kovacs (2017) takes backing to be, itself, the explanation-relation. Sjölin Wirling (2020) identifies backing with grounding (cf. Kovacs 2019: sec. 3).

excluded as most of them routinely figure in explanations.[13] This feature will come into play later when (Functionalist Grounding) is evaluated in terms of (Pluralism).

At this point, my goal is to argue that the realizers of f-grounding form a natural, *functional*, kind. A group of entities forms a functional kind when they exhibit the same functional role. But, insofar as we want the relevant kind to be *natural*, not just any function will do. Functions come cheap. The relevant function must not only be specific and intelligible, but it must also concern a 'deep' feature of reality. To compare, consider the function of a heart to make noise. This function is not 'joint-carving' in any deep sense. If anything is distinctive about the heart's function it surely isn't the fact that it makes noise (rather, it is the fact that it has such-and-such evolutionary and physiological effects). For this reason, philosophers of biology typically use this case as an example of how *not* to understand the notion of a *proper* function (Millikan 1989).

What notion of joint carving-ness is at play here? A classic approach appeals to the notion of naturalness. As Lewis (1983: 347) notes, natural properties "make for resemblance". When is a property natural? I don't have the space to present a full theory of naturalness here. Still, there are two useful heuristics: a property is natural when it is involved in a well-confirmed epistemic domain, or, when it is defined in terms of properties that do so (Dunaway & McPherson 2016: 3.5).

Non-causal explanatoriness is a highly natural function in this sense. The set of non-causal explanations can be defined in terms of two other, very natural, entities (explanation simpliciter and causation).[14] If explanation simpliciter is highly natural and

---

[13] Examples might include interlevel mechanistic explanation (Craver 2009) or truthmaking explanation (Asay 2017).

[14] There are three caveats I need to mention here. First, when a property is defined in terms of other natural properties the relevant definition needs to be relatively short. How short? It's hard to say. Still, the

its subset, the set of causal explanations, is also natural, then what remains should also be natural. The principle I am appealing to is the following:

> **(P)** If a set P is natural, and its proper subset, Q, is natural, then not-Q is also natural.

(P), like every metaphysical principle, is controversial. Still, I have yet to encounter a convincing counterexample to it. Consider the following case. Let G be an imaginary kind of being identical to an animal that is not a tiger. Being an animal and being a tiger are both reasonably natural. But is G also natural? Being an animal is perhaps more natural than being G. But it doesn't follow that G is not natural simpliciter. In fact, examples of such properties can be traced in scientific practice. Take G to be the property of having an autoimmune disease that is not Celiac disease. This property follows the initial principle of being defined in terms of a natural property and its natural proper part. Insofar as scientific practice is a heuristic for naturalness, it would be controversial to deny that G is natural.

Whether such cases extrapolate in a way that completely vindicates the initial thesis is a matter for further discussion. What is clear, however, is that (P) is more plausible than its negation.[15] For this reason, I take the naturalness of non-causal

---

definition of f-grounding is surely short *enough* insofar as it involves only two entities. Secondly, saying that entities have definitions is shorthand for saying that their corresponding *predicates* have definitions. Thirdly, note that I'm appealing to a graded notion of naturalness as the entities in terms of which f-grounding is defined are not perfectly natural under Lewis's original account (for one, neither 'causation' nor 'explanation' figures in fundamental physics). I take this qualification to be harmless and not without precedent (Sider 2011: 130-33).

[15] Appealing to (P) indicates that non-causal explanatoriness is natural in virtue of the entities that figure in its definition. But even if one remains skeptical towards (P), we can take non-causal explanatoriness to be *directly* natural regardless of its constituents. As mentioned, a plausible heuristic for the naturalness of P is whether P figures in a well-confirmed epistemic domain. According to that heuristic, non-causal

explanatoriness to be secured. At the very least, its naturalness is *more* established than the naturalness of other candidates in the literature such as metaphysical and synchronic explanation.

How about other functionalist proposals? I will focus on what I take to be the most developed functionalist account in the literature. Rettler (2017) defines grounding in terms of a cluster of roles, one of which is the role of being able to "relate the fundamental with the less fundamental level" (2017: 6).[16] The first thing to note is that the notion of fundamentality itself is highly contested and the target of new and ongoing research. More importantly, according to many putative definitions (e.g. Leuenberger 2020) fundamentality is defined in terms of grounding. But grounding was what we were trying to figure out in the first place! Or consider another role that Rettler proposes: appealing to grounding to "investigate relations of metaphysical dependence" (2017: 6). Is this a sufficiently natural role? I doubt it. The literature on dependence makes it clear that dependence is a vast phenomenon and covers many distinct relations (Tahko 2018). In this sense, (Grounding Functionalist) is more developed and plausible than its rivals.

---

explanatoriness is natural since it is routinely used in the sciences. In fact, the use of non-causal explanations is ubiquitous in science in the same way causal explanations are (Reutlinger & Saatsi 2018).

[16] Although it should be noted that the aim of his paper is not to present a functionalist theory but a kind of grounding monism that is broader in scope (Rettler 2017: sec. 4). Other pluralist accounts have been advanced by Smithson (2018: 3) and Richardson (2018). Smithson's account is incomplete as he doesn't specify what the unifying feature of different instances of grounding is supposed to be. Richardson (2018) identifies metaphysical explanation as the unifying feature but doesn't define metaphysical explanation itself (a term which, as mentioned, I find vague and unhelpful). Wilson (2014: 567-8) also mentions a functionalist theory of this sort.

3.2. Spelling out the account

What about the notion of realization at play in the definition of (Functionalist Grounding)? Thankfully, we don't have to reinvent the wheel. A minimal characterization of functional realization is the following:[17]

> **(R)** A property P functionally realizes a property Q if and only if for some role functional R (i) Q is the property of having a property that occupies R, and (ii) P is a property that occupies R. (Baysan 2015: 6)

Modifying the proposal in relational terms is a matter of simple substitution. Also, note that nothing in (R) requires that the relevant role-occupants play R *causally*. This is a good thing as, typically, most grounding relata occur synchronically (although they don't *have* to under the current proposal).

What about the realizers themselves? One thing to notice is that we won't have a full account of the nature of the relevant realizers until we have a complete theory of non-causal explanation. For example, if one's theory of explanation takes a successful explanation to be one that cites an asymmetric determination relation, then realizers of f-grounding will be those determination relations that are asymmetric.[18] Still, defining

---

[17] Numerous more fine-grained definitions of functional realization have been proposed (e.g. Wilson 2011) but running my presentation along such lines would needlessly complicate things.

[18] It should be noted, however, that under this relaxed view about the formal features of grounding, f-grounding doesn't have interesting general formal features: it is non-symmetric, non-transitive, and non-reflexive (assuming that most of the cases in fn. 9 are genuine metaphysical possibilities). In contrast to Bennett (2017: 2.5.) I take this to be feature of my view (vis-à-vis Pluralism and the challenge from explanatory unificationism) not a bug. Rather, it is a problem for the thesis that grounding (or "building" according to her ideology) is *general* (2017: 22, 28). But f-grounding *does* have a general modal profile and its realizers are simply instances of it. Also note that this kind of generality isn't trivial since it entails interesting philosophical consequences (see section 2).

grounding without building a particular account of non-causal explanation into the definition is a feature, not a bug.

For one, we don't *need* to have a full theory of non-causal explanation to know that there *are* instances of successful non-causal explanations. And insofar as those explanations involve relations, then there are instances of f-grounding (I will talk about the possibility of explanations not being backed by relations later). A non-grounding example might help to illustrate. Some philosophers argue that phenomenal experiences form a natural kind even though we do not know everything there is to be known about them. We have reasons to posit the existence of qualia based on what qualia *do* (they help agents orient themselves, they have evolutionary purposes, etc.) (Lycan 1987). Knowing everything about P is not a necessary condition for knowing that P exists. Analogously, a functional characterization of grounding is insightful even if we do not have full knowledge of all the possible extensions of the term 'f-grounding'. Not only do we know that there *are* genuine non-causal explanations backed by relations, but we also have many plausible *candidates* that can play this role. Many of those relations, I claim, can be found in the set of what Wilson calls 'small-g' relations (see also fn 3).[19]

Secondly, staying non-committal towards the ultimate theory of explanation doesn't compromise the unity of grounding. (Functionalist Grounding) is compatible with the possibility of radical multiple realization. An entity is radically multiply realizable iff it can be realized by entities that have no first-order properties in common. In this sense, it could be that distinct instances of grounding share nothing substantive in common at the first-order level. Or it might turn out that every grounding relation

---

[19] I say "many" and not "all", since Wilson can be interpreted as claiming that some small-g relations are causal (i.e. causal composition) (2014: 540). Consider also Bennett's list of building relations which explicitly includes causation (2017: 4.2.) (cf. Schaffer 2019).

has some set of fixed formal features which are yet to be discovered.[20] Either way, second-order similarity (with the caveats mentioned in 3.1.) is enough for (Unity).

Finally, the debate surrounding the right formal features of explanation is nuanced and ongoing. In this sense, not committing to a particular theory of explanation is the only non-parochial and prudent move at this point. Also, one can be neutral towards the details of the final theory of explanation and still recognize that we have independently good reasons to think that many paradigmatic small-g relations are genuinely explanatory and, hence, genuine realizers of f-grounding. In this sense, it is important to highlight that the definition of f-grounding should not be necessarily understood, also, as an *epistemology* of f-grounding. There might we other reasons according to which a relation is judged to be a realizer of f-grounding (e.g. from pre-theoretical intuitions to abductive reasons concerning scientific practice).

Before I move to the next section, I want to quickly address two worries. First, someone might object that (Functionalist Grounding) is circular. A relation is a grounding relation iff it can *realize* the non-causal explanatoriness role. But functional realization is also a grounding relation (since it is plausible that functional realization is a realizer of f-grounding). So, f-grounding is defined in terms of itself.

I have two responses. First, there are ways to account for functional-talk without appealing to a realization *relation*. For example, we could say that grounding is identical to the disjunction of all of its possible realizers (Clapp 2001).[21] Still, regardless of the

---

[20] I mention this because a contentious issue in the relevant literature concerns the kinds of constraints that are at play on the realizers of some higher-order phenomenon. For example, it is one thing to say that a given mental state is multiply realizable by many distinct neural states, and quite another to claim that the set of realizers involves radically qualitatively distinct entities like ectoplasmic states, etc.

[21] If the set of f-grounding realizers is a mere disjunction, then in what sense are they importantly unified? Kim (1992: 13), in a similar context, argues that "disjunctive properties [...] do not guarantee similarity

nature of the realization relation, the circularity worry is confused. Even if f-grounding is defined in terms of realization, it is not defined in terms of realization *qua grounding relation*. Rather, once we discover that instances of realization can back explanations *then* we come to recognize that an ingredient of the definition of grounding is also an instance of grounding.

Analogously, causation can be defined (functionally) as the relation that backs certain explanations. But it might turn out that backing itself should be understood causally according to some causal theory of reference: a grounding-fact backs its corresponding explanation insofar as the former causally regulates the latter (e.g. as per Boyd [1999: 70]). In this sense, causation would be defined in terms of backing which, in turn, is understood as a causal relation. But, again, causation is not defined in terms of backing *qua causal relation*. The reason for this is that we have a pre-theoretical grasp of the backing relation and its function. Saying that causation backs certain explanations *is* informative even if it turns out that backing is an instance of causation.

Secondly, someone might worry about the open-endedness of the account. It might be objected that the account is *too* liberal. What about non-causal explanations that don't seem to involve relations that are typically understood as grounding relations? Consider a case of what Smithson has recently called "conceptual grounding": the fact that x is a vixen holds in virtue of the fact that x is a female fox (2018: 4). According to (Functionalist Grounding), insofar as there are genuine cases of such conceptual

---

for instances falling under them". I agree that it is not *necessary* for a disjunction to have objectively similar instances. But under the current account, similarity is preserved. The unity of f-grounding realizers is secured in virtue of them playing the non-causal explanatoriness role. The key behind the disjunctive characterization concerns, rather, the fact that the functional state at play is *identical* to whichever realizer successfully implements it (for this sort of "filler" functionalism see McLaughlin [2006]).

explanations then the underlying relation is a realizer of f-grounding. In a sense, this is a revisionary result. Relational views typically take grounding to be worldly whereas conceptual explanations are typically underwritten by relations between interpreted sentences. Similar remarks can be made concerning mathematical explanations (Lange 2016), equilibrium explanations (Sober 1983), and the like.

At this stage I am prepared to bite the bullet. (Functionalist Grounding) accommodates both constraints from section 2 while also having other virtues. This allows me to treat outliers such as conceptual and mathematical grounding as discoveries, not counterexamples. If this means that I need to countenance a relation of (say) mathematical non-causal determination relation, then so be it.[22]

4. Grounding and Explanatory Unificationism

(Functionalist Grounding) establishes a link between explanation and grounding. Kovacs (2017: sec. 5; 2020: 1) and Baron & Norton (2020: 3) appeal to explanatory unificationism to argue against this.[23] According to unificationism,

---

[22] I don't have the space to go through every type of non-causal explanation that appears in the literature (for an overview see Reutlinger & Saatsi [2018]). A worry at this point is that there might be non-causal explanations which do not seem to involve *any* underlying determination relations (e.g. geometrical explanations). I will talk about the possibility of explanations not involving relations in the final section. Still, it's worth mentioning that such cases can be interpreted as either involving a combination of underlying non-causal relations (for an example of how this could work see Kovacs [2017: 19]) or as being causal explanations in disguise (Skow 2014).

[23] Kovacs (2017) presents this challenge against the backdrop of the argument from explanatoriness for the existence of grounding. According to this argument, grounding exists because only by appealing to grounding we can make sense of non-causal explanations. It is widely held that this is the grounding theorist's most powerful argument. So, an argument that compromises the connection between grounding and explanation lands a very powerful blow to the existence of grounding. For simplicity, I

explanations are certain kinds of arguments. The relevant explanantia are found in the premises and the explanandum is found in the conclusion. According to Kitcher's influential account, explanations are arguments that figure in a maximally unified set. Roughly, for something to be a genuine explanation it must be the case that a candidate systematization, in virtue of including that explanation, generates the largest possible set of *conclusions* using the *smallest* number of *patterns* (which must be as *stringent* as possible).

At this point we need to present some parts of Kitcher's ideology.[24] A pattern is a set with three components: a *schematic argument* (a sequence of schematic sentences; i.e. expressions where, typically, some of the non-logical symbols are replaced with dummy letters), a set of *filling instructions* (directions for replacing dummy letters), and a *classification* (a descriptions of the inferential characteristics of the schematic argument) (Kitcher 1989: 4.2.). Finally, the more stringent an argument pattern is, the more difficult it is to be instantiated. For illustration purposes, consider a candidate systematization E and the following derivation (1989: 432):

(Test)

(1) For x_1, x_1 is homozygous for the sickling allele, iff, x_1 develops anemia.

(2) Organism x_1 is homozygous for the sickling allele.

---

will grant that this is the case. It is also worth noting that in his 2017 paper, the framing is a bit different (specifically, the relata of argument patterns are *themselves* metaphysical explanations). I will ignore this complication.

[24] Kitcher's account is very developed and complicated. In what follows I will present those pieces of ideology which are necessary for the relevant argument to be understood and I will relegate additional information in the footnotes. Still, it is worth noting that, at this point, what's important is the core idea behind unificationism, not its accompanying ideology. Kitcher specifically mentions that the logical notions he employs could be replaced with more sophisticated machinery (Kitcher 1989: 501 fn. 18).

(C) So, organism x_1 will develop anemia.

(Test), considered in isolation, doesn't significantly contribute towards the unifying power of E for it only concerns one entity (x_1). Still, (Test) is plausibly subsumed by the following pattern:

(Test*)

(1') For all x, x is homozygous for the sickling allele, iff, x develops anemia.

(2) Organism x_1 is homozygous for the sickling allele.

(C) So, organism x_1 will develop anemia.

(Test) instantiates the general argument pattern (Test*) which, in turn, *is* an important contribution towards the unifying power of E. (Test*) can generate many different conclusions (for all kinds of different organisms) and is highly stringent (not any kind of state of affairs can instantiate its components).

Analogously, for f-grounding to exist, an explanation involving f-grounding should appear in the maximally unifying systematization. The unificationist argues that we have good reasons to believe that it doesn't. Consider a derivation involving functional realization:

(Realization)

(1) For x_1, P, Q, x_1 having brain state type P *functionally realizes* x_1 having mental state type Q.

(2) x_1 has a token of P.

(C) So, x_1 has a token of Q.

The natural move for the f-grounding theorist is to say that (Realization) is an instance of a more general pattern:

(F-grounding)

(1') For all x, P, Q, x having P *f-grounds* x having Q.

(2) x_1 has a token of P.

(C) So, x_1 has a token of Q.

Does (Realization) instantiate (F-grounding)? It seems that it does. (1') is an instance of (1), assuming functional realization is a realizer of f-grounding (similar remarks apply to patterns involving other paradigmatic small-g relations).[25] So now the question is whether (F-grounding) figures in the maximally unifying systematization.[26]

---

[25] To be more precise, for Kitcher (1989: 432-3), a particular derivation D instantiates a general argument pattern D* iff (i) D and D* have the same number of *terms,* (ii) each sentence in D can be obtained from D* using the relevant filling instructions, (iii) D and D* have the same classification. (i) is clearly met. (ii) and (iii) are also met given that functional realization is an instance of f-grounding.

[26] I want to quickly dispel two technical worries about (F-grounding). First, someone might object that (F-grounding) is not a genuine pattern because (1') is not a law of nature (a requirement that originates back to Hempel's deductive-nomological theory of explanation). Given that we are asked to apply the unificationist framework in the metaphysical domain, some amendments would have to be allowed. In this case, this involves replacing laws of nature with laws of metaphysics (or something of the sort) (e.g. Kment 2014: ch. 6). It is worth noting that Kitcher (1989: 447) takes this requirement loosely since he allows for non-exceptionless laws (or, what he calls, '*mini laws*'). Secondly, (F-grounding) is logically invalid unless we modify the relevant inference rule involved in the classification. Specifically, it would have to be admitted that every case of f-grounding is itself an instance of an entailment involving the same relata (i.e. if P f-grounds Q, then P entails Q). But then (F-grounding) is committed to some form of grounding necessitarianism. Is this a problem? One the face of it, it is. It would be nice to be able to countenance cases of contingent or indeterministic f-grounding in accordance with (Pluralism). Thankfully there is a way to do this under the unificationist framework if one allows for non-deductive arguments to figure in the best systematization. In fact, many happily make this amendment (e.g. Railton 1978; Schweder 2005: 433) as it doesn't compromise the key motivation behind explanatory

This is where the unificationist challenge comes into play. The unificationist argues that, regardless of its merits, (F-grounding) should be replaced with a simpler pattern:

(Simple)

(1) For all x, P, Q, x has P, iff, x has Q.

(2) x_1 has a token of P.

(C) So, x_1 has a token of Q.

(Simple) dispenses with the f-grounding operator and replaces it with a simple biconditional.[27] Simply put, a systematization with (Simple) will be more unified than a systematization with both (Simple) and (F-grounding). Other things being equal, the systematization with the lowest number of patterns is the more unifying one. If this is the case, then f-grounding is not compatible with explanatory unificationism. Non-causal explanations would not be backed by grounding-like facts. Rather, if P explains Q, then P *merely entails* Q (where the relevant entailment-claim is not backed by a grounding relation).

There are many ways in which the f-grounding theorist can respond to this challenge. Perhaps a promising strategy would be to deny explanatory unificationism altogether. After all, the unificationist must assume the truth of unificationism in order to advance her argument. But is it true that unificationism is independently motivated as a theory of explanation? Many philosophers disagree. For example, many argue that determination-based theories have emerged victorious (de Regt 2006; Woodward

---

unificationism (namely, the incorporation of complex phenomena under less complex, and more basic, phenomena).

[27] Formalizing scientific laws and similar principles using biconditionals is the default view in the literature (e.g. Dretske 1977; Friend 2016).

2003). Similarly, others argue that unificationism and determination-based views (whether causation or grounding-based) are compatible: unificationist explanations operate at a higher level of generality than grounding-based ones but there is no reason to believe that these explanations are incompatible with one another.[28] It is uncontroversial that higher and lower-level explanations can (at least sometimes) complement each other. Why assume that unificationism excludes other forms of explanation? Perhaps, unificationism shouldn't be understood in global terms but as concerning a particular subset of explanations. Still, I think the f-grounding theorist can do better. My strategy will be to argue that, even by the unificationist's own lights, at least some argument must appeal to f-grounding relations, thus showing that (Functionalist Grounding) is compatible with explanatory unification.

4.1. The Supplementation Strategy

My first attempt at introducing f-grounding tried to *add* an argument pattern in the relevant systematization.[29] The main challenge concerned the size of the systematization. A systematization with (Simple) has less argument patterns than a systematization with (Simple) *and* (F-grounding). Still, there are reasons to be optimistic about this strategy.

For one, (F-grounding) is highly stringent. It is more difficult for a derivation to instantiate (F-grounding) than (Simple). This is obvious considering (F-grounding) has one extra term than (Simple): the f-grounding operator. To illustrate, consider a case where P and Q modally covary without there being an intimate connection between them. A derivation involving this metaphysical scenario would instantiate (Simple)

---

[28] See Railton (1978: 208); Kim (1994: 69); De Regt (2006).

[29] Kitcher calls the set that includes every argument pattern, the 'generating' set (Kitcher 1981: 519-20; 1989: 434).

without instantiating (F-grounding). On the contrary, a derivation involving the relation of (say) constitution, would instantiate both argument patterns.

Secondly, consider a desideratum for unification that is usually omitted from the relevant discussion: the set of beliefs that a given systematization involves is supposed to include beliefs that are in some *default* state. This is what Kitcher calls the *explanatory store*. This store includes a 'reserve' of beliefs and arguments "which we may tap as need arises" (1981: 512, 519). Typically, this explanatory store includes scientific beliefs but for my purposes I will assume that it also involves beliefs about metaphysics. For example, there is a lengthy and developing literature on the nature of constitution. Constitution theorists disagree about its features and try to deal will classic puzzles involving, say, coincident entities. In their domain, the existence of constitution is assumed. In this sense, a systematization that accommodates many (say) constitution-beliefs is superior to its revisionist counterpart, other things being equal.[30]

Finally, a systematization than includes (F-grounding) has a straightforward way of dealing with a notorious problem that unificationist theories typically face: the problem of asymmetry. The typical example involves the explanation of the length of a flagpole's shadow in terms of the flagpole's height, the fact that there is a sun, and certain ancillary conditions involving the relevant laws of nature. Because laws typically have a biconditional form, it seems that one can derive the relevant law of nature from the flagpole's shadow and vice versa (Bromberger 1966). But this doesn't seem right. The correct result involves deriving the flagpole's shadow from the relevant law of nature and other ancillary assumptions, not the other way around.

---

[30] This is one of the reasons that (Functionalist Grounding) is particularly fit to deal with the challenge from unificationism. Perhaps, other pluralist accounts could deliver the same result. Still, as it stands, I take (Functionalist Grounding) to be the only plausible account on the market that can simultaneously meet (Unity) and (Pluralism).

The problem generalizes to non-causal explanations. Consider the explanation of a mental state in virtue of its corresponding brain state and the relevant psychophysical law. Similarly, due to the biconditional nature of the relevant law, it seems that we can derive (and, thus, *explain*) one's brain state partly in virtue of one's corresponding mental state.

Unificationists are aware of the problem. Typically, they argue that the intuitively correct derivation is also the one that importantly contributes to the unification of the relevant systematization.[31] Still, the f-grounding theorist can simply appeal to an asymmetric realizer of f-grounding to break the symmetry. In this sense, the relevant psychophysical law would be enriched via an f-grounding operator which would fix the order of determination. In the current example, the law would specify that, necessarily, mental states are functionally realized by brain states (but not vice versa).[32]

---

[31] For example, see Kitcher (1981: sec. 7; 1989: 436, 7.3). For an application of this strategy to the metaphysical domain see Baron & Norton (2020: sec. 4).

[32] It could be objected that even if we grant that mental states (M) are (say) functionally realized by brain states (B), given the relaxed view I adopted concerning the formal features of grounding (section 2), it might turn out that there's another realizer of f-grounding which runs towards the opposite direction (that is, from M towards B). This is definitely possible. But I don't take this to be a serious problem. If M somehow non-causally determines B (which is doubtful in itself), that would simply indicate that there's another non-causal explanation to be backed alongside the initial one (the one running from B to M). That would be a non-standard scenario, and it is a feature of the f-grounding framework that it can accommodate such cases (see fn. 9). Perhaps the worry concerns the possibility of explanatory circularity. If M f-grounds B, and B f-grounds M, then M grounds itself. Again, I don't think this is a serious worry. It is independently plausible that explanations can 'chain' in this way only if certain constraints are met (e.g. if the relevant contrast classes align properly) (for a different proposal see Hicks & van Elswyk [2015]).

Compared to the typical unificationist response, this is a preferable strategy for at least two reasons. The first reason is epistemological. For the unificationist the putative unifying power of argument patterns is judged holistically. This means that we won't be able to tell whether a given argument pattern is genuinely unifying *unless* we also have a sufficiently large portion of the relevant optimal systematization. To compare, a response that can appeal to an f-grounding relation is straightforward. In the example at hand, for example, we have independently plausible *empirical* reasons to think the determination runs from the neural level towards the mental level.[33]

Secondly, (F-grounding) provides fine-grained information about the underlying connection between the relevant entities. This point becomes clearer once we understand f-grounding relations as the non-causal analogues of causal mechanisms. Causal mechanisms make ordinary causal relations more precise by providing fine-grained details about the way the causal connection runs. For example, instead of saying "the fact that I didn't water the plants caused them to wither" we can supply a detailed story taking as a set-up position an event that doesn't include my intention to water the plants and the event of their withering as a termination condition (Machamer & Darden & Craver 2000). Trogdon (2018) has recently argued that grounding relations have a similar function to mechanisms: they can specify the relevant determination relation (although, in the grounding case, the relata are not causes and effects). For example, instead of saying that a given brain state grounds its corresponding mental state, we can say that the former *functionally realizes* the latter. The latter story is explanatorily richer than the former (Wilson 2014: sec. II).

---

[33] I have in mind certain counterfactual tests that are routinely used in the special sciences (Woodward 2003). For a similar solution to the asymmetry problem (that appeals to a notion of nomic determination) see Wilsch (2016: 4, 15-16). For a classic response that abandons unificationism see Salmon (1989). For a response in the context of unificationism see Bangu (2016: 11, 16).

Still, I think we can do better. Even with all these benefits, the supplementation strategy still countenances an extra argument pattern in addition to (Simple). So, on the one hand we have a systematization (call it S) that involves (Simple), and on the other we have a systematization (call it S*) that involves (F-grounding) in addition to (Simple). S wins in terms of having fewer argument patterns whereas S* is superior in terms of stringency, dealing with the asymmetry problem, and being less revisionist towards the explanatory store. Trying to compare S and S* in terms of their overall unifying power is a very difficult task. Kitcher (1989: 433, 435, 481) admits that it is not clear how the relevant desiderata are supposed to compete with one another. Even though I think that the supplementation strategy is promising, it doesn't definitively counter the unificationist challenge.

4.2. The Substitution Strategy

I propose moving forward in the following way: (Simple) should be *substituted* by (F-grounding). This way, a systematization with (F-grounding) will be definitively more unifying than a systematization with just (Simple), other things being equal.

For this move to work, (Simple) needs to be an instance of (F-grounding). One might be skeptical of this. (Simple), as motivated by the anti-relational unificationist, is not backed by a non-causal determination relation.[34] But, if (Simple) is replaced by (F-grounding), then the best systematization won't include explanations which do not involve any underlying relations. For example, a version of non-naturalism might say that moral properties supervene on non-moral properties without the latter grounding

---

[34] Kovacs (pers. comm.) is not committed to this (as he doesn't fully accept the argument from explanatoriness; also see fn 23). Rather, he holds that even if such arguments are backed by determination relations it is not in virtue of those relations that they exhibit their unifying power.

the former. Or, perhaps, one can appeal to certain mathematical facts supervening on other mathematical facts.

Notice, however, that not every supervenience thesis is created equal. Some supervenience theses (like moral supervenience) have explanatory ambitions, whereas others are trivial (like the fact that 2+2=4 supervening on the fact that 3+3=6). I take it that the objector wishes to accommodate the former kind of supervenience, not the latter. But it is plausible that explanatory supervenience theses *are* underwritten by grounding-like relations (i.e. candidate realizers of f-grounding). It is no accident that, in the relevant literature, the *explanation* of moral supervenience is a datum that every metaethical theory should be able to meet (Väyrynen 2017). In this sense, even non-naturalists must appeal to a *relation* that underwrites moral supervenience (even though such a relation is often cashed out in primitive terms) (Rosen 2017).[35]

Time to take stock. On the one hand we have S which includes (Simple) and on the other S* which includes (F-grounding) *having replaced* (Simple). Even if we assume that S and S* have the same amount of argument patterns and generate the same number of conclusions, it is still the case that moving from S to S* constitutes the pareto optimal move. So, it is plausible that (F-grounding) appears in the best

---

[35] What about so-called *conceptual* explanations of moral supervenience? Presumably, moral properties supervene on natural properties as a matter of conceptual necessity. I am unconvinced by such solutions as they fail to accommodate the right modal status of moral supervenience (specifically, the modality indicated by inner metaphysical necessity operator) (Dreier 2019). At any rate, even if such explanations were to go through, I am still confident that they can understood as involving non-causal relations (recall my remarks on conceptual grounding in section 3).

systematization of the phenomena, which means that the existence of f-grounding is compatible with explanatory unificationism.[36]

5. Conclusion

I conclude that f-grounding is a theory that deserves serious consideration. First, I argued that it meets two important constraints that pull in opposite directions: (Monism) and (Pluralism). F-grounding is functionally realized by determination relations that fulfill the non-causal explanatoriness role. Exemplars of such realizers are usually dubbed as 'small-g relations' in the relevant literature and include relations such as composition, constitution, and functional realization.

Secondly, I argued that (Functionalist Grounding), in virtue of meeting (Pluralism) and (Monism), can meet a powerful objection that threatens every theory of grounding: the objection from explanatory unificationism. According to this worry, the realizer set of f-grounding is empty since every successful non-causal explanation isn't backed by a determination relation. However, I argued that, even by the unificationist's

---

[36] So far, I have been assuming that f-grounding's existence depends on whether it appears in the best systematization of the phenomena. Is this assumption warranted? Looking closely into the way Kitcher treats the relation of causation it seems to follow that even if a relation figures in a systematization, it doesn't mean that that relation exists. For Kitcher, causal relations between phenomena simply reflect our inferential practices (1989: 4.5). Fortunately, nothing in the theory of explanatory unificationism *itself* entails antirealism about explanation. Rather, if explanatory unificationism has an antirealist component it is in virtue of what Kitcher calls 'the top-down approach' to explanation (1989: 439). A top-down approach to explanation, roughly, proceeds by defining explanation first and then trying to cash out talk about worldly items (such as causation, grounding, and mechanisms) in terms of explanation (rather than the converse). For a recent attempt to apply this idea to the case of grounding see Kovacs (fc.).

own lights, an argument pattern involving f-grounding is more unifying than its anti-relational counterpart.

References

Asay, Jamin (2017). Run Aground: Kit Fine's Critique of Truthmaker Theory. *Philosophical Quarterly* 67 (268):443-463. https://doi.org/10.1093/pq/pqw072

Audi, Paul (2012). Grounding: Toward a Theory of the In-Virtue-Of Relation. *Journal of Philosophy* 109 (12):685-711. https://doi.org/10.5840/jphil20121091232

Baron, Sam & Norton, James (2020). Metaphysical Explanation: The Kitcher Picture. *Erkenntnis*:1-21. https://doi.org/10.1007/s10670-018-00101-2

Baron, Sam ; Miller, Kristie & Tallant, Jonathan (2020). Grounding at a distance. *Philosophical Studies* 177 (11):3373-3390.

Baysan, Umut (2015). Realization Relations in Metaphysics. *Minds and Machines* (3):1-14. https://doi.org/10.1093/analys/anx086

Bennett, Karen (2017). *Making Things Up*. Oxford University Press.

Bernstein, Sara (2020). The metaphysics of intersectionality. *Philosophical Studies* 177 (2):321-335. https://doi.org/10.1007/s11098-019-01394-x

Bliss, Ricki & Priest, Graham (eds.) (2018). *Reality and its Structure: Essays in Fundamentality*. Oxford University Press.

Boyd, Richard (1999). Kinds, Complexity and Multiple realization. *Philosophical Studies* 95 (1):67-98. https://doi.org/10.1023/A:1004511407133

Bromberger, Sylvain (1966). Questions. *Journal of Philosophy* 63 (20):597-606. https://doi.org/10.2307/2024253

Clapp, Lenny (2001). Disjunctive Properties: Multiple Realizations. *Journal of Philosophy* 98 (3):111-136. https://doi.org/jphil20019834

Craver, Carl F. (2009). Mechanisms and natural kinds. *Philosophical Psychology* 22 (5):575-594. https://doi.org/10.1080/09515080903238930

Dasgupta, Shamik (2017). Constitutive Explanation. *Philosophical Issues* 27 (1):74-97. https://doi.org/10.1111/phis.12102

de Regt, Henk W. (2006). Wesley salmon's complementarity thesis: Causalism and unificationism reconciled? *International Studies in the Philosophy of Science* 20 (2):129 – 147. https://doi.org/10.1080/02698590600814308

Dowe, Phil (1996). Backwards causation and the direction of causal processes. *Mind* 105 (418):227-248. https://doi.org/10.1093/mind/105.418.227

Dreier, Jamie (2019). Is there a supervenience problem for robust moral realism? *Philosophical Studies* 176 (6):1391-1408.

Dretske, Fred (1977). Laws of nature. *Philosophy of Science* 44 (2):248-268.

Dunaway, Billy & McPherson, Tristram (2016). Reference Magnetism as a Solution to the Moral Twin Earth Problem. *Ergo: An Open Access Journal of Philosophy* 3.

Emery, Nina (2019). Laws and their instances. *Philosophical Studies* 176 (6):1535-1561. https://doi.org/10.1007/s11098-018-1077-8

Epstein, Brian (2015). *The Ant Trap: Rebuilding the Foundations of the Social Sciences.* Oxford University Press.

Fine, Kit (2012). Guide to Ground. In Fabrice Correia & Benjamin Schnieder (eds.), *Metaphysical Grounding*. Cambridge University Press. pp. 37--80.

Friedman, Michael (1974). Explanation and scientific understanding. *Journal of Philosophy* 71 (1):5-19.

Friend, Toby (2016). Laws are conditionals. *European Journal for Philosophy of Science* 6 (1):123-144.

Gillett, Carl (2016). *Reduction and Emergence in Science and Philosophy.* Cambridge University Press.

Hicks, Michael Townsen & van Elswyk, Peter (2015). Humean laws and circular explanation. *Philosophical Studies* 172 (2):433-443.

Kim, Jaegwon (1992). Multiple realization and the metaphysics of reduction. *Philosophy and Phenomenological Research* 52 (1):1-26.

Kim, Jaegwon (1994). Explanatory knowledge and metaphysical dependence. *Philosophical Issues* 5:51-69. https://doi.org/10.2307/1522873

Kitcher, Philip (1981). Explanatory unification. *Philosophy of Science* 48 (4):507-531. https://doi.org/10.1086/289019

Kitcher, Philip (1989). Explanatory unification and the causal structure of the world. In Philip Kitcher & Wesley Salmon (eds.), *Scientific Explanation.* Minneapolis: University of Minnesota Press. pp. 410-505.

Kment, Boris (2014). *Modality and Explanatory Reasoning.* Oxford University Press.

Kovacs, David Mark (2017). Grounding and the argument from explanatoriness. *Philosophical Studies* 174 (12):2927-2952. https://doi.org/10.1007/s11098-016-0818-9

Kovacs, David Mark (2019). Four Questions of Iterated Grounding. *Philosophy and Phenomenological Research.* https://doi.org/10.1111/phpr.12591

Kovacs, David Mark (2020). Metaphysically explanatory unification. *Philosophical Studies* 177 (6):1659-1683. https://doi.org/10.1007/s11098-019-01279-z

Kovacs, David Mark (forthcoming). An explanatory idealist theory of grounding. *Noûs.*

Lange, Marc (2016). *Because Without Cause: Non-Causal Explanations in Science and Mathematics.* Oxford University Press USA.

Leuenberger, Stephan (2020). The fundamental: Ungrounded or all-grounding? *Philosophical Studies* 177 (9):2647-2669. https://doi.org/10.1007/s11098-019-01332-x

Lewis, David (1983). New work for a theory of universals. *Australasian Journal of Philosophy* 61 (4):343-377. https://doi.org/10.1080/00048408312341131

Litland, Jon Erling (2018). In Defense of the Disunity of Grounding. *Thought: A Journal of Philosophy* 7 (2):97-108. https://doi.org/10.1002/tht3.273

Lycan, William G. (1987). *Consciousness.* MIT Press.

Machamer, Peter, Darden, Lindley & Craver, Carl F. (2000). Thinking about mechanisms. *Philosophy of Science* 67 (1):1-25.

McLaughlin, B., 2006. "Is Role-Functionalism Committed to Epiphenomenaliam?", *Consciousness Studies,* 13 (1–2): 39–66.

Melnyk, Andrew (2008). Conceptual and linguistic analysis: A two-step program. *Noûs* 42 (2):267–291. https://doi.org/10.1111/j.1468-0068.2008.00680.x

Menzies, Peter (1996). Probabilistic causation and the pre-emption problem. *Mind* 105 (417):85-117. https://doi.org/10.1093/mind/105.417.85

Millikan, Ruth Garrett (1989). In defense of proper functions. *Philosophy of Science* 56 (6):288-302. https://doi.org/10.1086/289488

Nolan, Daniel (2018). Cosmic Loops. In Ricki Bliss & Graham Priest (eds.), *Reality and Its Structure*. Oxford, UK: Oxford University Press. pp. 91-106.

Pereboom, Derk (2002). Persons and Bodies: A Constitution View. *Philosophy and Phenomenological Research* 64 (3):615-622. https://doi.org/10.1111/j.1933-1592.2002.tb00165.x

Polger, Thomas W. & Shapiro, Lawrence A. (2016). *The Multiple Realization Book.* Oxford University Press UK.

Psillos, Stathis (2008). Causal Pluralism. In Robrecht Vanderbeeken & Bart D'Hooghe (eds) *Worldviews, Science and Us: Studies of Analytical Metaphysics: A Selection of Topics From a Methodological Perspective*, World Scientific Publishers, pp. 131-151, 2009.

Railton, Peter (1978). A deductive-nomological model of probabilistic explanation. *Philosophy of Science* 45 (2):206-226. https://doi.org/10.1086/288797

Rettler, Bradley (2017). Grounds and 'Grounds'. *Canadian Journal of Philosophy* 47 (5):631-655. https://doi.org/10.1080/00455091.2017.1282306

Reutlinger, Alexander & Saatsi, Juha (2017). Introduction: Scientific Explanation Beyond Causation. In Alexander Reutlinger & Juha Saatsi (eds.), *Explanation Beyond Causation: Philosophical Perspectives on Non-Causal Explanations*. Oxford: Oxford University Press.

Richardson, Kevin (2018). Grounding Pluralism: Why and How. *Erkenntnis*:1-17.

Rodriguez-Pereyra, Gonzalo (2015). Grounding is not a strict order. *Journal of the American Philosophical Association* 1 (3):517-534. https://doi.org/https://doi.org/10.1017/apa.2014.22

Rosen, Gideon (2010). Metaphysical Dependence: Grounding and Reduction. In Bob Hale & Aviv Hoffmann (eds.), *Modality: Metaphysics, Logic, and Epistemology.* Oxford University Press. pp. 109-36.

Rosen, Gideon (2017). Metaphysical Relations in Metaethics. In Tristram McPherson & David Plunkett (eds.), *The Routledge Handbook of Metaethics.* Routledge. pp. 151-169.

Saenz, Noël B. (2015). A grounding solution to the grounding problem. *Philosophical Studies* 172 (8):2193-2214. https://doi.org/10.1007/s11098-014-0405-x

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

Schaffer, Jonathan (2010). Monism: The Priority of the Whole. *Philosophical Review* 119 (1):31-76. https://doi.org/10.1215/00318108-2009-025

Schaffer, Jonathan (2012). Grounding, transitivity, and contrastivity. In Fabrice Correia & Benjamin Schnieder (eds.), *Metaphysical Grounding: Understanding the Structure of Reality.* Cambridge University Press. pp. 122-138.

Schaffer, Jonathan (2020). Taking causing out of Bennett's Making Things Up. *Inquiry: An Interdisciplinary Journal of Philosophy* 63 (7):722-744.

Schweder, Rebecca (2005). A defense of a unificationist theory of explanation. *Foundations of Science* 10 (4):421-435. https://doi.org/10.1007/s10699-004-5250-5

Shoemaker, Sydney (2007). *Physical Realization.* Oxford University Press UK.

Sider, Theodore (2011). *Writing the Book of the World.* Oxford University Press.

Skiles, Alexander (2015). Against Grounding Necessitarianism. *Erkenntnis* 80 (4):717-751. https://doi.org/10.1007/s10670-014-9669-y

Skow, Bradford (2014). Are There Non-Causal Explanations (of Particular Events)? *British Journal for the Philosophy of Science* 65 (3):445-467.

Smithson, Robert (2018). Metaphysical and Conceptual Grounding. *Erkenntnis* 1:1-25. https://doi.org/10.1007/s10670-018-0088-3

Sober, Elliott (1983). Equilibrium explanation. *Philosophical Studies* 43 (2):201 - 210.

Sveinsdóttir, Ásta (2015). Social Construction. *Philosophy Compass* 10 (12):884-892. https://doi.org/10.1111/phc3.12265

Tahko, Tuomas E. (2018). "Fundamentality", *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2018/entries/fundamentality/>.

Taylor, Elanor (2018). Against explanatory realism. *Philosophical Studies* 175 (1):197-219. https://doi.org/10.1007/s11098-017-0862-0

Trogdon, Kelly (2018). Grounding-mechanical explanation. *Philosophical Studies* 175 (6):1289-1309. https://doi.org/10.1007/s11098-017-0911-8

Väyrynen, Pekka (2017). The Supervenience Challenge to Non-Naturalism. In Tristram McPherson & David Plunkett (eds.), *The Routledge Handbook of Metaethics.* Routledge. pp. 170-84.

Wallner, Michael (2018). The ground of ground, essence, and explanation. *Synthese.* https://doi.org/10.1007/s11229-018-1856-y

Wilsch, Tobias (2016). The Deductive-Nomological Account of Metaphysical Explanation. *Australasian Journal of Philosophy* 94 (1):1-23.

Wilsch, Tobias (2020). The governance of laws of nature: guidance and production. *Philosophical Studies*:1-25. https://doi.org/10.1007/s11098-020-01465-4

Wilson, Jessica M. (2011). Non-reductive realization and the powers-based subset strategy. *The Monist (Issue on Powers)* 94 (1):121-154. https://doi.org/10.5840/monist20119417

Wilson, Jessica M. (2014). No Work for a Theory of Grounding. *Inquiry: An Interdisciplinary Journal of Philosophy* 57 (5-6):535-579. https://doi.org/10.1080/0020174X.2014.907542

Sjölin Wirling, Ylwa (2020). Is backing grounding? *Ratio* 33 (3):129-137.

Woods, Jack (2018). Emptying a Paradox of Ground. *Journal of Philosophical Logic* 47 (4):631-648. https://doi.org/10.1007/s10992-017-9442-1

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

Chapter 2 | Backing as Truthmaking

Abstract

Separatists about grounding take explanations to be separate from their corresponding grounding-facts. Grounding-facts are supposed to underlie, or back, such explanations. However, the backing relation hasn't received much attention in the literature. The aim of this paper is to provide an informative definition of backing. First, I examine two prominent proposals: backing as explaining (Kovacs's 2017; 2019a) and backing as grounding (see Sjölin Wirling 2020). Finally, I put forward my own proposal. I argue that under plausible assumptions about the role of backing and the nature of explanation, backing should be understood as a form of truthmaking, minimally construed.

1. Introduction

Grounding is a worldly non-causal determination relation that is widely thought to be linked with a particular kind of non-causal explanation.[1] Typical examples of grounding-facts include mental states being grounded in their physical features, chairs by their constituents, and sets by their instances. Unionists about grounding take grounding-facts to be identical to their corresponding explanations.

---

[1] These explanations are usually called *metaphysical explanations* (Fine 2012). Spelling out the nature of metaphysical explanations is hard and a matter of ongoing discussion (see e.g. A. Wilson 2020). In the context of this paper, I will treat metaphysical explanations as species of non-causal explanations (perhaps, they are non-causal explanations that are of some particular interest to metaphysicians). I will also take 'explanation' to refer to *non-causal* explanation unless stated otherwise.

Separatists, on the other hand, argue that explanations are distinct from grounding-facts: grounding-facts underlie, or *back*, explanations (Raven 2015).[2] Explanations of the form 'A explains B' are backed by facts of the form [A grounds B].[3] Still, the backing relation has not received much attention in the literature. The aim of this paper is to provide an informative definition of backing.

Why is this important? First, the backing-locution appears in many contexts other than the grounding literature. For example, causal or constitutive explanations are said to be backed in an analogous manner but, again, backing-talk is left unspecified.[4] Providing a unified account of backing would allow us to make sense of backing-talk across many different contexts. Secondly, separatists are at a disadvantage in comparison to unionists when they cannot specify what backing-talk is supposed to designate. If we expect to do any sort of progress vis-à-vis the unionism/separatism debate, we should know more about what backing is supposed to be.[5] Finally, a definition of backing might be important for functionalist theories of grounding. Such theories do not define

---

[2] Proponents of separatism include Schaffer (2016) and Wilson (2016). On the unionist camp we could include Rosen (2010), Fine (2012), and Litland (2018a) (still, see fn. 5).

[3] For readability, I assume that grounding and explanation-claims involve one-to-one relations. This is unusual. Most grounding theorists take the grounding relation to be many-to-one (Trogdon & Bliss 2016) (cf. Litland 2016).

[4] See e.g. Kim (1988: 226; 1994: 58, 60, 67), Ruben (1990: 232), Glennan (2002: 342).

[5] Two clarifications. First, I do not wish to imply that an informative definition of backing would *settle* the debate. It could be that there are independent reasons for why someone should be a separatist (Maurin 2019: 20). Still, the question about the nature of backing plays a significant dialectical role in the relevant debate (see, e.g., Wallner 2018: 3). Secondly, I am also assuming that the separatism/unionism debate is a substantial one. For the view that the debate is largely verbal see Dasgupta (2017: 94 fn. 8), Skiles & Trogdon (2019: 160), and Kovacs (2019a: 4). Note that if it turns out that unionists also need to countenance a backing relation, then they would also need to provide an account of backing.

grounding in terms of their first-order formal features but in terms of their functional role. The most obvious role that grounding relations can play is their ability to back non-causal explanations (Rettler 2017). A primitive or unspecified notion of backing would count significantly against such views.

In section 2 I propose and motivate some constraints about the nature of backing. In section 3 and 4, I examine two prominent proposals: backing as explaining (B=E) and backing as grounding (B=G). In short, I argue that both (B=E) and (B=G) generate more questions than they answer. In section 5, I put forward my own proposal. Briefly, I argue that under plausible assumptions about the role of backing and the nature of explanation, backing should be understood as truthmaking, minimally construed. Finally, I tackle some objections.

2. Platitudes about backing

Separatism is the view that explanations are separate from the facts that make them obtain. In this sense, separatism is a minimal thesis (although some of its versions might be more plausible than others).[6] Still, since the focus of this paper is the backing relation that obtains between these two separated entities, I will bring up plausible features of separatism that are relevant to my discussion when appropriate.

In this section I will propose a functional definition of backing based on minimal assumptions about its nature. Then, I will clarify and motivate these features,

---

[6] Sjölin Wirling (2020: 2-3) takes separatism to be primarily driven by the idea that explanation necessarily involves some epistemic or pragmatic dimension. But unionists also recognize that feature and actively try to accommodate it, either by adopting some non-monotonicity clause or by countenancing multiple (often primitive) grounding relations (e.g. Litland 2018b).

thus showing that they can act as prior constraints for any proposed theory of backing. The functional characterization I want to propose is the following:

> **(Functional)** [A grounds B] (G) backs 'A explains B' (E), iff, there is a relation R with the following features:
>
> (1) R holds non-trivially between G and E,
>
> (2) R is cross-categorical, and,
>
> (3) R's obtaining (at least partially) makes it the case that E's success-conditions are met.

A couple of specifications are in order. First, the non-triviality clause makes sure that relations such as 'being in the same world as' do not come out as backing-relations. Secondly, the success-conditions of an explanation are those conditions that, once they are met, necessitate that the explanation is successful. When a grounding-fact backs an explanation, then that fact contributes to the satisfaction of the success-conditions of that explanation.[7] Thirdly, this characterization involves grounding-facts but there is no reason to think that it cannot be extended to other kinds of facts as well.[8] I begin my

---

[7] Of course, I do not need to commit to the view that a given grounding-fact is *sufficient* for the relevant success-conditions to be fully met. Perhaps, there are facts other than grounding-facts (e.g. pragmatic/epistemic facts as per Sjölin Wirling [2020: 2-3, 5]) which are also necessary. Or, it could be that grounding-facts are sufficient on their own given that the relevant pragmatic/epistemic conditions are *built into* the relevant grounding-fact (e.g. by taking grounding to be a quaternary relation ranging over contrast classes).

[8] The standard view is that explanations are backed by determination-facts (broadly construed) (see Taylor [2018]). But there might be other ways to back an explanation. For example, Lange (2016) claims that some explanations are distinctively modal (in cases were the explanandum holds with a weaker modality than the explanans). Glazier (2020) also mentions explanations backed by *nomic* facts. Kovacs (2019a: section 4.3; 2019b) argues that non-causal explanations are backed holistically. deRosset (2013:

discussion by focusing on grounding since backing has received attention primarily in the context of grounding (Raven 2015; Schaffer 2016). Even though, as mentioned, the backing-locution can also be found in other domains, it is fair to expect that the main source of my audience are people interested in grounding.[9] Finally, I take backing to be cross-categorical in the sense that backers (in this case, grounding-facts) are different in *kind* from the explanations they back. This is perhaps the most contentious proposed feature of backing and for this reason it requires further discussion.

Why think that backing is cross-categorical? First, I take cross-categoricity to be a *default* feature of backing. Taking at face value the way the term 'backing' is used in the literature, entails that backing is cross-categorical. Grounding-facts are taken to be worldly and broadly stance-independent, whereas explanations are *about* such facts (thus, representational in character). This is particularly salient if one considers the relation between causal events and the explanations which they back. As Schaffer (2016: 36) puts it:

> [O]ne wants to distinguish between causation—a concrete relation in the world—and causal explanation—an abstract pattern over facts or sentences. And one wants to connect these notions by allowing causal relations in the world to back causal explanations among facts or sentences. Or so orthodoxy has it, and so I take for granted here.

Secondly, it is independently plausible that explanations are ontologically different from their backers. Consider causation again. Causal explanations are apt to

---

12) takes explanations to be backed by arguments. Taylor (2020) has also argued that there is room for an anti-realist version of backing.

[9] I will transition to a more generalized conception of separatism when dialectically appropriate (e.g. sections 4, 5, 7; see also fn. 8).

pragmatic/epistemic constraints whereas the causal relation itself typically is not.[10] To compare, separatists about grounding make analogous remarks (Sjölin Wirling 2020: 2-3). In this sense, it is no accident that separatists about grounding model their view according to the relation between causation and causal explanation.

There are ways to resist the cross-categoricity of backing. Sjölin Wirling (2020: 4, 5, 6), in an interesting move, argues that *explanations* are not the proper kind of relatum for the backing relation. Rather, she argues, a grounding-fact like [A grounds B] backs the *fact* that 'A explains B' is an explanation. In other words, backers back the explanatoriness of that which represents them. In this sense, backing is not cross-categorical since both of its relata are facts.

It could be argued that one *can* translate explanation-talk into fact-talk: saying 'a grounding-fact G backs an explanation E' can be translated into 'G backs the fact that E is an explanation'. But it is unclear if one *should* do so. First, translating talk of one ontological category to another does not come for free. To compare, consider operator-based views of grounding being translated into relational-talk. This is certainly possible, but one would need to make some substantive ontological assumptions along the way (i.e. countenancing a *relation* of grounding). In this sense, it is unclear if the results of such translation procedures can be trusted. More importantly, it seems to me that the explanation *'A explains B'* and the fact that *'A explains B' is an explanation,* are importantly different entities. The former is an explanation, whereas the latter is a fact *about* that explanation. To my mind, backing concerns the former.

Identifying the facts that determine the fact that *'A explains B' is an explanation,* is an interesting (albeit different) question. There are two versions of this question. The

---

[10] The point that causal explanation and causation are significantly different kinds of entities goes all the way back to Davidson (1967). Still, there are unorthodox exceptions (cf. Kim 1981).

first reading concerns the fact that 'A explains B' is a *successful* explanation (Sjölin Wirling [2020: 4] seems to adopt this reading). But this fact is not 'backed' in the original sense. Rather, it holds in virtue of the fact that 'A explains B' is *fully backed* (in the sense that its success conditions are fully met). The second reading concerns the fact that *'A explains B' is an explanation* (successful or not). I cannot settle this question here, but I am inclined to think that facts about the explanatoriness of a putative explanation are grounded in certain of its structural features. For example, on Kim's proposal (1994), explanations are propositions that have three 'slots': one for the explanans (or explanantia), one for the explanandum, and one for the determination-relation that connects the underlying phenomena (more on this later). At any rate, I take the question concerning the nature of the relations that are involved in these two readings to be an open question (and distinct from the issue concerning the nature of backing).

Finally, it could be argued that even if Sjölin Wirling is wrong about the exact details of the explanation-relatum, backing is still mono-categorical in the following sense (bracketing translatability worries): Backing could be the relation between *the fact that A grounds B* and *the fact that A explains B*.[11]

---

[11] It could be suggested that given that grounding and explanation typically relate facts, we can infer that backing relates facts. I have two responses. First, this inference can be doubted since backing doesn't relate the *relata* of a given explanation but that explanation *as a whole*. To compare, explanations might relate propositions, but it doesn't follow that explanations are propositions *themselves*. They could, for example, be arguments as per Hempel (1965). Secondly, we should resist the claim that explanation relates facts. Bracketing controversial views, standard accounts in the literature take the relevant relata of explanation to be sets of interpreted sentences (Kitcher 1989) or constituents of propositions (Kim 1994). The so-called ontic view of explanation *might* be a plausible candidate for the view that the relata of the explanation-relation are facts (in saying this I assume that such facts would be *worldly* in the same way the typical relata of the grounding relation are). Still, as it will become apparent, the ontic view is incompatible with separatism (see section 2.1.).

I have the following response. The cross-categoricity of backing should be understood in a general, non-fetishistic, way. A relation is cross-categorical when it relates entities that are importantly different from one another, even if those entities belong (strictly speaking) to the same ontological category. As noted, grounding relations are worldly items whereas explanations are representational entities about those items. In this sense, even if explanations are facts, the kinds of facts that they are differ significantly from the kinds of facts that grounding-facts are.[12]

2.1. On the nature of explanations

(Functional), in its current state, is non-trivial but not very informative. Its content will become more determined once we answer the following question: What sort of *entity* is an explanation? There are two plausible types of answers to this question. The first view takes explanations to be vessels that report or represent determination-facts. For example, some take explanations to be arguments (Kitcher 1989) or sets of propositions (Kim 1994). Alternatively, according to the so-called ontic conception, explanations are determination-facts *themselves*. In this sense, explanations are (quite literally) discovered in the same way causal events are discovered.[13]

---

[12] To compare, the mono-categoricity of explanation goes beyond the sameness in the ontological category of its relata. Take the classic Hempel view (1965) according to which an explanation is a relation between a set of sentences involving reference to a law of nature, and a sentence having a nomologically expected state as its content. Both relata occupy the same ontological category, while also being importantly similar: they both have a representational function vis-à-vis the phenomenon under examination.

[13] Conceptions of explanations should be distinguished from *theories* of explanation (Bokulich 2016: 263; Wright & van Eck 2018: 998). Theories propose a model of how explanations are supposed to work (e.g., law-based theories, à la Hempel, take reference to laws to be a constitutive feature of explanations).

The ontic conception cannot accommodate the cross-categoricity of backing and is incompatible with separatism. According to the ontic view, the explanation of B in terms of A is *identical* to a fact like A grounds B. If explanations *just are* determination-facts, then there is nothing separating them from those very determination-facts. So, it seems that the ontic view is a unionist-friendly account. Instead, the thesis that we should plug in to (Functional) to get a more substantive account is the one that identifies explanations with vessels that have some sort of representational or reporting capacity.[14]

Simply saying that explanations are vessels that report determination-facts does not get us very far since this idea can be cashed out in many different ways. For example, explanations could be arguments, sets of propositions, models, or something along those lines. Still, I think there is a way forward by taking explanations to be sets of propositions. This is not a costly assumption. For one, it is not clear that the competitors of such a view have anything more to offer in terms of features. The only contestable feature I can think of concerns the datum that many explanations are complex. The view that takes explanations to be arguments can accommodate this datum rather naturally. An argument can have many different premises each capturing a different kind of explanans: singular causes, background conditions, laws of nature, etc.

Conceptions, on the other hand, concern the *metaphysics* of explanations (i.e. what explanations *are* in the most literal sense). For interpretative issues see Craver (2019).

[14] For this reason, we can reject Glazier's (2020: 125) recent proposal that backing is a version of the part-whole relation. If explanations were themselves worldly facts, then it could (perhaps) be said that an instance of a determination *relation* literally composes an explanation. But this view does not seem to deliver the right results once one understands explanations as vessels that report. It is a category mistake to claim that instances of determination-relations are literally proper parts of the entities that represent them.

But I see no reason why the explanations-as-propositions view cannot deliver the same result. Propositions can be complex or structured too.

3. Backing as Explaining

So what would the relation between an explanation and its backer be? The first proposal I will examine has been advanced by Kovacs (2017: 2934; 2019a: 6) who identifies 'backing' with 'explaining':

(B=E) [A grounds B] backs 'A explains B', iff, [A grounds B] *explains* 'A explains B'.[15]

For example, according to (B=E), an explanation of a specific mental state in virtue of a brain state, is *explained* by the fact that the latter grounds the former.

I have two worries against this proposal. The first worry concerns the cross-categoricity of backing. As mentioned, the grounding separatist view does not simply claim that explanations are distinct from their corresponding grounding-facts. Specifically, the claim is that explanations are separate from grounding-facts and different *in kind*. But (B=E) cannot accommodate this datum. I take it that the explanation-relation obtains between the explanandum and the relevant explanans (or explanantia). In this sense, every theory of explanation takes explanations to relate the same kind of entities. As mentioned, some views take explanations to be arguments:

---

[15] Kovacs cites Rodriguez-Pereyra (2005: 28) and Schnieder (2010: 326–8) as others who also share this view. But it is not so clear from these passages that they do. Rodriguez-Pereyra (2005: 18) takes explanations to be truth-apt and backing to be a form of primitive grounding. But primitive grounding is arguably distinct from the explanation-relation unless one is a unionist about grounding in this latter instance (which is bizarre and unmotivated). Similarly, Schnieder's view is that backing is grounding, but grounding is construed as a connective and not as a reified relation (2010: 327; see also MacBride [2020: 3.5]). I consider backing-as-grounding in the next section.

explanantia and explananda are both propositions or sets of sentences (Kitcher 1989). Others take the explanation-relation to be a connective: explanantia and explananda are non-reified, linguistic, entities (Schnieder 2010).[16] Or, perhaps, explanations are relations between worldly items as per the so-called ontic conception of explanation.[17] No theory of explanation, which I know of, takes (or can take) the explanation-relation to be cross-categorical.[18]

My second worry against (B=E) appeals to a particular version of the separatist view. If the separatist view is that *every* successful explanation is backed, and if to be backed is to be explained, then we get the following thesis:

(Explanation Maximalism) Every successful explanation is explained.

But (Explanation Maximalism) leads to a regress since it requires the existence of an infinite series of explanations.

---

[16] One might worry that some representational accounts of explanations *are* cross-categorical. For example, a model of some phenomenon (a representation) would, presumably, explain that very phenomenon (a worldly entity). This is false. A model involves an explanandum and some explanantia in the same way an explanatory proposition or an explanatory argument does. In this sense, the explanation *relation* obtains between those entities (which, in turn, are literally *components* of that model), not between the explanation and the phenomenon itself.

[17] I addressed the ontic conception in the previous section.

[18] Sjölin Wirling (2020: 7) prefers an interpretation of (B=E) according to which the relevant 'is' is not the 'is' of identity. Rather, the slogan "backing is explanation" should be read elliptically as the claim that the backing-relation is explanatory. But this, in turn, is compatible with the claim that such explanations are underwritten by some other relation (with which backing *would* be identical). I agree that the explanatoriness of backing can be cashed out in this way (I return to this point in section 7). But Kovacs's view is that backing is, *literally*, the explanation-relation. After all, this is what allows him to respond to one version of the meta-grounding problem (i.e. the problem of what explains facts of the form: [A explains B]) in a new and interesting way (2019a: sec. 4.3.).

The proponent of (B=E) might respond by saying that the resulting regress is benign. But this move does not come cheap. First, it is not that clear what counts as a vicious regress. Some believe that vicious regresses are regresses that have some specific structural feature. Others disagree and take viciousness to be context-relative. So, it seems that in order for the proponent of (B=E) to claim that the relevant regress is benign, additional (and controversial) assumptions about the nature of regresses must be made.[19] Secondly, even if we accept that (Explanation Maximalism) leads to a benign regress, it is still the case that (B=E) has to countenance a regress. And even benign regresses have implications for one's ontology. Specifically, (Explanation Maximalism) leads to the existence of infinite explanations. This is still a cost for (B=E).[20]

The obvious move for the proponent of (B=E) would be to reject (Explanation Maximalism) and claim that only *some* successful explanations are explained. Such a view is entailed by the weaker separatist view that only some successful explanations are backed. This move avoids the regress worry. It does so, however, at a significant cost. Again, more questions are raised. Which explanations are not backed? And (once we identify them) why is it the case that *these* explanations are as such that they do not require a backer? Again, I take these questions to not have obvious or easy answers.

---

[19] The most straightforward way of doing this would involve claiming that a regress is benign in virtue of its objective features *and* showing that (Explanation Maximalism) leads to a regress with such features. Alternatively, one could adopt a context-relative account of benign regresses and claim that (Explanation Maximalism) leads to a benign regress *in the relevant contexts*. For example, Maurin (2013) argues that a series is vicious insofar as the task that that series was *supposed* to fulfil isn't met. So, now the question arises: what kind of task is a second-order explanation supposed to fulfil? Again, this indicates that, either way, the proponent of (B=E) needs to do a lot of additional work in order to claim that (Explanation Maximalism) leads to a benign regress.

[20] Of course, evaluating the severity of such a cost requires *further* assumptions about the relevant metric of parsimony. Also, it is worth mentioning that Kovacs, in other work, accepts that countenancing regresses (regardless of their form) counts against a theory (2019b: 7-8).

Finally, and perhaps more importantly, accepting that only some explanations are explained requires the existence of successful but *unbacked* (or *bare*) explanations (bare explanations, henceforth). Bareness should be distinguished from fundamentality. An explanation can be fundamental but still backed in the sense that it holds in virtue of a fundamental fact. Bare explanations, on the other hand, are 'naked': they do not hold in virtue of anything else. Some successful explanations would have to be unbacked for the regress worry to be avoided.

Bare explanations are controversial in themselves but, more importantly, they seem to be in tension with separatism. This is a significant cost for (B=E). Separatism wants to capture the intuition that successful explanations are higher-order, macro, phenomena that do not 'float-free'. Bare explanations go against this intuition. To be clear, I am not claiming that the existence of bare explanations is *incompatible* with separatism. If separatism is construed minimally as the thesis that explanations are separate from their backers, then bare explanations do not pose a threat (bracketing their controversial nature). My point is that positing successful unbacked explanations is in tension with a philosophically *attractive* version of separatism. Such a view would say that *every* successful explanation is backed, in the same way physicalists say that *every* state is nothing over and above a physical state. In this sense, proponents of (B=E) who wish to argue that only some successful explanations are explained, need to adopt an unattractive version of separatism.[21]

---

[21] The proponent of (B=E) would need to show in what way, if any, the principle that every explanation is backed should be rejected. So far, I have argued that the rejection of that principle generates more questions than it answers. Also, a philosophically attractive version of separatism takes every successful explanation to be backed. In this sense, it is not just that backing has some pre-theoretical content that puts pressure on (B=E). There are also philosophical reasons to *prefer* a notion of backing with the features I have highlighted.

4. Backing as Grounding

The natural alternative to (B=E) is the view that backing is a form of grounding. As previously noted, explanations hold *in virtue of* the facts that back them. Since in-virtue-of talk and priority-talk are often taken to be indicators for the existence of grounding relations, backing-as-grounding is a view that deserves consideration:

> **(B=G)** [A grounds B] backs 'A explains B', iff, [A grounds B] *grounds* 'A explains B'.

Saying that the explanation of a mental state in terms of its corresponding brain state is backed by the relevant grounding-fact, means that the former is grounded by the latter. Still, the proponent of (B=G) needs to do a lot of additional work in order to make (B=G) an illuminating definition of backing.

Before critically evaluating (B=G) I should make the following clarification. The most developed version of (B=G) has been recently proposed by Sjölin Wirling (2020).[22] However, her account is different from my understanding of (B=G) in two ways. First, as already noted, she takes backing to relate facts of the form [A grounds B] with facts concerning the explanatoriness of the proposition 'A explains B'. Secondly, she defines backing as *partial* grounding.

I have addressed the first point in section 2. Concerning the second point, I agree with Sjölin Wirling that if backing is a form of grounding, then grounding-facts *partially* ground their corresponding explanations.[23] It is plausible that for an

---

[22] Wilson (2016), Thompson (2018), and Kovacs (2019a) also consider (B=G).

[23] Naturally, Sjölin Wirling would frame things using her preferred backing relatum (i.e. facts about the explanatoriness of propositions, instead of explanations themselves). Still, my formulations are

explanation's success-conditions to be fulfilled certain epistemic/pragmatic conditions should also be in place (in addition to the relevant grounding-facts). But I see no reason to insist that the first relatum of backing should *always* be a grounding-fact. In this sense, these epistemic/pragmatic conditions would *also* be partial grounds. And such conditions *together* with the relevant grounding-fact would *fully* ground the relevant explanation. In this sense, backing can't be *defined* as partial grounding since there are clear instances of backing being underwritten by full grounding relations (i.e. the ones involving the *conjunction* of a grounding-fact with its appropriate epistemic/pragmatic conditions).[24]

That being said, I take these two differences to not be particularly important. The core of Sjölin Wirling's theory (i.e. that backing is a form of grounding) is independent from these auxiliary assumptions and deserves serious consideration. Also, as it will become apparent, the challenges I will raise against (B=G) also apply to Sjölin Wirling's specific view. For this reason, my focus on (B=G) is dialectically acceptable.

Appeals to grounding can be understood in many ways. I will divide them into two groups: reductive and non-reductive accounts. The non-reductive view takes grounding to be primitive (Audi 2012). Proponents of this view argue that grounding resists analysis but is still a valuable metaphysical tool. Perhaps the most important merit of this view is that it can accommodate the thesis that every successful explanation

---

dialectically acceptable since I take the claim that backing is partial grounding to be largely orthogonal to whether backing relates explanations or facts about such explanations.

[24] The point that such epistemic/pragmatic conditions are partial grounds is in line Sjölin Wirling's own remarks (2020: 5). Still, it could be argued that even if such conditions are partial grounds, they should not be understood as *backers* of the relevant explanation (the same worry might also apply to the *conjunction* of these conditions and the relevant grounding-fact). Presumably this is because the first relatum of backing is supposed to be restricted to grounding-facts or determination-facts broadly construed. But, again, I do not see the motivation for this restriction.

is backed without generating a regress. The resulting picture would simply be the view that every successful explanation is grounded.

Still, non-reductive versions of (B=G) have an important shortcoming: they are uninformative. First, if grounding is primitive and backing is a form of grounding then (B=G) simply collapses to the view that backing is ultimately unanalyzable.[25] In this sense, it simply is not illuminating enough to say that backing is a form of grounding. Instead, we should prefer views which reduce backing to more familiar phenomena. Secondly, the grounding literature is still relatively young in comparison to other literatures. This unsurprisingly leads to a lack of consensus concerning many of the features of grounding (cf. Rodriguez-Pereyra 2015). In this sense, it would be better to appeal to a less controversial entity when we wish to define backing.

At this point it could be objected that, for separatists, the existence of grounding is already assumed. In this sense, separatists would not take grounding to be unfamiliar. This is certainly true for *some* separatists. But, as noted in section 1, backing is a phenomenon that appears in other domains as well (e.g. causal explanation). One needn't believe in the existence of grounding in order to be a separatist. In this sense, a separatist about causal explanation who is also a grounding-skeptic would find (B=G) unconvincing.

Finally, another way to illustrate the fact that (B=G) is uninformative is by considering that a characterization of grounding *in terms of backing* would be explanatorily underwhelming.[26] Even if one does not aim to *define* grounding in terms

---

[25] Of course, taking backing itself to be primitive is different from taking backing to reduce to grounding and *then* taking grounding to be primitive. In the latter instance, backing reduces to something 'more familiar' which, among other things, has a primitive nature. In this sense, (B=G) is superior to the view that backing is primitive simpliciter in terms of ontological parsimony.

[26] Versions of this objection can be found in Wilson (2016) and Thompson (2018).

of its backing-role towards explanation, it would be nice to be able to *characterize* grounding in terms of its function. Grounding is the relation that, among other things, backs explanations. If backing is a form of grounding, then that way of illuminating grounding would not be available anymore.

I should note that I agree with Sjölin Wirling (2020: 7) that a circular characterization of grounding is *permissible*. She presents the following case: the property of being a blood relative is realized by the property of parenthood. But it is plausible that in characterizing blood relatedness, one would refer to the property of parenthood. Indeed, it would not "*obstruct* understanding" (Sjölin Wirling [2020: 7]; my emphasis) to say that grounding is that which grounds certain kinds of explanations. My weaker point is that it would be *desirable* to be able to illuminate grounding in this *additional* way (i.e. by appealing to its function to back explanations). In this sense, a characterization of grounding in terms of backing qua grounding would be uninformative (albeit, possible).

What about reductive versions of grounding? A promising way of reducing grounding is by appealing to the set of, what Wilson calls, 'small-g' relations (2014). A small-g relation is, roughly, a determination relation that has a very similar profile to grounding. For example, the relation of constitution is also said to capture priority-talk and is also taken to be explanatory.

A reductive account of grounding in terms of small-g relations can be articulated in two ways. The first one would be to say that backing is a form of grounding and grounding is multiply realizable by different small-g relations (Wilson 2014: 567-8; Rettler 2017). In this sense, backing would be identical to the *realized*

entity which would entail that backing *itself* is multiply realizable by small-g relations.[27] This option is very implausible. Many paradigmatic small-g relations are not plausible realizers of backing (e.g. set-formation).

A more promising way of understanding backing as a kind of reduced grounding is by taking backing to be identical to one of the (potentially many) *realizers* of grounding. In this sense, (B=G) would be the view that backing is identical to some small-g relation. Still, it is impossible to assess this proposal without specifying which small-g relation backing is supposed to be. For this reason, (B=G) is unilluminating and, at best, only the first step towards an informative definition of backing.

5. Backing as Truthmaking

In this section I propose a new definition of backing. According to (Functional), the backing relation connects grounding-facts with explanations. Also, explanations, as mentioned in section 2, are certain kinds of propositions. But what kind of relation connects propositions with grounding-facts? A very plausible candidate for this role is the truthmaking relation.

> **(B=T)** [A grounds B] backs 'A explains B', iff, 'A explains B' is *made true by* [A grounds B].

(B=T) satisfies all three clauses of (Functional). It is a non-trivial and substantive issue whether a given worldly fact serves as a truthmaker for a proposition. Also, truthmaking is a cross-categorical relation par excellence (Tahko 2016). Finally, it is plausible that in

---

[27] I think this is the most promising reductive theory of grounding currently in the literature (hence my focus). Another option would involve reducing grounding to essences (e.g. Correia & Skiles 2019). I do not have the space to properly discuss this interesting alternative.

order for an explanation to be successful it is (at least) necessary for that explanation to be true.

What is the nature of the truthmaking relation? According to Armstrong's (2004) classic picture, truthmaking is the internal relation of necessitation. As a quick example, take the following psychophysical explanation:

> **(M)** The fact that Mary has mental state M_1 is non-causally explained by the fact that Mary has brain state B_1 and that a psychophysical law connecting M-type mental states with B-type brain states obtains.

What backs (M)? Plausibly, worldly facts such as the relevant psychophysical law partially grounding the fact that Mary has mental state M_1. Adopting Armstrong's account, this means that Mary's brain state B_1 together with the relevant psychophysical law necessitate the truth of (M). On the face of it this seems plausible. Of course, there are well known objections to the view that truthmaking is the relation of necessitation which I am not going to rehearse here (MacBride 2020: 1.2.). Thankfully, (B=T), on its own, takes no particular stance towards the specifics of the truthmaking relation. In this sense, I intend for my proposal to be minimal. Truthmaking, in this context, is supposed to designate whichever relation is posited by one's background truthmaker theory.

Deferring to one's background truthmaker theory is dialectically acceptable. Figuring out which truthmaker theory is the correct one is a substantive philosophical problem that requires our attention *anyway*, independently of the discussion around backing, grounding, and the separatism/unionism debate. In this sense, the fact that the

final details of my proposal are 'hostage' to one's background truthmaker theory is not a significant cost.[28]

The minimality of (B=T) is worth emphasizing in at least three additional ways. For one, nothing in (B=T) entails a robust notion of truthmakers. Opponents of truthmaker theory usually argue that one can make sense of truth being dependent on reality without having to posit separate entities called 'truthmakers'. (B=T) is neutral towards this objection. One is free to adopt a thin understanding of truthmakers according to which such entities are not separate and distinct *facts* but are simply those bits of reality that account for the truth of the relevant proposition.[29]

Secondly, there is an ongoing discussion about the level of *grain* of the truthmaking relation. For example, some opponents of the correspondence theory of truth argue that a robust understanding of correspondence (perhaps, as causal regulation) is explanatorily superfluous. Again, (B=T) is neutral towards this discussion. Following Armstrong, proponents of (B=T) can take the truthmaking relation to be an internal relation without reifying it by appealing to some external determination relation.

Thirdly, so far, I have been focusing on backing as a relation that holds between explanations and grounding-facts. But, as mentioned in section 2, there is no reason why backing should be constrained in this way. Truthmaking, as such, can be used to

---

[28] The relevant literature has made a number of proposals concerning the nature of truthmaking. Perhaps, truthmaking is a version of the supervenience relation (Heil 2016) or, as mentioned, some kind of necessitation (Armstrong 2004). Truthmaking could also be a version of grounding (cf. notes 29 and 33).

[29] See Rodriguez-Pereya (2005: 24) and MacBride (2014: 378) for the distinction I am appealing to. Truthmaking relations which do not seem to require a robust ontological category of truthmakers include most traditional correspondence-relations (e.g., Barnard & Horgan 2006: 29) and deflationary truthmaking (Hornsby 2005). For discussion see Asay & Baron (2019: 10).

define diverse metaphysical theses (from anti-realism to realism about a given domain). Analogously, an explanation could be backed by determination-facts (broadly construed), nomic facts, or facts about essences. But they could also be backed by facts which are not worldly, at least in the traditional sense. For example, it could turn out that explanations are true in virtue of the fact that they are appropriately related to other explanations. In a case like this, such explanations would be backed by the fact that they figure in an appropriately defined set of explanations.[30]

Finally, (B=T) is more informative than both construals of (B=G) while staying minimal towards the details of the truthmaking relation itself and pluralistic concerning the different kinds of entities that can serve as truthmakers. In this sense, (B=T) could be compatible with (B=G) *if* one takes grounding to be reducible and truthmaking to be one of its realizers. Still, such assumptions are controversial, and I take it that (B=T) has merit independently of its connection to grounding. This is particularly evident if one considers that skepticism about truthmaking is, in general, more controversial (and less widespread) than skepticism about grounding. As mentioned, grounding is a relatively new metaphysical tool whereas my notion of truthmaking is minimal. Under these considerations, I take it that (B=T) should be the default view concerning the nature of the backing relation. In the following two sections I consider two objections against my view.

6. The Inheritance Worry

Grounding theorists often appeal to the nature of explanation to justify the fact that grounding has certain formal features. Grounding is supposed to be irreflexive

---

[30] As Daly (2005: 97) correctly notes, coherentists about truth can argue that a true proposition is 'backed' by the fact that this proposition is a member of a maximally coherent set of propositions. For the analogous view on explanation see Kovacs (2019b).

because explanation is irreflexive (grounding *inherits* its formal features from explanation) (Maurin 2019). However, it could be objected that (B=T) is implausible because the truthmaking relation (in any of its forms) cannot plausibly fulfil the inheritance function. Specifically, it is implausible that truthbearers have the same formal features as their truthmakers.

To illustrate, compare truthmaking to relations that plausibly *do* license inheritance claims. The identity relation is the most obvious example. If explanations are identical to grounding-facts, then it follows that they have the same formal features. A less trivial case can be found in the literature on mental causation. Some philosophers argue that mental states hold in virtue of brain states and, because of that, the causal profile of the latter is transmitted to the former.[31] But backing is neither the identity relation nor the relation that holds between minds and brains.

There are two ways in which (B=T) can be defended against this. The first strategy involves directly meeting the challenge by arguing that truthmaking *can* license inheritance claims. To that effect, it could be claimed that truthmaking should be understood as an *isomorphic* relation. More specifically, truthmaking could be isomorphic in the sense that truthbearers are structurally identical to their corresponding truthmakers. The *proposition* that the cat is on the mat involves a certain relation between its constituents that 'mirrors' the structure of the *fact* that the cat is on the mat. If this is true, then those explanations that are backed plausibly share the same formal features with the grounding-facts that back them.

Against this move, it should be noted that knowing about the features of explanation cannot tell us anything about grounding-facts that *do not* serve as

---

[31] Such inheritance theses are usually articulated by appealing to functional realization or immanent causation (Robb & Heil 2018: 6.4).

truthmakers. Looking into the features of explanation (asymmetry, irreflexivity, etc.) can only serve to illuminate the features of those (and only those) grounding-facts which back explanations. In this sense, even if one assumes that truthmaking is isomorphic, this at best provides only limited access to the inner workings of grounding. Also, the very claim that truthmaking is isomorphic is far from obviously true (Schipper 2020).

So far, I have been mostly focusing on how explanations relate to their corresponding grounding or determination-facts. But, as previously mentioned, there is no good reason to restrict backing in this way. Backing is supposed to be the relation that connects explanations with *whichever* entity makes them true. Assuming isomorphism would beg the question against views that allow for successful explanations to be true but cash out their success holistically (to mention one example). Such views would say that an explanation of the form 'A explains B' is not made true by a grounding-fact of the form [A grounds B] but by a plurality of different facts that bear no isomorphic relation to the original explanation.

My preferred strategy involves denying that we should expect to find out about the nature of grounding by looking at the explanations which grounding-facts back. Grounding should be examined in the same way every other phenomenon is examined: by hypothesizing and testing such hypotheses against a background of a well-established body of beliefs. The analogy with causation is helpful at this point.[32] It is uncontroversial that there are causal events which are not explanatory. Still, there are other ways in which philosophers have proceeded to learn about the nature of causation. Being able to back causal explanations is *one* of the roles that causal relations have. Other roles include, for example, figuring in our best scientific and folk theories.

---

[32] For the tight connection between causation and grounding see Fine (2012), Schaffer (2016), A. Wilson (2018).

Philosophers, based on certain folk and scientific platitudes about causation, propose certain hypotheses about what the causal relation *really* is (e.g. counterfactual dependence, energy transfer, etc.). At best, looking into the nature of causal explanations can be one of the many ways in which we can uncover the true nature of causation.

7. The Regress Worry

Recall that one of my worries against (B=E) was that, given (Explanation Maximalism), a regress is generated. (B=T) is supposed to avoid this worry. But it could be plausibly argued that (B=T) also leads to a regress if one takes truthmaking to be an explanatory relation. If truthmaking is explanatory, then (B=T) seems to collapse into (B=E). A way to raise this worry is to take truthmaking to be a form of grounding and then take grounding to be explanatory. But the objector does not even need to make this move.[33] The link between truthmaking and explanation is usually highlighted in the relevant literature, so I take the claim that truthmaking is explanatory to be well-motivated independently of whether it is a grounding relation.

What does it mean to say that truthmaking is explanatory? Asay (2017: 10) identifies two possible answers: either truthmaking is *itself* an instance of the explanation-relation, or explanation is simply a constraint on truthmaking (i.e. P makes <P> true only if P *explains* <P>). The first option is a form of unionism about truthmaking whereas the second is a form of separatism about truthmaking.

---

[33] Especially since cashing out the relation between grounding and truthmaking is controversial (see Liggins [2012]; Tahko [2016]; Asay [2017]). For recent and direct rejections of the thesis that truthmaking is a case of grounding see Griffith (2014), Saenz (2018), and Audi (2019).

Ruling out the first option seems fairly easy. It is simply bizarre to be a separatist about grounding and a unionist about truthmaking. Even if one does not take truthmaking to be a form of grounding it could still be argued that the two notions are similar enough. So, separatism can be motivated for both relations on similar grounds.

Is grounding similar to truthmaking? It is not obvious that it is. As previously mentioned, it is controversial that truthmaking is a form of grounding. Perhaps, as Audi (2019) notes, truthmaking is not a grounding relation because it should not be understood as a *determination* relation. But if that's the case, then the putative similarity between the two relations becomes thin. Against this, I propose that separatism about both grounding and truthmaking should be motivated differently. Rather, separatism is primarily a view about *explanation*. Successful explanations are certain kinds of propositions which do not float free. So, insofar as one is a separatist about explanations involving grounding one should also be a separatist about explanations involving truthmaking. In other words, separatism about grounding and truthmaking are simply instances of separatism simpliciter.

The second option is trickier. According to this form of separatism, truthmaking and explanation are distinct, but every truthmaking-fact ([P makes <P> true]) has a corresponding explanation (<P explains <P>>). This version of separatism is different from the kind of separatism I have been considering so far (call my version, 'standard' separatism). An important and desirable feature of standard separatism is that grounding-facts can exist in the absence of explanation. In an analogous sense, separatism about truthmaking should allow for certain truthmaking-facts to exist in the absence of a corresponding explanation. Does this compromise the well-established link between truthmaking and explanation? It does not. It can still be said that truthmaking is explanatory even if it is not the case that *every* truthmaking-fact figures in an

explanation. Causation is also explanatory by anyone's lights even though there are clear cases of causal events which are not involved in causal explanation.

Still, one could wonder whether standard versions of separatism coupled with (B=T) can successfully deal with the regress worry. Can truthmaking-facts exist in the absence of explanation? In essence, the answer to this question depends on one's view about the ontological status of explanations (i.e. the ontological status of propositions). According to one influential view, propositions are reified abstracta that are expressed by particular sentences. Such views take explanations to be stance-independent in the sense that explanations exist independently of whether they are expressed or conceived by agents. But if that's the case, then the regress worry looms large again. Every truthmaking-fact would have a real, stance-independent, corresponding explanation. Such truthmaking-facts (e.g., the fact that [P makes <P> true] makes <P explains <P>> true) would, in turn, have their own corresponding explanations, and so on, *ad infinitum*. So, it seems that (B=T) faces the same worry as (B=E) in terms of generating regresses.

One thing to immediately notice is that even if we accept that (B=T) generates a regress, the *kind* of regress it generates is significantly different from the one (B=E) generates. (B=E) coupled with the thesis that every explanation is backed entails that every explanation is explained. This means that if an explanation exists, then an infinite series of explanations is generated. Under closer inspection, this is a regress that runs 'downwards'. A second-order explanation is more 'fundamental' than the explanation it explains (in the sense that first-order explanations *depend* on second-order explanation, second-order explanations depend on third-order explanations, and so on, *ad infinitum*). This means that the infinite regress (B=E) generates runs towards the fundamental level. (B=T), on the other hand, leads to an 'upwards' regress. In this case, the first element of the series is a truthmaking-fact which, in turn, gives rise to another

truthmaking-fact, and so on *ad infinitum*. This means that the order of dependence goes from a truthmaking-fact (of the form [P makes <P> true]) towards its corresponding 'meta' truthmaking-fact (of the form [[P makes <P> true] makes <P makes <P> true> true]). This is a significant difference because upward regresses are generally regarded as less problematic than downward regresses (Cameron 2008).

Still, it should be noted that there is a way for the separatist to deny that (B=T) generates a regress altogether. Specifically, the separatist should deny that propositions exist stance-independently.[34] Notice that if the existence of explanations directly depends on the deliberation of conscious agents, then there will not be an infinite series of explanations unless an agent can generate one. If a truthmaking-fact makes its corresponding explanation true, this presupposes that that explanation *exists*. The demand for a truthmaker arises only once the relevant proposition is formed. But according to the view I am considering, whether an explanation is formed depends on the relevant agents and their explanatory demands. To compare, (B=E) faces a regress-worry that cannot be avoided using similar means. It is independently plausible that every successful explanation requires a backer. (B=T) can meet that demand: explanations, insofar as they are formed, require truthmakers. Rather, the proponent of

---

[34] It could be objected that my solution is plausible insofar as explanations are propositions. But what if they are physical structures representing some phenomenon (say, two rocks representing the sun and the moon)? Surely, the rocks would continue to exist even in the absence of an explanation involving them. It is certainly possible to construe explanations in this way. But, sticking to the framework that explanations are propositions (entailing that the existence of that explanation is nothing over and above the existence of that proposition), it could be said that such physical systems illustrate or *express* the content of the relevant explanation (without being, themselves, explanations).

(B=T) should deny the converse thesis, namely, that every truthmaker has a corresponding explanation.[35]

I do not have the space to fully defend a stance-dependent view about the existence of propositions. Still, it suffices to say that such a view is both powerful and already assumed by many philosophers working on explanation. Kim (1988; 1999), for example, adopts a view of propositions akin to contemporary conceptualist theories (i.e. propositions as abstractions of mental tokenings) (King 2007).[36] The alternative would be to understand propositions as entities whose existence-conditions are completely independent of human agency. But such views are notorious for their inability to explain how they get their truth-conditions and their representational capacities. In comparison, the agent-dependent view has a clear-cut explanation: propositions get their truth-conditions and their ability to represent in virtue of the capacities of conscious agents.

There are ways to resist these points. But it should be noted that whichever way the relevant dialectic plays out, at least as far as regresses are concerned, (B=T) is superior to (B=E). If (B=T) generates a regress, then it is a regress that is more palatable than the one (B=E) generates. Finally, I have highlighted a plausible, albeit not completely non-parochial, way for the proponent of standard separatism to completely bypass the regress worry.

A final point about (B=T) and regresses. I have argued that (B=E) leads to a downward regress, whereas (B=T) leads to an upward one. But it could be argued that

---

[35] This also means that A grounding B is compatible with there being an *explanatory gap* between A and B (see Wilson 2016: 11). The fact that A grounds B does not entail, on its own, that A explains B.

[36] Philosophers adopting similar views include Armstrong (1997), Wright (2012: 375-6, 376), Koslicki (2012: 212-13), Wright & van Eck (2018: 1005). One can also trace this line of thinking in the literature on scientific representation (Giere 2010).

(B=T) leads to a downward regress *as well.* After all, we can still ask: what explains why [A grounds B] makes 'A explains B' true? And, then, we can ask what explains that further truthmaking-fact, *ad infinitum.* In this sense, (B=T) might not be superior to (B=E) in that regard after all.

Note that in order for (B=T) to generate a regress in the above sense, (B=T) needs to be coupled with the two following triggering statements:[37]

(1) Every truthmaking-fact has a ground.

(2) Every grounding-fact involving a truthmaking-fact (of the form "P grounds the fact that Q makes <Q> true") has a corresponding explanation (of the form "P explains the fact that Q makes <Q> true").

Both statements can be contested. (2) presumably follows from the principle that determination-facts always have explanations which they back. But I have already indicated that an attractive version of separatism would want to accept the possibility of determination-facts occurring in the absence of explanation. I also proposed a plausible way to deliver this result by taking propositions to be stance-independent entities.

(1) can also be resisted. Even if some truthmaking-facts have grounds, it is not obvious that this should apply to *every* truthmaking-fact. For all we know, some truthmaking-facts might be brute or fundamental. Perhaps, (1) can be motivated by appealing to a principle of sufficient reason according to which *everything* has a ground. Compare with (B=E). The triggering statement that (B=E) needs to be coupled with is, as noted, simply a plausible version of separatism: "Every successful explanation is backed.". But, surely, that principle is much less controversial than the principle that everything has a ground.

---

[37] A triggering statement is a statement that ensures that the relevant series of elements will not have a last member (Maurin 2013).

Separatism coupled with (B=T) entails that every successful explanation has a truthmaker, whereas separatism coupled with (B=E) entails that every successful explanation needs to be explained (which results in a regress). For this reason, the downward regress problem is avoided by (B=T) and is particularly pressing for the proponent of (B=E).[38]

8. Conclusion

The aim of this paper was to argue that instances of backing are instances of truthmaking. I examined and rejected two prominent proposals: backing as explaining and backing as grounding. Then I proposed that backing should be understood as an instance of truthmaking. I also understood truthmaking minimally. Finally, I responded to two objections against my proposal: the inheritance worry and the regress worry.

---

[38] It is also not so clear whether (B=T) coupled with (1) and (2) generates a *downward* instead of an *upward* regress (the latter of which, as noted, is generally understood as less controversial). The proposed regress runs as follows. First, a truthmaking-fact (TM) is grounded by some ground (G). Then, a grounding-fact of the form [G grounds TM] is formed. Then, that grounding-fact makes its corresponding explanation true. Then, that truthmaking-fact has a further ground, and so on, *ad infinitum*. But the dependencies involved in this regress run in multiple directions. The first element of the series (TM) *depends* on the second element (G). But, then, the emerging grounding-fact *gives rises* to its corresponding explanation (which, in turn, *depends* on a further ground, and so on). To compare, on a paradigmatic downward regress the first element depends on the second element, which depends on the third element, *ad infinitum*. So, even if one brackets the worries that I raised against (1) and (2), the resulting regress is not paradigmatically downward.

References

Armstrong, D. M. (1996). *A World of States of Affairs*. Cambridge University Press.

Armstrong, D. M. (2004). *Truth and Truthmakers*. Cambridge University Press.

Asay, Jamin & Baron, Sam (2020). Deflating Deflationary Truthmaking. *Philosophical Quarterly* 70 (278):1-21. https://doi.org/10.1093/pq/pqz036

Asay, Jamin (2017). Run Aground: Kit Fine's Critique of Truthmaker Theory. *Philosophical Quarterly* 67 (268):443-463. https://doi.org/10.1093/pq/pqw072

Audi, Paul (2012). Grounding: Toward a Theory of the In-Virtue-Of Relation. *Journal of Philosophy* 109 (12):685-711. https://doi.org/10.5840/jphil20121091232

Audi, Paul (2019). Why Truthmaking Is Not a Case of Grounding. *Philosophy and Phenomenological Research*. https://doi.org/10.1111/phpr.12605

Barnard, Robert & Horgan, Terence (2006). Truth as Mediated Correspondence. *The Monist* 89 (1):28-49. https://doi.org/10.5840/monist200689132

Bliss, Ricki and Trogdon, Kelly, "Metaphysical Grounding", The Stanford Encyclopedia of Philosophy (Winter 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2016/entries/grounding/>.

Bokulich, Alisa (2016). Fiction As a Vehicle for Truth: Moving Beyond the Ontic Conception. *The Monist* 99 (3):260-279. https://doi.org/10.1093/monist/onw004

Cameron, Ross P. (2008). Turtles all the way down: Regress, priority and fundamentality. *Philosophical Quarterly* 58 (230):1-14. https://doi.org/10.1111/j.1467-9213.2007.509.x

Correia, Fabrice & Skiles, Alexander (2019). Grounding, Essence, And Identity. *Philosophy and Phenomenological Research* 98 (3):642-670.

Craver, Carl F. (2019). Idealization and the Ontic Conception: A Reply to Bokulich. *The Monist* 102 (4):525-530.

Daly, Chris (2005). So where's the explanation? In Helen Beebee & Julian Dodd (eds.), *Truthmakers: The Contemporary Debate*. Clarendon Press. pp. 85.

Dasgupta, Shamik (2017). Constitutive Explanation. *Philosophical Issues* 27 (1):74-97. https://doi.org/10.1111/phis.12102

Davidson, Donald (1967). Causal relations. *Journal of Philosophy* 64 (21):691-703.

deRosset, Louis (2013). Grounding Explanations. *Philosophers' Imprint* 13. http://hdl.handle.net/2027/spo.3521354.0013.007

Fine, Kit (2012). Guide to Ground. In Fabrice Correia & Benjamin Schnieder (eds.), *Metaphysical Grounding*. Cambridge University Press. pp. 37--80.

Giere, Ronald N. (2010). An agent-based conception of models and scientific representation. *Synthese* 172 (2):269–281. https://doi.org/10.1007/s11229-009-9506-z

Glazier, Martin (2020). Explanation. In Michael J. Raven (ed.), *Routledge Handbook of Metaphysical Grounding*. London: pp. 121-132.

Glennan, Stuart (2002). Rethinking Mechanistic Explanation. *Philosophy of Science* 69 (S3):S342-S353. https://doi.org/10.1086/341857

Griffith, Aaron M. (2014). Truthmaking and Grounding. *Inquiry: An Interdisciplinary Journal of Philosophy* 57 (2):196-215. https://doi.org/10.1080/0020174X.2013.855655

Heil, John (2016). Truthmaking and fundamentality. *Synthese*:1-12. https://doi.org/10.1007/s11229-016-1292-9

Hempel, Carl (1965). Aspects of scientific explanation. In *Philosophy and Phenomenological Research*. Free Press. pp. 504.

Hornsby, Jennifer (2005). Truth without truthmaking entities. In Helen Beebee & Julian Dodd (eds.), *Truthmakers: The Contemporary Debate*. Clarendon Press. pp. 33.

Kim, Jaegwon (1973). Causation, nomic subsumption, and the concept of event. *Journal of Philosophy* 70 (8):217-236. https://doi.org/10.2307/2025096

Kim, Jaegwon (1981). Causes as explanations: A critique. *Theory and Decision* 13, 293–309. https://doi.org/10.1007/BF00126965

Kim, Jaegwon (1988). Explanatory Realism, Causal Realism, and Explanatory Exclusion. *Midwest Studies in Philosophy* 12 (1):225-239. https://doi.org/10.1111/j.1475-4975.1988.tb00167.x

Kim, Jaegwon (1994). Explanatory knowledge and metaphysical dependence. *Philosophical Issues* 5:51-69. https://doi.org/10.2307/1522873

King, Jeffrey C. (2007). *The Nature and Structure of Content*. Oxford University Press.

Kitcher, Philip (1989). Explanatory unification and the causal structure of the world. In Philip Kitcher & Wesley Salmon (eds.), *Scientific Explanation*. Minneapolis: University of Minnesota Press. pp. 410-505.

Koslicki, Kathrin (2012). Varieties of ontological dependence. In Fabrice Correia & Benjamin Schnieder (eds.), *Metaphysical Grounding: Understanding the Structure of Reality*. Cambridge University Press. pp. 186.

Kovacs, David Mark (2017). Grounding and the argument from explanatoriness. *Philosophical Studies* 174 (12):2927-2952. https://doi.org/10.1007/s11098-016-0818-9

Kovacs, David Mark (2019a). Four Questions of Iterated Grounding. *Philosophy and Phenomenological Research*. https://doi.org/10.1111/phpr.12591

Kovacs, David Mark (2019b). Metaphysically explanatory unification. *Philosophical Studies*:1-25. https://doi.org/10.1007/s11098-019-01279-z

Lange, Marc (2016). *Because Without Cause: Non-Causal Explanations in Science and Mathematics*. Oxford University Press USA.

Liggins, D. (2012). Truthmakers and dependence. In F. Correia, & B. Schnieder (Eds.), Metaphysical Grounding: Understanding the Structure of Reality (pp. 254-271). Cambridge: Cambridge University Press.

Litland, Jon (2016). Pure Logic of Many-Many Ground. *Journal of Philosophical Logic* 45 (5):531-577. https://doi.org/10.1007/s10992-015-9386-2

Litland, Jon Erling (2018a). Could the grounds's grounding the grounded ground the grounded? *Analysis* 78 (1):56-65.

Litland, Jon Erling (2018b). In Defense of the Disunity of Grounding. *Thought: A Journal of Philosophy* 7 (2):97-108.

MacBride, Fraser (2014). How truth depends upon being. Analysis 74 (3):370-378. https://doi.org/10.1093/analys/anu057

MacBride, Fraser, "Truthmakers", The Stanford Encyclopedia of Philosophy (Spring 2020 Edition), Edward N. Zalta (ed.), forthcoming URL = <https://plato.stanford.edu/archives/spr2020/entries/truthmakers/>.

Maurin, Anna-Sofia (2013). "Infinite Regress Arguments'" in C. Svennerlind, J. Almäng & R. Ingthorsson (eds.), *Johanssonian Investigations: Essays in Honour of Ingvar Johansson on His Seventieth Birthday*, New Jersey: Transaction Books, pp. 421–438.

Maurin, Anna-Sofia (2019). Grounding and metaphysical explanation: it's complicated. *Philosophical Studies* 176 (6):1573-1594. https://doi.org/10.1007/s11098-018-1080-0

Raven, Michael J. (2015). Ground. *Philosophy Compass* 10 (5):322-333. https://doi.org/10.1111/phc3.12220

Rettler, Bradley (2017). Grounds and 'Grounds'. *Canadian Journal of Philosophy* 47 (5):631-655. https://doi.org/10.1080/00455091.2017.1282306

Robb, David and Heil, John, "Mental Causation", The Stanford Encyclopedia of Philosophy (Summer 2019 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2019/entries/mental-causation/>.

Rodriguez-Pereyra, Gonzalo (2005). Why Truthmakers? In Helen Beebee & Julian Dodd (eds.), Truthmakers: the contemporary debate. Oxford University Press. pp. 17-31.

Rodriguez-Pereyra, Gonzalo (2015). Grounding is not a strict order. *Journal of the American Philosophical Association* 1 (3):517-534. https://doi.org/10.1017/apa.2014.22

Rosen, Gideon (2010). Metaphysical Dependence: Grounding and Reduction. In Bob Hale & Aviv Hoffmann (eds.), *Modality: Metaphysics, Logic, and Epistemology*. Oxford University Press. pp. 109-36.

Ruben, David-Hillel (1990). Explaining Explanation. London: Routledge

Saenz, Noël Blas (2018). An account of truthmaking. *Synthese*:1-23. https://doi.org/10.1007/s11229-018-1894-5

Schaffer, Jonathan (2016). Grounding in the image of causation. *Philosophical Studies* 173 (1):49-100. https://doi.org/10.1007/s11098-014-0438-1

Schipper, Arthur (2020). Aboutness and ontology: a modest approach to truthmakers. *Philosophical Studies* 177 (2):505-533. https://doi.org/10.1007/s11098-018-1192-6

Schnieder, Benjamin (2010), "A Puzzle About 'Because'," Logique et Analyse, 53: 317–343

Skiles, Alex & Trogdon, Kelly (2019). Maurin on Grounding and Explanation. In Stenwall, Robin & Hansson Wahlberg, Tobias (eds.) (2019). *Maurinian Truths : Essays in Honour of Anna-Sofia Maurin on her 50th Birthday*. Department of Philosophy, Lund University.

Tahko, Tuomas E. (2016). Armstrong on Truthmaking and Realism. In Francesco F. Calemi (ed.), Metaphysics and Scientific Realism: Essays in Honour of David Malet Armstrong. De Gruyter. pp. 207-218.

Taylor, Elanor (2018). Against explanatory realism. *Philosophical Studies* 175 (1):197-219. https://doi.org/10.1007/s11098-017-0862-0

Taylor, Elanor (2020). Backing without realism. *Erkenntnis.* https://doi.org/10.1007/s10670-020-00249-w

Thompson, Naomi (2018). Metaphysical Interdependence, Epistemic Coherentism and Holistic Explanation. In Bliss, R. and Priest, G. (eds) *Reality and its Structure* (Oxford: Oxford University Press).

Trogdon, Kelly (2013). An introduction to grounding. In Miguel Hoeltje, Benjamin Schnieder & Alex Steinberg (eds.), *Varieties of Dependence.* Munich, Germany: Philosophia Verlag. pp. 97-122.

Wallner, Michael (2018). The ground of ground, essence, and explanation. *Synthese.* https://doi.org/10.1007/s11229-018-1856-y

Wilson, Alastair (2018). Metaphysical Causation. *Noûs:*723-751. https://doi.org/10.1111/nous.12190

Wilson, Alastair. (2020). "Classifying Dependencies". In The Foundation of Reality: Fundamentality, Space and Time, eds. David Glick, George Darby and Anna Marmodoro. Oxford University Press.

Wilson, Jessica M. (2014). No Work for a Theory of Grounding. *Inquiry: An Interdisciplinary Journal of Philosophy* 57 (5-6):535-579. https://doi.org/10.1080/0020174X.2014.907542

Wilson, Jessica M. (2016). Grounding-Based Formulations of Physicalism. *Topoi* 37 (3):495-512. https://doi.org/10.1007/s11245-019-09644-7.

Sjölin Wirling, Ylwa (2020). Is backing grounding? *Ratio* 33 (3):129-137.

Wright, Cory & van Eck, Dingmar (2018). Ontic Explanation Is either Ontic or Explanatory, but Not Both. *Ergo: An Open Access Journal of Philosophy* 5:997–1029. http://dx.doi.org/10.3998/ergo.12405314.0005.038

Wright, Cory D. (2012). Mechanistic explanation without the ontic conception. *European Journal of Philosophy of Science* 2 (3):375-394. https://doi.org/10.1007/s13194-012-0048-8

Part 2 | Grounding and Metaethics

Chapter 3 | Nomic Moral Naturalness

Abstract: Moral realists often disagree about the nature of moral properties. These properties can be natural (as per naturalistic moral realism) or non-natural. But it is unclear how we should understand the notion of naturalness employed in these discussions. In this paper I propose a novel account of moral naturalness. I suggest that a property F is natural iff F falls within the scope of a natural law. In turn, a law is natural when it figures in a nomic nexus involving the laws of physics. In section 2 I present three desiderata that should constrain a theory of moral naturalness. In section 3 I present a "two-step" schema according to which a property F is natural iff F is identified with some paradigmatically natural "base" properties, or F is appropriately related to these base-properties. In section 4 I examine two grounding-based proposals concerning the second clause of this schema and argue for their inadequacy. In section 5 I present the view that moral naturalness should be defined in terms of natural lawhood, and the naturalness of a law should be defined in terms in whether that law figures in a nexus involving the laws of physics.

1. Introduction

The aim of this paper is to propose a taxonomy that distinguishes between naturalist and non-naturalist versions of moral realism. The demand for such a taxonomy arises from the curious fact that there is no apparent consensus on what

exactly differentiates these views. In this sense, the debate between different moral realists is compromised. It might turn out that many moral realists talk past each other given that they do not operate under a common set of definitions.[1]

I will take moral realism to be a view about the metaphysical status of the truth-makers of moral truths. The moral realist argues that at least some first-order moral judgments are true, and their truth holds independently of the stances of moral agents. In this context, stance-independence is to be understood (crudely) as independence from the epistemic situation, interests, beliefs, and theories of moral agents. The way a moral realist proceeds in characterizing the nature of the entities in virtue of which these truths hold determines the version of realism they adopt. The most straight-forward way of articulating that difference seems to be in terms of properties. For example, a non-naturalist moral realist would claim that at least some moral judgments are true independently of our stances, in virtue of instances of non-natural properties.

Of course, little progress would have been made without a theory that distinguishes natural from non-natural properties in a useful and substantive way. In this paper I propose a novel view according to which a property is natural when it falls within the scope of a natural law (roughly, a law that figures in a nomic nexus involving the laws of physics). In section 2, I present and defend three key desiderata that should constrain a theory of moral naturalness. Specifically, I argue that a taxonomy based on such a theory should be as metaphysically interesting, non-revisionary, and non-parochial as possible. Then, I propose a schema according to which a natural moral property is either identified with some appropriately

---

[1] It is plausible that the relevant notion of naturalness would also apply to non-realist views. But I will focus on moral realism as this is how the discussion is usually framed in the literature (see, e.g., McPherson [2015]).

defined *base*-type property, or it is defined *indirectly* in terms of its relation to these base-properties. In section 3, I consider several candidates that would serve as base-properties and conclude that a plausible candidate is the set of physical properties. In section 4, I consider a grounding-based version of this schema and argue for its inadequacy. Finally, in section 5, I develop my proposed notion of nomic naturalness. In section 6, I develop further my account by tackling some potential objections.

2. Desiderata and motivation

An intuition shared by many philosophers (e.g. Enoch 2011: 363; McPherson 2012: 6; Leary 2017: 78; Rosen 2017a: 151) is that the debate between different versions of moral realism is not merely linguistic. On the contrary, it is supposed to be a distinction which identifies a metaphysically interesting fact about different classes of properties. For example, according to the non-naturalist, moral properties form an autonomous class of properties that is discontinuous from the class of natural properties (Väyrynen 2017: 171).[2] In contrast, the naturalist argues that the class of moral properties is continuous with the class of natural properties given that the former is a subset of the latter (Sturgeon 2003). Hence, the first desideratum is the following:

> **(Metaphysical)** The distinction between natural and non-natural properties should be cashed out in a metaphysically interesting way (and not, for example, in a merely linguistic way).

---

[2] From now on I will refer to naturalistic moral realists as naturalists (ditto for non-naturalist moral realists). Arguably, supernaturalism would also be a distinct kind of non-naturalism (Adams 1999). I will bracket this complication (although see fn. 25).

Non-natural properties should be understood as metaphysically different from natural properties in a way which justifies calling them discontinuous with the latter. After all, non-naturalists often claim that moral properties are *sui generis,* radically different from other kinds of properties.

Arguably this result is not delivered by purely methodological accounts of moral naturalness. Cuneo (2007) argues in favor of a methodological taxonomy drawing influence from similar views outside metaethics (cf. Ney 2008). According to such a taxonomy one of the distinctive features of naturalism is its commitment to the so-called *external accommodation project.* This project aims to integrate or subsume ethical theory under our best scientific theories (Cuneo 2007: 854). On the other hand, non-naturalists are supposed to only be concerned with the accommodation of ordinary moral practice (2007: 854). But from the mere fact that two properties figure in projects with different methodological commitments it does not follow that they are discontinuous in any deep metaphysical sense. The internal accommodation project could still be true even if it turned out that there are no non-natural properties. I take this result to be obviously problematic.[3]

If some version of moral realism is true then the truth of that proposition should reflect a substantive truth about the fabric of reality, not a simply a fact about our conceptual apparatus or our methodology. These remarks are also supported by Moore himself who explicitly described non-naturalism as a substantive metaphysical view (Moore 1903: chapter IV; cf. Gibbard 2003: 31).

---

[3] Enoch (2011: 362) believes that even though Cuneo fails to meet (Metaphysical) he is still, sociologically speaking, "on to something". But I do not think that even this weaker claim is true. For one, some naturalists (if not all of them) seem eager to accommodate moral practice and phenomenology (Railton 1986: 171). On the other hand, non-naturalists seem to accept that our moral beliefs (up to some extent at least) should cohere with widely accepted scientific facts.

This indicates another desideratum. The taxonomy should take into serious consideration the current metaethical debate:

> **(Revision)** The taxonomy of natural and non-natural properties should accommodate, as much as possible, the views of paradigmatic naturalists and non-naturalists.

(Revision) is motivated by the plausible claim that the metaethical debate between different moral realists so far has not been entirely misplaced or misconstrued by most of its participants. Even if the final verdict about the correct characterization of these properties is still out, I will assume that there is an interesting phenomenon at hand which metaethicists have been tracking it in a non-trivial manner. For example, it would be suspicious if my taxonomy classified Peter Railton as a non-naturalist and G. E. Moore as a naturalist. The way the current metaethical landscape is organized should count as *evidence* for the development of my taxonomy, even if such a taxonomy is ultimately underdetermined by such evidence. This is not to say that revisions are to be avoided at all costs. For example, some claim that Shafer-Landau, a self-proclaimed non-naturalist, can be interpreted as a non-reductive naturalist.[4] Still, when (Revision) acts as a constraint one should expect that the final taxonomy will accommodate to some extent the current landscape of metaethics (for a similar demand see Rosen [2017a: 153]).

There are some taxonomies that plausibly accommodate both desiderata. For example, McPherson defines reductive naturalism as the view that there are some moral properties that are elite (roughly, a special type of properties that ground objective similarity between entities that have them) and that these properties are reduced to uncontroversial natural properties (which means that there

---

[4] For discussion see McPherson (2015: 224) and Väyrynen (2017: 177).

is "an elite (non-identity) function of only natural properties" which backs their "real definition" [2015: 17]). On the contrary, non-naturalism is the view that some moral properties are elite but not a part of the class of natural properties.

Even if McPherson's taxonomy correctly captures what is really at stake in the debate between different moral realists it does so at a great cost: philosophical parochiality. I understand the parochiality charge in the sense of Sturgeon (2009) (cf. Väyrynen 2009: sec. 6). A theory is parochial insofar as it can be accepted only by those who share some specific (and often controversial) background assumption. McPherson's background assumptions are clear. They include a commitment to neo-Lewisian metaphysics (Lewis 1983) as well as a theory of reduction in terms of real definition (Rosen 2015). This is not to say that these assumptions are not viable (or even plausible) on their own. The problem is that it seems plausible that someone could be a moral naturalist independently of her specific beliefs about metaphysics and reduction. In this sense, an ecumenical taxonomy which does not rely on any such assumptions will be clearly superior:

> **(Parochiality)** The taxonomy of natural and non-natural properties should avoid relying on parochial assumptions.

Someone might claim that such a constraint is overly demanding. After all, Sturgeon's own conclusion about moral supervenience is that there is simply no way of articulating such a view in a completely non-parochial way. This result seems to support the idea that every taxonomy will necessarily rely on some assumptions if it has any aspirations of even taking off the ground. However, it seems implicit in Sturgeon's discussion that parochiality is a *graded* phenomenon. Low degrees of parochiality are acceptable (for example, everyone agrees that a theory should be logically coherent). Higher levels of parochiality on the other hand occur when

these assumptions start to get more controversial. To my mind, it is obvious that neo-Humean metaphysics marks a high degree of parochiality.

The taxonomy that I am looking for should highlight what is at stake in the debate between moral realists from a neutral standpoint. For instance, if someone aspires to be a non-naturalist my taxonomy should help her make more salient the ways in which she can defend or vindicate her thesis (or, in negative terms, the ways in which she can attack her rivals). I will now turn to the evaluation of some suggested taxonomies in the literature and argue that they fail to meet the desiderata specified above.

3. A general schema

As already noted, the most natural way of developing a taxonomy of different versions of moral realism is to rely on an independent distinction between natural and non-natural properties. An initial thought would be to define natural properties as the properties that are examined by some naturalistically acceptable field of inquiry. Physics is a plausible contender:

> **(Physical)** F is a natural property iff F is a property that is tracked by physics.

But such a result is too strong. If what is necessary for moral properties to be identified as natural is for them to be tracked by Physics then this renders naturalism trivially false. A way to remedy this would be to introduce a second, recursive, clause:

> **(Physical)** F is a natural property iff
>
> (1) F is a property that is tracked by physics, or,

(2) F is (or can be) appropriately related to such properties.

Clause (2) makes (Physical) more viable. Even though moral properties are (obviously) not tracked by Physics perhaps they can be related to properties that are. But problems remain. Naturalism and physicalism do not necessarily go together. It would be nice, for example, if naturalism could still be true in non-physical worlds.[5] Of course, this is by no means a knock-down argument against (Physical) since it might, given a plausible construal of clause (2), be the only way to meet the desiderata set in section 2. Still, I will briefly review some alternatives.

One option involves broadening the scope of (1) so that it doesn't include merely physics but other paradigmatic scientific fields as well:

(Empirical) F is a natural property iff

(1) F is tracked by the empirical sciences, or,

(2) F is (or can be) appropriately related to such properties.

Even though this proposal fares better than (Physical) it is still problematic for it violates (Metaphysical). Natural properties are defined indirectly in virtue of their second-order property to be tracked by the appropriate sciences. Even if we are lucky enough and the empirical sciences do, in fact, successfully track the relevant natural properties, (Empirical) would be (at best) a heuristic. This is evident when we realize that we need an independent account of what deems a science as empirical, and thus apt for tracking natural properties. But if this is so, then why not skip directly to *that* characterization instead of relying on (Empirical)? (Copp 2003: 182; Väyrynen 2009: 288)

---

[5] For instance, see the ghost-world example in Sinhababu (2018: 35) (cf. Copp 2003: 180).

Another option would be to adopt an epistemological and not a merely disciplinary version of (1) (Copp 2003: 189):

(Epistemological) F is a natural property iff

(1) (a) it is possible for F to be instantiated and (b) there are propositions about the instantiation of F that are both synthetic and possibly true, and, (c) no such proposition is strongly a priori. Or,

(2) F is (or can be) appropriately related to such properties.[6]

A proposition P is strongly a priori iff (a) P is known a priori, and, (b) P is empirically indefeasible, which means that from the point of view of an ideal thinker, there can be no empirical defeaters against it. First, we should note that there are paradigmatic non-naturalists who deny that they are committed to the view that moral propositions are known strongly a priori (Cuneo 2007: 852); contra (Revision). Secondly, supernatural properties like the properties of miracles can plausibly meet (Epistemological) and would be falsely characterized as natural. Simply imagine the causal influence of angels: even though we could have empirical access to them, "being an angel" is a paradigmatic non-natural property.[7] Finally, (Epistemological) is compatible with mutually exclusive metaphysical bases; contra

---

[6] This option (or some of its variants) is quote popular (Moore 1903: 91; Smith 1994: 212; Copp 2003; Rubin 2015: 389-90; Väyrynen 2009: 289; 2017: 171).

[7] I should also note that this is a reason why purely causal accounts (e.g. Lewis 1983; Sturgeon 2003: 538) fail as well. Another reason is that they violate (Revision) since there are non-naturalists who deny that moral properties are causally inert (Wedgwood 2007; Oddie 2018) and there are naturalists who accept the possibility of causally inert natural properties (like quantum states or certain biological properties). Similar remarks seem to apply to proposals that identify the natural with the spatiotemporal (Shafer-Landau 2003: 59; Sturgeon 2003: 38-39).

(Metaphysical). Non-natural properties could still be sui-generis even though they are discovered in the manner of (Epistemological).

Unfortunately, not much progress has been made since (Physical). It seems that despite its problems it seems to be the option with the least number of problems. However, not much hangs upon this. If someone is not happy with (Physical) then they are welcomed to use their preferred "base-level" property sacrificing some parochiality points accordingly.[8] Still, I will stick to the current formulation of (Physical) for present purposes. The following section will examine the content of clause (2).

4. Grounding

Grounding is a non-causal determination relation that has recently gained a lot of traction in the literature. If a fact P grounds another fact Q, then Q exists, in some sense, in virtue of P. As a first pass, I will assume a "standard" theory of grounding according to which grounding is a primitive, asymmetric, transitive, and irreflexive, relation (Schaffer 2009; Rosen 2010; Audi 2012).

Grounding has two important features that make it a plausible candidate for clause (2). First, grounding seems like the ideal relation to explain moral supervenience. Secondly, in doing so, grounding accommodates the naturalistic intuition that moral facts obtain *because* of some non-moral natural facts. There

---

[8] Some other proposals include "descriptive" (Jackson 1998), "non-normative" (Rosen 2017a), and "non-normatively involving" (Ridge 2019). I do not have the space to fully evaluate these options but it seems that they beg the question against naturalists (like Sturgeon) who claim that moral properties could be normative (or captured only by normative concepts) but still natural in virtue of some other feature (see also Väyrynen [2021]).

seems to be a consensus among metaethicists that the set of moral properties supervenes on some set of non-moral properties (in our case, physical properties).[9] Q-properties supervene on P-properties iff (crudely) it is not possible for there to be a difference in P-properties without some corresponding difference in Q-properties. In its strongest formulation this modal covariation holds even when these two sets of properties are ascribed to entities which inhabit different possible worlds.

Even though strong supervenience is widely accepted, not everyone agrees about its implications. For example, Jackson (2017: sec. 4) argues that this sort of necessary covariation provides a strong reason to take P-properties to be *identical* to Q-properties. On the other hand, some non-naturalists accept moral supervenience but resist such identification (e.g. Shafer-Landau 2003: 105; Enoch 2011: 141). It seems that mere supervenience cannot distinguish between naturalism and non-naturalism. This is where grounding comes into play.[10]

(Grounding) F is a natural property iff

(1) F is a physical property, or,

(2) F is fully grounded by such properties.

Given (Grounding) the key difference between natural and non-natural properties is that the latter are not fully grounded in physical properties although they may supervene upon them. A grounding relation between two entities is supposed to

---

[9] There are exceptions. Some naturalists think that moral supervenience is in some sense trivial (see Sturgeon 2009), whereas some non-naturalists deny that it yields in every possible world (Fine 2002; Rosen 2020).

[10] Here's a list of philosophers that seem to be sympathetic to a grounding-based definition of naturalism: Rosen (2010); Väyrynen (2013); Dasgupta (2014); Maguire (2015); Atiq (2017: 10-3); Jackson (2017: 196-7); Bohn (2018); Wygoda Cohen (2018: 17); Hattiangadi (2018: 606); Berker (2018: 28).

establish a tighter metaphysical connection than mere modal correlation. If moral properties are grounded by natural properties then this indicates an explanatory relation. This is a result of the fact that standard grounding is asymmetric (in contrast to supervenience which is non-symmetric).[11]

I want to resist adopting (Grounding) for at least two reasons. First, grounding itself is very controversial and a highly contested relation. Some argue that it incoherent or that it does not exist at all (Daly 2012). Other claim that even if it exists it is not evident that it can do any useful metaphysical work (Wilson 2014; Koslicki 2015). Finally, even among grounding proponents there is great disagreement about many of its features with no clear prospects for a consensus. It should be possible to be a moral naturalist independently of one's commitment to grounding.[12]

Secondly, and relatedly, (Grounding) has an awkward implication for naturalism. If naturalism is the view that there are only natural facts then, arguably, this should also include facts about grounding itself. But under the present

---

[11] Many of the features I am appealing to are by no means uncontroversial. Still, I will make some assumptions to get the discussion going. For a challenge against the features proposed by what I call the "standard" view of grounding see Rodriguez-Pereyra (2015). I assume for now that properties can figure into grounding relations (Schaffer 2009). I also assume a distinction between full and partial grounding (more on this later). Finally, I adopt a *separatist* view about the relation between grounding and non-causal explanation (roughly the view that grounding *backs* non-causal explanations). To compare, so-called *unionists* take grounding to be *identical* to metaphysical explanation and not simply a relation that backs it (Raven 2015).

[12] Berker (2018) has recently suggested that it is hard to make sense of moral discourse without appealing to grounding since moral claims do not indicate mere modal covariations between moral and non-moral properties. Rather, the goodness of some action holds *in virtue of* some non-moral feature of that action. This latter claim is true, but it doesn't follow that the *only* way to capture the hyperintesionality of these contexts is by appealing to a grounding-based ideology.

formulation, for facts about grounding to come out as natural it must either be the case that they are physical facts, or that they are grounded upon facts that are. The first option seems unlikely, and the second option generates a regress. Of course, it might be the case that they are formidable ways to rescue the naturalness of grounding.[13] But it would be weird if it turned out that moral naturalists have to deal with obscure issues about the metaphysics of grounding in order to defend their view. Both points go against (Parochiality).[14]

A possible way forward would be to give up on using grounding altogether and use less controversial tools. Wilson (2014) famously argued that there is a family of relations that makes grounding obsolete. These are relations like composition, constitution, realization, set-formation etc., which are less controversial and share many of the features of grounding like asymmetricity and being explanatory. Also, their non-primitive nature doesn't generate problems like the one I develop in the previous paragraph. Take realization for example. According to one influential formulation realization can be further analyzed in causal terms: φ realizes ψ when the causal powers of the latter are a proper subset of

---

[13] For example, the naturalist could argue (in some way) that grounding-facts are physical after all, or she could deny that grounding facts are apt for grounding (Dasgupta 2014), or that they are grounded in way that does not generate a vicious regress (e.g. by appealing to zero-grounding) (Litland 2017). But these moves involve a commitment to *further* parochial assumptions. For a similar objection along these lines see Melnyk (2016).

[14] There are further issues which I will not defend for reasons of space. Still, I will briefly mention one reason why (Grounding) also goes against (Revision). The standard view takes grounding to always obtain with metaphysical necessity (cf. Leuenberger 2014; Skiles 2015). But some non-naturalists reject this (e.g. Rosen 2020). If grounding is used in the definition of non-naturalism (in one way or another), then this would exclude such views (which, even though they are controversial, deserve serious consideration).

the causal powers of the former (Wilson 2011). I will follow Wilson and call these relations "small-g" relations.

The small-g based proposal would presumably be that a property is natural iff it is either a physical property or a property that is *small-g* related to physical properties. A problem that immediately arises concerns the *identification* of the small-g relation that is best fitted for the definition of moral naturalness. Some relations (like set-formation) are clearly not well suited for that role. So, at best, a small-g based proposal is incomplete at this point. In a sense, this proposal trades non-parochiality for lack of informativeness.

This problem can be solved by further (hopefully, non-parochial) metaphysical assumptions. Assume, for example, that there are good reasons to think that functional realization is the right relation to figure in the definition of moral naturalness. But even if that move is granted, there is *yet another* sense in which the small-g based account is incomplete. The obvious way to distinguish between naturalism and non-naturalism would be to appeal to a distinction between full and partial realization.[15] A property F would be a natural property given that it is fully realized in terms of physical properties, whereas non-naturalism would be stated via a realization claim involving partial physical realizers. But what is this *further* thing (other than the physical realizers) that makes a non-natural property obtain?[16] In the next section I propose an answer to this question.

---

[15] Note that this distinction is similar but not the same as the standard distinction between core and total realization (Shoemaker 2007) (in the sense that not every partial realizer is necessarily a core realizer of a given phenomenon).

[16] Note that it cannot be said that a property F is non-natural when F is partially grounded in physical properties and *nothing else*. If so, it would be a mystery how F even obtains in the first place (cf. Enoch 2019: 4-5 fn. 10). To compare, if the only influence exerted upon an event E is an

5. Nomic Naturalness

According to the small-g based account I sketched in the previous section, a non-natural property F is only partially realized by physical properties. If so, what other component would be needed in order to instantiate F? Perhaps, as Enoch (2019) has recently suggested, that component is a *moral law*.[17] The resulting picture would be the following:

> **(1)** A property F is non-natural when F is partially realized by physical properties, together with a moral law connecting these two kinds of properties.[18]

The non-naturalist can claim that the involvement of a moral law in (1) is the hallmark of the non-natural. The fact a moral law is required to connect a moral property with a physical property arguably showcases that the former is *just too different* from the latter (Enoch 2011: 82; Enoch 2019) (cf. Rosen 2017).[19]

---

insufficient cause, then it would be bizarre to claim that E obtains nonetheless. Neither is this question begging against the moral particularist (more on this later).

[17] As it will become apparent, I take moral principles to be an instance of moral lawhood. Also, by 'moral law' I mean '*pure* moral law' (in contrast to *mixed,* or derivative, ones) unless stated otherwise.

[18] There is a technical complication about how exactly a thesis like (1) is supposed to be formulated. On the grounding literature there are at least two ways in which principles or laws can play a role. Several philosophers (e.g. Maguire 2015; Leary 2017; Rosen 2017a; 2017b; 2017c; Enoch 2019) introduce laws as parts of the grounds. Another option is to introduce laws as grounds of the grounding-fact taken altogether (Schaffer 2017) (Skow 2016; Bader 2017; Berker 2018; cf. Maguire 2015: 198 fn. 22). I don't think anything significant hangs on this and I will run my discussion according to the first reading.

[19] This is not to say that this is the only way to cash out the "just too-different" intuition. For an excellent discussion see Paakunainen (2017).

However, there is no apparent reason why a *naturalist* cannot accept (1) as well. For example, the naturalist could claim that their thesis also involves a moral law in the same way special-science claims involve bridge-laws. Facts about certain states of pain obtain in virtue of facts about certain neurological states insofar as there is a general law connecting these two states. So, the *mere involvement* of a moral law in the instantiation of a property is not enough for that property to be characterized as non-natural. Instead, there must be a *principled difference* between natural and non-natural moral laws for them to do the required work. I will argue that there is such a difference. Specifically, natural laws have a feature which is not shared by non-natural laws: they figure in a particular kind of *network* of laws. That network is the one that includes the laws of physics. I will call this the *nomic nexus* (NN).[20]

A law figures in NN when it either is a physical law or it is dependent upon them. This is not to say that the only way to make sense of NN is to accept that the laws of (say) psychology necessarily *reduce* to the laws of physics. Accepting the existence of NN really comes down to the more modest claim that natural laws are not "nomological danglers" (Smart 1959): laws figuring in NN have interesting metaphysical connections between them and they do not "float free" (i.e. they are not brute). What would these connections be? It is hard to say without committing to a particular theory concerning the metaphysics of lawhood. For example, Lewis famously suggested that laws are a certain type of regularity, and the relation between laws is that of *subsumption* (Psillos 2014). According to this framework, the regularity describing Kepler's principles of planetary motion are *subsumed* by

---

[20] I borrow the term from Hempel (1965: 488) and Railton (1986: 184). Feigl's (1956) "nomological network" is a similar idea. As it will become apparent, my way of understanding the nomic nexus is metaphysically weightier.

Newton's law of gravitation. Or, perhaps, laws are relations between universals as per Armstrong (1983) and the relation between laws is the relation of *instantiation*: higher-level laws are expressed by relations between so-called "complex" universals which are, in turn, instantiated in terms of more basic universals.[21] Things get complicated depending on which theory of lawhood one adopts. But independently of which theory one chooses, it will still be the case that if a law L figures in NN there will be an interesting metaphysical connection between L and the laws of physics. *That* is the distinctive feature of non-physical natural laws.

A way to further illustrate this is by considering how typical lawlike propositions appear in the context of scientific *theories*. Moral laws are the moral analogues of bridge-laws figuring in the special sciences: they connect entities from different levels of reality (roughly, the moral with the non-moral domain). But there is a consensus in the relevant literature that bridge-laws are intimately connected with other laws figuring in other theories. To see this, consider that propositions about bridge-principles sometimes act as premises and other times as conclusions. Take, for example, the Boyle-Charles law (a law connecting pressure, volume, and temperature). It can be used as a premise to (roughly) derive the truth of the kinetic theory of gases (alongside other auxiliary assumptions, for example that molecules are point particles). But it can also be used as a conclusion (using the principles involved in the kinetic theory as a premise instead) (Ager et al. 1974: 121).[22]

---

[21] For the intricacies of exactly how this is supposed to work see Fisher (2018).

[22] This feature of bridge-laws is well-established in the literature (Nagel 1961: 354; Sklar 1967: 118; Ager et al. 1974; Causey 1976: 334, 337; Horgan 1978: 231; Blachowicz 1997: 451; Boyd 2003: 519; Dizadji-Bahmani et al. 2010: 408; Bangu 2011: 116; Shapiro 2018: 5-7). I also take it that a realist stance about the structure of theories is dialectically permissible at this point given that my point of focus concerns the notion of natural in the context of moral realism.

Moral naturalists should hold that moral laws are structurally and metaphysically analogous to bridge-laws like the Boyle-Charles law. Consider the basic moral law posited by Boyd (1988: 329):

> **(Consequentialism)** Something (an act, etc.) is good iff, and because, it brings about the satisfaction of some important human needs (specified appropriately).[23]

If (Consequentialism) is a natural fact then it is dependent on how the rest of the NN is organized. This is not say that (Consequentialism) cannot obtain in worlds with radically different laws. It might turn out that (Consequentialism) is highly invariant across different possible nexuses. Or it may not be invariant at all and it might be dependent on whether some very specific physical law obtains. This is an empirical issue and a question of first-order normative ethics. The important point is that (Consequentialism) does not float-free from paradigmatic natural laws.[24]

To compare, non-natural laws would be laws holding in their *own* network in a way that has nothing to do with the organization of NN. This is the kind of idea that paradigmatic non-naturalists seem to want to capture.[25] Non-naturalists

---

[23] The exact logical form of moral law is a contentious issue in the literature (Berker 2018: sec. 2) but orthogonal to point I am trying to illustrate.

[24] Sturgeon (1985: 241) argues that our posited moral laws will be vindicated by the relevant evidence in the same way reductive theories of colour are vindicated by facts about optical theory and psychology. For similar remarks see (Brink 2001: 169; Boyd 2003: 525).

[25] What about *supernatural* laws? A paradigmatic supernatural law involves the *negation* of a natural law in a way that produces a miraculous event (Luck 2011). In this sense, the distinctive feature of supernatural laws is that they *violate* natural laws. Another view about the manifestation of miracles is due to Augustine who thought that miracles were the result of the acceleration of natural laws. The direct intervention view is usually ascribed to Aquinas and Hume (Larmer 1988). There is also a complication with deistic views such as the view of Descartes according to which natural laws hold in

understand the metaphysical status of moral laws as something separate and insulated (so to speak) from the natural domain. Some take these principles to be analogous to mathematical facts (e.g. Ross [1930: 29-30]), ante rem universals (e.g. Cohen [2003: 291]; Skarsaune [2015]), and other types of facts that are typically understood as holding independently of the natural domain.[26]

To illustrate, consider Rosen's (2020: 219) recent suggestion that non-natural moral laws are *fact-independent*. Something is fact-independent, in his sense, when it would still have been the case independently of the non-normative facts (2020: 219 fn. 21). Under my proposal, Rosen's notion of fact-independence would be qualified in the following way: Fact-independence should be understood as independence *from NN* (i.e. the network of laws involving the laws of physics). This qualification is important since, for all that I have said, the laws figuring in NN *may or may not* be non-normative. In this sense, even essentially normative laws (whether moral or not) would be natural if they figure in NN. Still, the important non-naturalist intuition is captured: if a moral law does not figure in NN, then there is a clear sense in which that principle is metaphysically insulated.

---

virtue of God's immutability and are metaphysically necessary. This view implies that the laws of nature are, in a sense, supernatural since they are generated by God herself. This is not the place to settle the issue but there is a sense in which, in that case, God would simply be a node in NN. This would be a version of moral naturalism under my definition. But this is not necessarily a problematic result. I take it that what is interesting about a putative account of supernaturalist moral realism is that moral goodness obtains *regardless* of the organization of the natural domain. For general discussion see Ott (2009: chapter 7).

[26] This characteristic of (pure) non-natural moral laws is recognized by paradigmatic and self-proclaimed non-naturalists across the board (e.g. Moore [1903: 193]; Nagel [1986: 138]; Hampton [1998: 105]; Shafer-Landau [2003: 46]; Oddie [2005: 210]; Wielenberg [2009: 32]; Fitzpatrick [2011: 27]; Enoch [2011: 146; 2019]; Scanlon [2014: 121; 2017]; Rosen [2017b: 287]). This goes in favour of (Revision).

6. Fleshing out the account

According to the general schema I proposed in section 2, a property F is natural when it is either a physical property or it is appropriately related to physical properties. I examined an explication of the second clause of that definition by appealing to standard grounding and small-g non-causal determination. I concluded that standard grounding is too philosophically parochial whereas appeals to small-g relations are uninformative. There should be a principled way of identifying the small-g relation that connects the moral and the natural, as well as the nature of that *further* partial ground that makes a given moral property non-natural. Appealing to NN avoids both problems.

> **(Nomic)** F is a natural property iff F falls within the scope of a law figuring in NN (i.e. the network of laws involving the laws of physics).

(Nomic) defines moral naturalness directly in terms of subsumption under a natural law (which, in turn, is defined in terms of figuring in a network of laws involving the laws of physics). In this sense, there is no *need* to appeal to either standard grounding or some small-g relation. To be clear, it could be that the logical *form* of moral laws involves a grounding operator or something of the like (cf. Berker 2018: sec. 1; Emery 2019). But this is orthogonal to the issue at hand. What makes a moral law natural is whether it figures in NN, not its logical form.

I take the idea behind NN to be non-parochial. It is common ground between every major theory of natural lawhood that every natural law can be traced back to the same *source*. Lewisian metaphysics take laws to be ultimately based at the Humean mosaic (Lewis 1994). Similarly, necessitarians argue that laws do not operate in isolation (cf. Wilson 2005: sec. 2.2.2.). This could be so because they are

generated by the same primitive source (Maudlin 2007: 12-15), or they correspond to some common constellation of universals (Fisher 2018), or they are a part of a single set (Lange 2009: 37-40).

Nor does the idea behind NN involve a commitment to ethical reductionism. The fact that a moral law figures in NN (thus making that principle, and the properties that fall within its scope, natural) does not necessarily entail that that principle *reduces* to the laws of physics. For a law to figure in NN, as mentioned, there simply needs to be an interesting metaphysical connection between that law and the rest of the laws in NN. Presumably, ethical reductionism would be true if the logical form of moral law involves the *identity* of moral properties with their non-moral base. Still, it is plausible that moral properties are multiply realizable which would entail that the relation that figures in that moral law to be functional realization. But this is still an interesting metaphysical connection which makes that moral law naturalistically acceptable.[27]

Another worry might concern the fact that the notion of a moral *law* is built into the definition of moral naturalness. But shouldn't moral particularism also be an option for the naturalist? I agree that moral naturalists can be particularists. But I disagree with the idea that particularists reject the existence of moral laws. What they should reject is the existence of moral *principles.* So far, I have been taking a moral law to simply be a general fact involving the connection between a moral and

---

[27] Note that, as mentioned in the previous section, an interesting metaphysical connection is that connection between laws which is posited by one's background theory of lawhood. For all I have said, a moral law could strongly *emerge* from laws figuring in NN (for something along these lines see Stringer [2018]). Even in such an extreme case, there is still *a* relation at play between a moral law and the rest of NN.

a non-moral entity (in this case, a physical property).[28] But there is no reason to think that that law will have the right features in order to be characterized as a moral principle. Moral principles (construed as standards and not mere guides) are supposed to be explanatory and counterfactually robust in a way that indicates that they are metaphysical *uniform* across a variety of different circumstances. A typical particularist response would be to deny that there are such principles given the phenomenon of reasons holism: roughly, something might be reason to φ in one particular set of circumstances, even though it may not be a reason at all in some different set of circumstances (this is supposed to extrapolate to other normative entities like moral goodness) (Dancy 1993: 60). But even so, it would still be the case that there are general moral facts (i.e. moral *laws*, in my sense) which are highly disjunctive: an act would be morally good under a multitude of non-moral circumstances which are not unified in any important sense.[29]

Finally, consider the epistemology of moral naturalness. How do we *know* whether a moral law figures in NN? Someone might worry that evaluating whether a law P figures in NN would require evaluating counternomics of the form 'Had not-P been the case, the organization of NN would be different in such-and-such a way'. Presumably, this would be a problem given the controversial nature of these counterfactuals: counterfactuals involving particular facts are usually evaluated by holding fixed the relevant laws of nature. But this would not apply to counternomics of the above form.

---

[28] For a less minimal definition which includes these features *into* the definition of moral lawhood see Rosen (2017c: 146).

[29] I take all this to be uncontroversial since it is widely accepted that even moral particularists should accept moral supervenience. Insofar as such a supervenience thesis is not brute, moral particuralists would accept a moral law of the sort I have just sketched. For general discussion see Väyrynen (2018).

In response, I want to suggest that one can stay agnostic about the way these counternomics are evaluated.[30] Instead, the question of whether a moral law figures in NN ultimately collapses to the question of whether there is an interesting metaphysical connection holding between that principle and some component of NN. For example, if according to one's background theory of lawhood the relation holding between laws is that of grounding (cf. Emery 2019) then the epistemology of moral naturalness is nothing but an instance of the epistemology of grounding. But then (Nomic) does not face a *special* epistemological problem. If P is natural then there will be a metaphysical relation holding between P and a component of NN. But how one tracks that relation concerns one's background theory of *that* relation, not (Nomic) in particular. Under these considerations, I conclude that (Nomic) does not face any special problems in terms of its assumptions.

7. Taking stock

In this paper I developed and defended a novel account of moral naturalness: a property F is natural iff it falls within the scope of a natural law. In turn, I defined the naturalness of a law in terms of whether that law figures in a nomic nexus involving the laws of physics (what I called NN). I argued that

---

[30] It could even be argued that these counternomics are not as controversial as they might appear. Tan (2017) has recently suggested that such counternomics are routinely used in the sciences and they can be evaluated using mundane experimental observations. Analogously, first-order moral discourse arguably appeals to counterfactuals involving moral discourse. For example, a deontologist might argue that if utilitarianism is true, then some intuitively morally abhorrent moral principle would be true. In a sense, such a deontologist takes the counterfactual 'Had utilitarianism been the case, such-and-such moral principle would be true' to be non-trivial. This seems to suggest that possible worlds which are governed by different (pure) moral principles are conceivable (for discussion see Hattiangadi [2018]).

(Nomic) is superior to its rivals in terms of the desiderata I introduced in section 2. Paradigmatic non-naturalists accept that non-natural laws are metaphysically insulated. Also, (Nomic) does not really on any parochial assumptions (e.g. assuming a grounding-like relation) while identifying a deep metaphysical difference between the natural and the non-natural domain. Finally, I further defended (Nomic) by arguing that it doesn't beg the question against the moral particularist nor does it necessarily entail ethical reductionism.

References

Adams, Robert (1999). *Finite and Infinite Goods: A Framework for Ethics*. Oxford University Press USA.

Ager, Tryg A. & Aronson, Jerrold L. & Weingard, Robert (1974). Are bridge laws really necessary? *Noûs* 8 (2):119-134.

Armstrong, D. M. (1983). *What is a Law of Nature?*. Cambridge University Press.

Atiq, Emad H. (2018). On Ground as a Guide to Realism. *Ratio* 31 (2):165-178.

Audi, Paul (2012). Grounding: Toward a Theory of the I n-Virtue-Of Relation. *Journal of Philosophy* 109 (12):685-711.

Bader, Ralf M. (2017). The grounding argument against non-reductive moral realism. *Oxford Studies in Metaethics* 12.

Bangu, Sorin (2011). On the Role of Bridge Laws in Intertheoretic Relations. *Philosophy of Science* 78 (5):1108-1119.

Berker, Selim (2018). The Explanatory Ambitions of Moral Principles. *Noûs* 53 (4):904-936.

Blachowicz, James (1997). Reciprocal justification in science and moral theory. *Synthese* 110 (3):447-468.

Bohn, Einar (2018). Normativity all the way down: from normative realism to pannormism. *Synthese* 195 (9):4107-4124.

Boyd, Richard (1988). How to be a Moral Realist. In G. Sayre-McCord (ed.), *Essays on Moral Realism.* Cornell University Press. pp. 181-228.

Boyd, Richard (2003). Finite beings, finite goods: The semantics, metaphysics and ethics of naturalist consequentialism, part I. *Philosophy and Phenomenological Research* 66 (3):505–553.

Brink, David O. (2001). Realism, Naturalism, and Moral Semantics. *Social Philosophy and Policy* 18 (2):154.

Causey, Robert L. (1976). Identities and reduction: A reply. *Noûs* 10 (3):333-337.

Cohen, G. A. (2003). Facts and Principles. *Philosophy and Public Affairs* 31 (3):211-245.

Copp, David (2003). Why naturalism? *Ethical Theory and Moral Practice* 6 (2):179-200.

Cuneo, Terence (2007). Recent Faces of Moral Nonnaturalism. *Philosophy Compass* 2 (6):850-879.

Daly, Chris (2012). Scepticism about Grounding. In Fabrice Correia & Benjamin Schnieder (eds.), *Metaphysical Grounding: Understanding the Structure of Reality.* Cambridge University Press. pp. 81.

Dancy, Jonathan (1993). *Moral Reasons.* Blackwell.

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

Dizadji-Bahmani, Foad & Frigg, Roman & Hartmann, Stephan (2010). Who's Afraid of Nagelian Reduction? *Erkenntnis* 73 (3):393-412.

Emery, Nina (2019). Laws and their instances. *Philosophical Studies* 176 (6):1535-1561.

Enoch, David (2011). *Taking Morality Seriously: A Defense of Robust Realism.* Oxford University Press UK.

Enoch, David (2019). How Principles Ground. *Oxford Studies in Metaethics* 14.

Feigl, Herbert (1956). Some major issues and developments in the philosophy of science of logical empiricism. In Herbert Feigl & Michael Scriven (eds.), *Minnesota Studies in the Philosophy of Science.* , Vol. pp. 3-37.

Fine, Kit (2002). Varieties of Necessity. In Tamar Szabo Gendler & John Hawthorne (eds.), *Conceivability and Possibility.* Oxford Up. pp. 253-281.

Fisher, A. R. J. (2018). Structural universals. *Philosophy Compass* 13 (10):e12518.

FitzPatrick, William J. (2011). Ethical non-naturalism and normative properties. In Michael Brady (ed.), *New Waves in Metaethics.* Palgrave-Macmillan.

Frank, Jackson (2017). In Defence of Reductionism in Ethics. In Peter Singer (ed.), *Does Anything Really Matter?: Essays on Parfit on Objectivity.* Oxford: Oxford University Press. pp. 195-211.

Gibbard, Allan (2003). *Thinking How to Live.* Harvard University Press.

Hampton, Jean E. (1998). *The Authority of Reason.* Cambridge University Press.

Hattiangadi, Anandi (2018). Moral supervenience. *Canadian Journal of Philosophy* 48 (3-4):592-615.

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

Horgan, Terence E. (1978). Supervenient bridge laws. *Philosophy of Science* 45 (2):227-249.

Jackson, Frank (1998). *From Metaphysics to Ethics: A Defence of Conceptual Analysis.* Oxford University Press.

Koslicki, Kathrin (2015). The Coarse-Grainedness of Grounding. *Oxford Studies in Metaphysics* 9:306-344.

Lange, Marc (2009). *Laws and Lawmakers Science, Metaphysics, and the Laws of Nature.* Oxford University Press.

Larmer, Robert (1988). Against 'against miracles'. *Sophia* 27 (1):20 - 25.

Leary, Stephanie (2017). Non-naturalism and Normative Necessities. *Oxford Studies in Metaethics* 12.

Leuenberger, Stephan (2014). Grounding and Necessity. *Inquiry: An Interdisciplinary Journal of Philosophy* 57 (2):151-174.

Lewis, David (1983). New work for a theory of universals. *Australasian Journal of Philosophy* 61 (4):343-377.

Lewis, David (1994). Humean Supervenience Debugged. *Mind* 103 (412):473--490.

Litland, Jon (2017). Grounding Grounding. *Oxford Studies in Metaphysics* 10.

Maguire, Barry (2015). Grounding the Autonomy of Ethics. *Oxford Studies in Metaethics* 10.

Maudlin, Tim (2007). *The Metaphysics Within Physics*. Oxford University Press.

McPherson, Tristram (2012). Ethical Non-Naturalism and the Metaphysics of Supervenience. In *Oxford Studies in Metaethics Vol 7*. pp. 205.

McPherson, Tristram (2015). What is at Stake in Debates among Normative Realists? *Noûs* 49 (1):123-146.

Melnyk, Andrew (2016). Grounding and the Formulation of Physicalism. In Ken Aizawa & Carl Gillett (eds.), *Scientific Composition and Metaphysical Ground*. Palgrave. pp. 249-269.

Moore, G. E. (1903). *Principia Ethica*. Dover Publications.

Nagel, Ernest (1961). *The Structure of Science: Problems in the Logic of Scientific Explanation*. Harcourt, Brace & World.

Nagel, Thomas (1986). *The View From Nowhere*. Oxford University Press.

Ney, Alyssa (2008). Physicalism as an attitude. *Philosophical Studies* 138 (1):1 - 15.

Oddie, Graham (2005). *Value, Reality, and Desire*. Clarendon Press.

Oddie, Graham (2018). Non-Naturalist Moral Realism, Autonomy and Entanglement. *Topoi* 37 (4):607-620.

Ott, Walter (2009). *Causation and Laws of Nature in Early Modern Philosophy*. Oxford University Press.

Paakkunainen, Hille (2018). The "Just Too Different" Objection to Normative Naturalism. *Philosophy Compass* 13 (2):e12473.

Psillos, Stathis (2014). Regularities, Natural Patterns and Laws of Nature', *Theoria* 79: 9-27.

Railton, Peter (1986). Moral realism. *Philosophical Review* 95 (2):163-207.

Raven, Michael J. (2015). Ground. *Philosophy Compass* 10 (5):322-333.

Ridge, Michael, "Moral Non-Naturalism", *The Stanford Encyclopedia of Philosophy* (Fall 2019 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2019/entries/moral-non-naturalism/>.

Rodriguez-Pereyra, Gonzalo (2015). Grounding is not a strict order. *Journal of the American Philosophical Association* 1 (3):517-534.

Rosen, Gideon (2010). Metaphysical Dependence: Grounding and Reduction. In Bob Hale & Aviv Hoffmann (eds.), *Modality: Metaphysics, Logic, and Epistemology*. Oxford University Press. pp. 109-36.

Rosen, Gideon (2015). Real Definition. *Analytic Philosophy* 56 (3):189-209.

Rosen, Gideon (2017a). Metaphysical Relations in Metaethics. In Tristram McPherson & David Plunkett (eds.), *The Routledge Handbook of Metaethics*. Routledge. pp. 151-169.

Rosen, Gideon (2017b). Ground by Law. *Philosophical Issues* 27 (1):279-301.

Rosen, Gideon (2017c). What is a Moral Law? *Oxford Studies in Metaethics* 12.

Rosen, Gideon (2020). What is normative necessity? In M. Dumitru (Ed.), Metaphysics, Meaning and Modality: Themes from Kit Fine (pp. 205-233). Oxford: Oxford University Press.

Ross, David (1930). *The Right and the Good*. Oxford University Press UK.

Rubin, Michael (2015). Normatively Enriched Moral Meta-Semantics. *Philosophy and Phenomenological Research* 91 (2):386-410.

Scanlon, T. M. (2014). *Being Realistic About Reasons*. Oxford University Press.

Scanlon, T. M. (2017). Normative realism and ontology: reply to Clarke-Doane, Rosen, and Enoch and McPherson. *Canadian Journal of Philosophy* 47 (6):877-897.

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

Schaffer, Jonathan (2017). Laws for Metaphysical Explanation. *Philosophical Issues* 27 (1):302-321.

Shafer-Landau, Russ (2003). *Moral Realism: A Defence*. Oxford University Press.

Shapiro, Lawrence (2018). Reduction redux. *Studies in History and Philosophy of Science Part A* 68:10-19.

Shoemaker, Sydney (2007). *Physical Realization*. Oxford University Press UK.

Sinhababu, Neil (2018). Ethical Reductionism. *Journal of Ethics and Social Philosophy* 13 (1):32-52.

Skarsaune, Knut Olav (2015). How To Be a Moral Platonist. *Oxford Studies in Metaethics* (10).

Skiles, Alexander (2015). Against Grounding Necessitarianism. *Erkenntnis* 80 (4):717-751.

Sklar, Lawrence (1967). Types of inter-theoretic reduction. *British Journal for the Philosophy of Science* 18 (2):109-124.

Skow, Bradford (2016). *Reasons Why*. Oxford University Press UK.

Smart, Jjc (1959). Sensations and brain processes. *Philosophical Review* 68 (April):141-56.

Smith, Michael (1994). *The Moral Problem*. Blackwell.

Stringer, Ryan (2018). Realist Ethical Naturalism for Ethical Non-Naturalists. *Philosophical Studies* 175 (2):339-362.

Sturgeon, Nicholas (1985). Moral Explanations. In James Rachels (ed.), *Ethical Theory 1: The Question of Objectivity*. Oxford University Press.

Sturgeon, Nicholas L. (2003). Moore on ethical naturalism. *Ethics* 113 (3):528-556.

Sturgeon, Nicholas L. (2009). Doubts about the Supervenience of the Evaluative. In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics*. pp. 53-92.

Tan, Peter (2017). Interventions and Counternomic Reasoning. *Philosophy of Science* 84 (5):956-969.

Väyrynen, Pekka (2009). Normative Appeals to the Natural. *Philosophy and Phenomenological Research* 79 (2):279 - 314.

Väyrynen, Pekka (2013). Grounding and Normative Explanation. *Aristotelian Society Supplementary Volume* 87 (1):155-178.

Väyrynen, Pekka (2017). The Supervenience Challenge to Non-Naturalism. In Tristram McPherson & David Plunkett (eds.), *The Routledge Handbook of Metaethics*. Routledge. pp. 170-84.

Väyrynen, Pekka (2018). Reasons and Moral Principles. In Daniel Star (ed.), *The Oxford Handbook of Reasons and Normativity*. pp. 839-61.

Väyrynen, Pekka (2021). Normative Naturalism on Its Own Terms. *Organon F: Medzinárodný Časopis Pre Analytickú Filozofiu*.

Wedgwood, Ralph (2007). *The Nature of Normativity.* Oxford University Press.

Wielenberg, Erik J. (2009). In Defense of Non-Natural, Non-Theistic Moral Realism. *Faith and Philosophy* 26 (1):23-41.

Wilson, Jessica (2005). Supervenience-based formulations of physicalism. *Noûs* 39 (3):426-459.

Wilson, Jessica M. (2011). Non-reductive realization and the powers-based subset strategy. *The Monist (Issue on Powers)* 94 (1):121-154.

Wilson, Jessica M. (2014). No Work for a Theory of Grounding. *Inquiry: An Interdisciplinary Journal of Philosophy* 57 (5-6):535-579.

Wygoda Cohen, Shlomit (2018). Not All Partial Grounds Partly Ground: Some Useful Distinctions in the Theory of Grounding. *Philosophy and Phenomenological Research* 100 (1):75-92.

Chapter 4 | Neo-Humean Moral Contingentism

Abstract: The neo-Humean account of moral principles (HMP) is a theory about the metaphysical status of moral principles. According to HMP, moral principles (in the genetic, non-propositional, sense) are entities that supervene on the Humean mosaic (the set of fundamental, non-modal, physical properties). Specifically, HMP holds that moral principles are nothing over and above the regularities that figure in the Humean mosaic. In this paper, I do two things. First, I present the positive case for HMP. HMP provides two ways of explaining moral principles and a way of explaining moral supervenience which is dialectically superior to its rivals. Secondly, I defend the contingentist aspect of HMP. The properties that figure in the Humean mosaic are governed by a principle of recombination (Lewis 1986). In this sense, moral principles are contingent and differ depending on the distribution of the base properties of each possible world. This makes HMP susceptible to a normative authority problem (Väyrynen 2021), an epistemological problem (Dreirer 2019), and modal miracle problem (Fogal & Risberg 2020). I show that HMP can meet these challenges thus making it superior to other moral contingentist proposals (cf. Rosen 2020; 2021).

1. Introduction

There is regularity in the world. Electric and magnetic fields operate according to Maxwell's equations. Objects are subject to Newton's law of gravitation. Biological inheritance follows the principles of Mendel. Realist moral generalists hold that *moral*

reality is organized in an analogous way. According to their view, there are moral facts, and at least some of these are general in nature.[1] For example, if the principle of utility holds, then any action that maximizes utility is morally right.

I will call general moral facts of this sort, moral principles. We can distinguish between moral principles qua propositions, and moral principles qua truthmakers of such propositions (moral principles in the genetic sense) (e.g., as per Robinson 2011). Analogously, scientific principles (qua propositions) describe or represent the relevant laws of nature, which are understood as worldly entities. I will also take moral principles to be *standards* and not mere guides of moral conduct (McKeever & Ridge 2006; Väyrynen 2018). In this sense, moral principles are supposed to be explanatory of particular moral facts.

In this paper I defend a neo-Humean theory about the metaphysical status of moral principles. Moral generalists typically focus on the logical form of moral principles (in the propositional sense). For example, it is a substantive issue whether moral principles involve separate slots for enablers and disablers, or ceteris paribus clauses. However, the question concerning the proper metaphysics of moral principles has received little attention.

The few views that have been defended in the literature are *substantivist* in nature (as I will call them). They posit moral principles (in the genetic sense) as something *over and above* the particular moral instances that fall within their scope. For example, Robinson (2011) adopts a dispositional theory of moral principles. According to this model, moral principles are "real, irreducibly dispositional properties of individuals that are responsible for and thereby explain the moral properties of (e.g.) agents and actions, moral properties such as my obligation to remain faithful to my wife

---

[1] Moral generalism and moral realism can, of course, come apart (cf. Hare 1963).

and the wrongness of John Lennon's murder." (2011: 291). Similarly, Rosen (2017a) takes moral principles to just be relations between universal-like entities which *govern* particular moral phenomena. My proposal is non-substantivist in nature. According to the neo-Humean view, moral principles (in the genetic sense) are nothing over and above the instances that figure in them. To illustrate, the core principle of consequentialism, so construed, would consist in all and only those acts that are morally right in virtue of having outcomes that maximize value.[2]

My paper has two aims. First, I will argue that the neo-Humean view of moral principles (NHM) has important virtues and should be taken seriously. Secondly, I will defend the *contingentist* aspect of NHM. NHM entails that the fundamental moral principles are metaphysically contingent (in the sense that there are possible worlds that differ in the moral principles they contain). I will argue that the contingentist aspect of NHM is not only defensible, but also a *feature* of the view.

In section 2 I will talk about the neo-Humean program in general and show how NHM is a particular instance of that program. I will also spell out the metaphysics and logical form of neo-Humean moral principles. In section 3 I will motivate NHM by showing how it provides plausible answers to two important questions: (1) What explains moral principles, and (2) What explains moral supervenience. In section 4, I motivate the contingentist aspect of NHM. Finally, in sections 5 and 6, I tackle three

---

[2] The neo-Humean view I have in mind should be distinguished from the Best System Account of lawhood (Lewis 1973: 73-7). According to the Best System Account, laws of nature are those regularities which are expressed by propositions that figure in a deductive system constrained by the desiderata of simplicity and informativeness (cf. Loewer 2020). But it is not *constitutive* of lawhood that they are expressed by propositions that figure in the best system (Lewis is explicit about this vis-à-vis the *ratbag idealist* problem) (Lewis 1994: 479). At best, a best-system analysis provides an *epistemology* of laws (i.e. we have good reasons to take a regularity R to be a law of nature, if R figures in the best system) (cf. Jaag & Loew 2018).

recent objections against contingentist theories of moral principles: Väyrynen's (2021) normative authority problem, Dreier's (2019) epistemological worry, and Fogal & Risberg's (2020) modal miracle worry.

2. Neo-Humean moral principles

According to neo-Humeanism everything that exists at a given world is either a part of the so-called *Humean mosaic*, or it is appropriately related to it. What is the Humean mosaic? There is disagreement among Humeans (Loewer 2007). Still, it will suffice for my purposes to take the Humean mosaic (henceforth, the *base*) to involve non-modal properties that are typically tracked by Physics.

The non-modal feature of these properties is particularly important as this is what makes NHM distinctively *Humean.* According to Hume's dictum, there are no necessary connections in nature. There are many ways to interpret this claim but, according to the typical Lewisian view, the non-modal character of base-properties entails a recombination principle on modal space (Wilson 2015). Since a given base-property is metaphysically independent from any other base-property, it follows that they are recombinable. Roughly put, for every permutation that the base-properties can exhibit, there is a possible world involving that permutation (more on this in the next section).

Not everything that exists is part of the base, such as middle-sized physical goods (chairs, tables, etc.) and entities posited by the special sciences (biological species, economies, minds, etc.). Such macro-properties are not among the base-properties but supervene on them. To be more precise, the property of (say) having a mental state, belongs to an entity that is composed out of entities that bear base-properties. In this sense, the supervenience of macro-properties on the relevant base-properties is not *bare.*

Rather, it is being underwritten by a composition relation (McPherson 2013; Jaag & Loew 2020).[3]

What about the moral domain? Moral properties are plausibly not among the base-properties. What is plausible (although the precise details are not entirely uncontroversial) is that moral properties supervene on base-properties.[4] Concerning the metaphysics of moral principles, it is helpful to see how neo-Humeans understand scientific laws. The substantivist about scientific laws, takes laws to be something over and above their instances. For example, Maudlin's (2007) account takes scientific laws to be metaphysically primitive and irreducible to the instances they govern. Neo-Humean laws, on the other hand, are simply the sum of every instance that falls within their scope. As Psillos (2014) notes, scientific laws are composite objects: they are the fusion of the sum of their instances. Analogously then, moral principles are nothing but the fusion of the particular moral entities that fall within their scope.

What does this entail about the logical form of moral principles? Consider moral principles in the abstract. As any kind of lawlike generalization, moral principles have the following form:

---

[3] Two caveats are worth mentioning here. First, there could be other relations *alongside* composition which connect macro-properties with base-properties (e.g. other "small-g" relations as per Wilson [2014] or some Humean-friendly relation of grounding [e.g. Sider 2018: 10-1]). Secondly, it is controversial what the *bearers* of base-properties are supposed to be. Traditionally, they are spacetime points (cf. Ladyman & Ross 2007). Other Humean accounts posit more liberal objects to accommodate cases of quantum entanglement (Bhogal & Perry 2017). Thankfully, I don't have to take a particular stance towards this.

[4] Roberts (2017: 198 fn. 3) interprets Sturgeon's view as allowing the possibility of fundamental moral properties. I think this claim is too implausible to consider.

(*) (Necessarily) (For every x) (Fx fully because Gx)[5]

But according to NHM, (*) fully reduces to the following formulation:

> **(List)** (Necessarily) [(Fx$_1$ fully because Gx$_1$) and (Fx$_2$ fully because Gx$_2$) and … (Fx$_n$ fully because Gx$_n$)][6]

This would not be the case if moral principles referred to entities that are wholly distinct from their instances (see also Berker [2018]). For this reason, I will say that, at the propositional level, moral principles are mere generalizations.

I take it that if there's anything philosophically distinctive about moral generalizations it's their non-substantivity which entails that (*) fully reduces to (List). Note that, for someone who accepts the existence of moral principles, there is nothing particularly controversial about (*). The existence of the necessity-operator is trivial. Everyone accepts that, if there are moral principles, they hold with some kind of necessity. Summaries of moral phenomena can hold with metaphysical, normative (Fine 2002; Rosen 2017; 2020; 2021) or even nomological necessity (Dretske 1977). Nor is

---

[5] This is a simplification. As noted, it is plausible that the form of moral principles involves separate clauses for enablers/disablers, right/bad-makers, and, perhaps, ceteris paribus clauses (as per Väyrynen's [2009] 'hedged' view) (cf. Ridge & McKeever 2006). It could also turn out that moral principles are not universally quantified conditionals to begin with (e.g. they could be 'defeasible generalizations' as per Lance & Little [2004]) (cf. Scanlon 2014). But such complications are orthogonal to the point I am trying to illustrate.

[6] It might be objected that, regardless of one's view concerning the metaphysics of moral principles, (*) *trivially* entails (List) (cf. Schaffer 2018: sec. 4.1.). In response, I would say that even if it is true that (*) can be *translated* into (List), it does not follow that (List) would fully express the content of the relevant view. If, following Rosen (2017), one takes moral principles to be relations between universal-like entities, then it would be false to claim that such a view can be adequately expressed in terms of (List). In this sense, when I claim that, according to the moral-generalizations view, (*) fully reduces to (List), I mean that (List) fully captures the content of that view.

the because-operator particularly suspect. Intensionalists can interpret the because-operator as a material conditional or biconditional, whereas hyperintensionalists can take "fully because" to generate a context where the substitution of intensionally equivalent entities doesn't necessarily preserve its overall truth-value.[7]

The view that moral principles are mere generalizations at the propositional level, and fusions of objects at the genetic level might seem striking to some. Many (Robinson 2011; Rosen 2017; Berker 2018) argue that if moral principles are nothing but the sum of their instances, then they cannot be explanatory towards those instances. This is a challenge for the proponent of NHM since the explanatoriness of moral principles is a key datum from first-order normative ethics.[8] Also, Robinson (2011: 298) argues that the inherent non-modal character of the base-properties upon which moral properties supervene, makes it difficult for NHM to have a plausible metaphysical story about the directionality of pro tanto reasons (i.e. the fact that a pro tanto reason to φ favours φ-ing even in the presence of defeaters). These are serious challenges that the proponent of NHM must face. But, as noted, in this paper I will defend the *contingentist* aspect of NHM. In this sense, a full defense of NHM must proceed in a piecemeal fashion.

---

[7] Concerning the intensionalist framework, I take it that fundamental or "pure" moral principles would involve a biconditional, whereas a derivative, or "mixed", principle would involve the material conditional. Value pluralists can either accept multiple principles involving a conditional, or one biconditional involving separate disjuncts for every type of value. Hyperintensionalists have at their disposal a variety of metaphysical determination relations (Rosen 2010; Wilson 2014).

[8] For references, see Robinson (2011: 290; 294 fn. 7) and Walden (2016). Fogal & Risberg (2020) take the genuine explanatoriness of moral principles to be a key datum that every theory of moral explanation should accommodate. See also what Schroeder calls the 'Standard Model' of normative explanation (2005: 24). I also take it that the genuine explanatoriness of moral principles is at least *one* of the things that are at stake in the debate between moral generalists and moral particularists.

One might worry, however, that given these problems, NHM is not even worth engaging with. I disagree. First, in the next section I argue that NHM has important virtues concerning moral supervenience and the explanation of moral principles. More importantly, NHM is a part of larger project: neo-Humeanism. This project has global ambitions: it concerns entities that reside at the base-level but also macro phenomena such as mentality and social facts. If there are moral facts, then such facts should also get the same neo-Humean treatment.

It is an understatement to say the neo-Humean project is a project worth taking very seriously. Neo-Humeanism is a strong contender in the literature on scientific laws and a powerful metaphysical framework overall. Its successes can be found in multiple areas in philosophy, ranging from metaphysics (Lewis 1983) to moral metasemantics (Dunaway & McPherson 2016). At the very least, neo-Humeanism about moral principles should be understood as a live conceptual possibility that is worth exploring.

Also, there is a sense in which the view that moral principles are mere generalizations shouldn't be considered as *particularly* controversial. Understanding moral principles as mere generalizations is a metaphysical theory about the status of moral principles. But any metaphysical theory about the status of moral principles is going to involve some level of controversy since metaphysical views, in general, do not immediately follow from first-order discourse. To illustrate, Robinson's (2011) view about moral principles appeals to essentialist facts about moral dispositions. But no facts about essences can be directly read off from the way moral discourse usually operates. First-order normative discourse, at least at a surface level, is metaphysically neutral. For this reason, I take NHM to be a serious candidate concerning the metaphysical status of moral principles.

3. The positive case for NHM

According to NHM, moral principles (in the propositional sense) are generalizations of moral phenomena. The principle of utility states that moral rightness holds in virtue of utility-maximization. Under NHM, the principle of utility translates as the fusion of every particular fact of the following sort: something (acts, policies, etc.) is morally right in virtue of maximizing utility. Before proceeding to make the case for NHM, I want to make the following clarifications.

First, so far, I've presented NHM as a generalist proposal in the sense that the instances figuring in (List) are objectively similar (i.e. by maximizing utility). But NHM is also open to a particularist reading. It could be that (List) involves radically heterogeneous base properties which, nevertheless, all give rise to instances of moral rightness. As it stands, nothing in the neo-Humean framework favours the generalist or the particularist account. This is a good thing since a metaphysical account about the nature of moral facts should allow for this sort of flexibility. Still, it is dialectically reasonable to keep framing things along generalist lines. This is because my main rivals in this debate are generalists (e.g. Robinson 2011; Rosen 2017b) and I also think that moral generalism is independently plausible. Also, many of the features of my account can be easily extrapolated in a way that accommodates particularism (I will return to moral particularism in the final section).

Secondly, the principle of recombination figuring in NHM needs to be explicated in more detail. As noted, base-properties are freely recombinable and depending on their distribution at a given possible world, different moral principles are exhibited. But it is crucial to stress that the principle of recombination only applies to base-properties. In this sense, NHM does not allow for the possibility of two possible worlds being base-identical but different in terms of their moral principles. This restriction of the principle of recombination is independently plausible. Moral

principles, according to NHM, are simply fusions of their instances. But it would be bizarre to claim that two entities can be different in terms of their moral properties but still base-identical (more on this in section 3.2.).

Thirdly, it is plausible that not just *any* sort of variation in the distribution of base-properties is going to result in a difference in moral principles. The difference in base-properties needs to be *morally relevant*. This is evident if one considers the multiple realizability of moral rightness. To compare, not just any difference is relevant concerning the instantiation of *scientific* laws. This is a good thing since we need to keep scientific laws *fixed* in order to evaluate certain counterfactuals. The same applies to moral principles.

Finally, there is a puzzle concerning the possibility of uninstantiated moral properties. It is perfectly reasonable to claim that there could be a world in which a moral principle holds but there are no instantiated moral properties. But NHM seems to not allow this possibility since (List) involves the fusion of every *instantiated* moral property. In response, the first thing to note is that (List) is not temporally restricted: it involves past, present, and *future* moral instances. What about worlds not involving such future instances? There should be room for worlds of this sort. But even such worlds could involve *dispositional* moral properties which would be triggered if the appropriate conditions were in place. Does this move go against the neo-Humean framework? It does not. The proponent of NHM can countenance moral dispositions as long as they are ultimately reduced in terms of base-properties. Of course, someone could be skeptical about whether neo-Humeans can accommodate such dispositional talk. But this is not a distinctive issue for NHM and I am prepared to pass the buck

concerning that problem.[9] After these clarifications, I now turn to the positive aspect of NHM.

3.1. Explaining moral principles

NHM seamlessly provides an answer to the question of what explains moral principles themselves. I will assume that there are many ways to explain a phenomenon. One of these ways involves the explanation of a phenomenon in terms of its parts. In the literature on mechanistic explanation there are many examples of mechanisms being explained in terms of their constituents (for example, the explanation of the mechanism of protein-synthesis in terms of the components and the processes which comprise it).

According to NHM, at least one way of explaining moral principles is in virtue of the particular instances which compose them. In this sense, the principle of utility is *explained* in terms of the moral goodness of every particular action and its ground. Another way to see this involves examining NHM principles at the propositional level (i.e. as mere generalizations). Even opponents of the NHM approach accept that generalizations are explained via their instances. The generalization that corvus albicollis ravens reside in Africa is explained in terms of every individual corvus albicollis raven residing in Africa.[10] The view that moral principles are mere generalizations might be objectionable for other reasons (e.g. explanatoriness reasons), but it is uncontroversial that if they *were* as such then they would explainable via their instances.

---

[9] It could be objected that there are worlds exhibiting moral principles in the *absence* of moral dispositions. But this move is either question-begging or entails a controversial ontology of brute moral principles.

[10] As Rosen (2010: 120-1) notes, the explanation of such generalization would plausibly also involve a *totality* fact (i.e. a fact stating that the relevant instances involve such-and-such features and *that's it*). I will ignore this complication for presentation purposes.

The question of what explains moral principles hasn't received much attention in the relevant literature. But this should not be taken to show that the question is unimportant. The explanation of moral principles is of obvious interesting for anyone who thinks that moral phenomena, in general, can be explained. After all, moral principles are, themselves, moral phenomena. If an explanation of them is available, then that would be a philosophically interesting result. Also, the *way* a moral principle is explained, has direct consequences for the status of the moral principle itself. An explanation of a principle in terms of a stance-dependent fact is surely different from an explanation that appeals to a fact that holds stance-independently.[11]

It could be objected that the question of what explains moral principles is ill-formed as moral principles have *no* explanation. Hattiangadi (2018: 608) notes that it could be that there is no explanation for why laws of nature are the way they are. Similarly, there is no explanation for why moral principles are the way they are (see also Shafer-Landau [2003: 47]). But neo-Humeanism goes against claim that laws of nature have no explanation. According to the neo-Humean program, laws of nature are fusions of base-objects (or objects whose constituents are base-objects). Similarly, the proponent of NHM should hold that moral principles are explainable via their instances. So, there is at least *one* way in which moral principles can be explained.

Perhaps it could be argued, instead, that moral principles are not even *apt* to be explained. In response, I take issue with the very notion of non-aptness for explanation. It is simply not clear what it means for something to be non-apt for explanation. The best way I can think of to spell out this sort of non-aptness is by appealing to the notion of zero-grounding (Fine 2012; Litland 2017; cf. Bader 2017). Roughly speaking, zero-grounding is a form of explanatory relation that connects some explanandum with an *empty* set of explanantia. In this Perhaps it could be claimed that saying that P is not apt

---

[11] For a similar point see Väyrynen (2018) and Rosen (2017b: 287).

for explanation is elliptical for saying that P is zero-grounded (in the sense that P is "explained" by an empty set of statements).

It should be noted that zero-grounding is a philosopher's posit with a specific purpose in mind: to provide a way to explain meta-explanations like "Z explains that P explains Q". The idea is that such meta-explanation facts are different from regular explanations and require special treatment (Wallner 2018). But even if zero-grounding is the proper way of dealing with this problem, it has little to do with moral principles. Moral principles are not like meta-explanations or anything of the sort. Even in their minimal form, moral principles are modalized generalizations and, thus, not the proper explanandum for the zero-grounding relation.[12]

I've argued that NHM entails a constitutive explanation of moral principles. But there is yet *another* way moral principles are explained under NHM. As noted, NHM is a part of the neo-Humean project: macro-phenomena are all composed out of base-objects and their properties. There is an implicit but dialectically powerful monist assumption at play here. In this sense, everything that exists is ultimately traced at the *same* base. This is important because it allows for top-down explanation of moral phenomena. There are plausible examples of such explanations (call them subsumption-explanations). Kepler's laws of orbit are subsumed under Newton's more general laws of motion. Analogously, in the moral domain, mixed moral principles are explained in this way by pure moral principles (Scanlon 2014; Rosen 2017b). Assume that every killing that occurs in the context of self-defense minimizes utility. If so, then the principle

---

[12] There is a unionist assumption at play here according to which grounding is identical to some form of explanation (cf. Raven 2015). For another notion of explanation non-aptness which fails for similar reasons see Dasgupta (2014).

"One ought not to kill unless one has to defend themselves" is subsumed by the principle of utility.[13]

But what about pure moral principles? After all, what is at stake is *their* top-down explanation, not simply the explanation of mixed moral principles. The most obvious way this could happen is by subsuming such pure principles in terms of the *most general* regularity holding at the Humean mosaic. Lewis sketches the idea of a predicate F which is supposed to apply to all the things that hold at the base-level of a given world (1983: 367). If so, a generalization of the form $\forall x(Fx)$ would express a regularity involving every other regularity holding at that world. In this sense, $\forall x(Fx)$ would subsume the principle of utility. Of course, this subsumption is relatively uninformative, but it still illustrates the idea that every regularity involving macro-phenomena can be traced back at the same Humean level.[14] Whether there are more informative subsumptions of the principle of utility (e.g. by appealing to a more 'basic' regularity involving normative reasons) is something that requires further assumptions which I will not undertake in this paper. But the fact that pure moral principles are subsumed by the Humean mosaic *in one way or another* is still a substantive result.[15]

---

[13] Other examples of principle-to-principle subsumption can be found in Mill (Brink 2013) and Ross (1939). I resist using the more familiar principle-to-instance type of explanation because, as noted in section 2, it is controversial whether NHM can accommodate such explanations.

[14] Lewis calls F a 'perverse' predicate posited by an ideal, final, theory of everything (1983: 367). Granted, it is dubious that there could ever be such a theory. But I take it that the idea behind it is conceptually coherent.

[15] The non-moral analogues of moral principles (in my sense) are bridge-laws figuring in the special sciences (e.g. psychophysical laws). Similarly, even though the exact details of the subsumption of these laws in terms of more general laws is a non-trivial empirical issue, nevertheless it is uncontroversial that they are dependent on the organization of the world at which they operate (e.g. Horgan [1978] takes the truth of a bridge-law to supervene on the truth of the reducing theory being involved in that bridge-law).

Overall, I take subsumption-explanations to be intuitive and plausible. Here's a potential worry. Subsumption-explanations operate by showing how an explanandum-fact is an instance of a more general phenomenon. It follows that maximally general phenomena cannot be subsumed under any other phenomena. It might then seem that pure moral principles cannot have an explanation of this sort. But, bracketing extreme views, moral principles (in the genetic sense) are not a part of the fundamental constituents of the universe. It is no accident that even non-naturalists like Rosen agree that such a view would be too controversial to defend (2017b: 138). But if moral principles are *not* fundamental in this sense, then it is not at all obvious why they would lack an explanation.

3.2. Explaining Moral Supervenience

I have argued that there are at least two ways to explain moral principles under NHM. As I will show, it is precisely this feature that allows NHM to have an explanation of moral supervenience that is superior to other more traditional proposals. The truth of moral supervenience is something that every metaethical theory needs to accommodate. Moral supervenience is a thesis concerning the modal covariation between a set of moral and non-moral properties:

> **(S)** (Necessarily) (For every x) (If there's a moral difference in x then, (Necessarily) there is a non-moral difference in x)

There are many ways to make (S) more precise. For example, it is assumed in the literature that both necessity operators need to indicate *metaphysical* necessity.[16] But, as noted, NHM denies the existence of necessary connections. So how could it explain

---

[16] It is true that many philosophers take the first operator to indicate *conceptual* necessity. I will bracket these views as I am convinced by Dreier (2019) that this view is untenable.

(S)? Since this issue ultimately bears on the modal status of moral principles, I will bracket it and return to in the relevant section. For now, I will focus on the way NHM explains (S) without worrying whether the relevant modality is explained.

(S) is easily explainable under NHM. In fact, (S) is simply an instance of Humean supervenience (i.e. the thesis that everything supervenes on base-properties). Given that the base-properties occupying the Humean mosaic are paradigmatically non-moral then, assuming the truth of Humean supervenience, (S) follows. As mentioned in section 2, Humean supervenience is underwritten (and, thus, *explained*) but a composition relation. In this sense, (S) is also underwritten (and, thus, explained) by a composition relation.

(S) is not the only moral supervenience thesis that needs to be explained given that the relevant set of supervenient properties can be made more fine-grained. It is plausible that moral properties also modally covary with some interesting, *specific*, supervenience base. For example, Boyd's (1988) preferred base would involve properties concerning a cluster of important human needs, Copp (1995) would focus on properties contributing to societal stability, Bloomfield (2018) on properties pertaining to some Aristotelian notion of well-being, etc. At any rate, there's an interesting, non-trivial, supervenience thesis that requires an explanation:

> **(S*)** (Necessarily) (For every x) (If there's a moral difference in x then, (Necessarily) there is a *specific* non-moral difference in x)

(S*) also gets a clear explanation under NHM. Moral properties supervene on some specific set of natural properties because there's an underlying moral principle at play. For example, the fact that moral properties, in general, are appropriately related to properties pertaining to important human needs (as per Boyd's [1988: 4.3.] homeostatic consequentialism) explains the fact that something cannot exhibit a

difference in its moral properties without a difference in its properties pertaining to needs.[17]

The strategy of appealing to moral principles (in whichever form) to explain moral supervenience is not novel (cf. McPherson 2021: 4.3.). My claim is, rather, that the proponent of NHM has a dialectical edge over their rivals. To see this, consider how the dialectic concerning the explanation of moral supervenience usually plays out. First, it is noted that a moral supervenience thesis like (S*) cannot be left unexplained. As McPherson (2012) notes, theories that posit unexplained phenomena are worse to those that don't, ceteris paribus. Then, a general metaphysical fact which supposedly underlies (S) is posited. But the proponent of this strategy faces a further problem. Even though, (S*) are plausibly explained by an underlying moral principle, this principle is itself unexplained. In this sense, the challenge from bruteness arises once again (as per McPherson's [2012] "bruteness revenge").

NHM avoids this worry as it involves at least two plausible explanations of moral principles. Moral principles, under NHM, are explained constitutively and via subsumption. At the initial level, NHM is explanatorily richer to its rivals because it posits an additional explanatory step (which, in turn, can be cashed out in at least two

---

[17] The explanation of moral supervenience is usually presented as a challenge against non-naturalist moral realists. But note that both (S) and (S*) require an explanation even assuming that moral properties are natural (compared to the trivial, for the naturalist, supervenience thesis that moral properties supervene on natural properties) (cf. Sturgeon 2009). It should also be noted that, it is not necessarily the case that neo-Humeanism is a naturalistic project (ditto for NHM). Even though the base-properties of the Humean mosaic are tracked by the physical sciences, it doesn't follow that everything that is related to that base is, itself, a natural property. Under neo-Humeanism, objects that bear moral properties are composed out of objects bearing base-properties. But, for all we know, this could be *partial* composition claim. For this reason, it is no accident that McPherson (2013) utilizes the neo-Humean framework to develop a taxonomy between different types of moral naturalism and non-naturalism.

different ways depending on the explanation one focuses on). But it should also be noted that these two types of explanations are importantly connected since they follow from the neo-Humean framework itself.[18] According to NHM, moral supervenience is simply an instance of Humean supervenience. Of course, one could still ask what explains the Humean mosaic. But I take it that the view that the Humean mosaic is brute is more palatable than the view that some moral principle is brute. Under these considerations, I take NHM to provide a powerful explanation of moral supervenience.

4. In Defense of Moral Contingentism

I have argued that NHM provides an explanation of moral supervenience theses, *bracketing their modal status.* But recall that a plausible way of interpreting the necessity operators figuring in (S) and (S*) appeals to metaphysical necessity. So, we have the following datum:

> **(Necessitarian)** Both necessity operators figuring in moral supervenience theses are metaphysical necessity operators.

NHM cannot accommodate (Necessitarian). Recall, that one of the key features of neo-Humeanism is the metaphysical *contingency* of the distribution of the Humean mosaic. This means that NHM must construe moral principles as metaphysically contingent: different possible worlds will have different moral principles depending on the relevant distribution of base-properties. In this sense, NHM would take the first necessity operator to signal metaphysical necessity, and the *second* necessity operator to signal a weaker kind of necessity (i.e. in terms of the moral principles that hold at a given world). I will call this second, relativized, kind of necessity *Humean* necessity:

---

[18] This is a case of explanation-integration (i.e. an instance where two explanations directly complement one another) which is a powerful theoretical virtue (Wayne 2018).

(S**) (Necessarily_{MET}) (For every x) (If there's a moral difference in x then, (Necessarily_{HUM}) there is a non-moral difference in x)

There are two ways to argue against (Necessitarian) and in favor of (S**). The first one involves direct arguments against (Necessitarian). The second one involves dealing with the claim that the *rejection* of (Necessitarian) entails objectionable consequences. I will address the second kind of arguments in the next section. In this section, I will argue that the rejection of (Necessitarian) is not only palatable, but also a *feature* of NHM. Specifically, there are three reasons one should be skeptical of (Necessitarian).

First, it has been argued (Roberts 2018; Hattiagandi 2018; Rosen 2020; 2021) that violations of strong moral supervenience are conceivable. On the assumption that there is a plausible link between conceivability and possibility, it follows that there are good reasons to be skeptical of (Necessitarian).[19] Additionally, Hattiagandi (2018: sec. 2) (following Hills [2009]) appeals to the psychological phenomenon of imaginative resistance to explain away our apparent difficulty to conceive pretheoretically morally abhorrent scenarios which in some possible worlds are morally good. For example, if we try to imagine a world "where Hitler did no wrong, we are prevented from doing so by a powerful feeling of moral disgust" (Hattiangadi 2018: 594).

I find the appeal to imaginative resistance powerful, but I think that it can be strengthened. Even if we grant that there's a strong link between conceivability and possibility, imaginative resistance can only explain away morally deviant worlds (cf. Walton 2006). But what about less extreme scenarios? Consider a particular moral fact P holding at the actual world. Then, consider a world which is more or less the same as

---

[19] Note that one need not appeal to the stronger, and more controversial, thesis that conceivability entails possibility to make this point (cf. Yablo 1993).

the actual world in terms of base-properties which, nevertheless, does not involve P. But surely there are worlds of this sort which are not morally abhorrent. There might by a world which is more or less the same as the actual world minus the fact that it is morally impermissible to eat asparagus. But this is not a morally abhorrent world.

Also, I have a general worry about the interpretation of the phenomenon of imaginative resistance. Simply put, explaining cases of imaginative resistance in terms of emotional inability or unwillingness (as per Driver [2008]) is controversial to begin with. There are prominent accounts of imaginative resistance according to which, for a case C, there's resistance in imagining C because C is metaphysically impossible. Walton (1994), for example, claims that our (presumed) resistance to imagine a scenario where moral supervenience fails, is *precisely* because moral supervenience holds with metaphysical necessity! So, appealing to imaginative resistance is, potentially, dialectically weak.

Instead, I would like to propose that cases of such imaginative resistance are less widespread than the objector would like us to believe. Specifically, it is hard to make sense of moral disagreement without assuming the conceivability of possible worlds with different moral principles. What I have in mind are cases of *deep,* or fundamental, moral disagreement. These moral disagreements are not disagreements about the non-moral facts. Rather, they are cases where two moral agents disagree about the fundamental moral principles at play. For example, the consequentialist fundamentally disagrees with the deontologist about the permissibility of some action *because* of their underlying difference in moral principles.

Of course, I take no stance as to whether their disagreement is in principle resolvable or not.[20] What is important is that, in such cases of disagreement, the parties involved conceive of worlds where their preferred moral principles are different. A typical way for the deontologist to object to the consequentialist, is to propose a consequentialist world where some pretheoretically morally objectionable action is morally permissible.[21] But in order for this move to make sense, the relevant scenario must be *conceivable*. And if it's conceivable, then this puts pressure on (Necessitarian).

Finally, another reason to be suspicious of (Necessitarian) involves the very notion of metaphysical necessity. For one, metaphysical necessity is controversial to begin with. Some claim that there are reasons to resist adopting such a distinctive notion of necessity (Priest 2018). Others adopt a pragmatic, anti-realist, treatment of metaphysical necessity which is arguably at odds with how (Necessitarian) is usually interpreted (Divers & Elstein 2010). More importantly, though, I am inclined to agree with Scanlon (2017: 892) that, in whichever form, metaphysical necessity does not

---

[20] It wouldn't be resolvable (in a non-trivial way, at least) if, following Wong (2006), the relevant scenario is morally ambivalent (i.e. it is a case where both moral principles are correct). Or, it could be resolvable in the sense that one of the two parties is mistaken, or in the case where both parties are referring to the same principle without realizing it (e.g. if deontology is consequentializable) (Brown 2011).

[21] It could be objected that all that is required is that the relevant parties can *reason under the supposition* that a rival moral principle holds (without having to *conceive* a world where that principle holds). In response, I want to suggest that reasoning under the supposition that P involves the conceivability of P. Conceivability should be distinguished from imaginability. It is certainly possible to reason under P without being able to imagine P (e.g. in cases of abstract mathematical reasoning). To compare, conceivability should be understood as a way of reasoning through the use of the appropriate concepts (for a more precise account see Hill [2016: sec. 2]). But I find it hard to understand what it would mean to reason under the supposition that P without employing the concept of P. For this reason, I take these mental processes to not be significantly different.

explicitly figure in our folk moral discourse (cf. Rosen 2021: 263). And if it doesn't, then it is unclear how (Necessitarian) could.

There are other direct arguments against (Necessitarian) which I cannot address here.[22] Still, the reasons I presented generate independently good reasons to be skeptical of (Necessitarian). This is a good thing for NHM since the rejection of (Necessitarian) is an essential component of the view. Of course, it could still be the case that, all things considered, the rejection of (Necessitarian) is worse than its adoption. I will address these challenges in the next section.

5. Normative Authority and Epistemic Access

It has been recently suggested that contingentist accounts cannot accommodate two key features of moral principles: their normative authority and the fact that they're epistemically accessible. On the face of it, these features seem importantly different. However, as it will become apparent, the objections that appeal to them are structurally the same.

Consider normative authority. Väyrynen (2021) rightly argues that a metaphysical theory of moral principles needs to be able to give an account of their normative standing. Moral principles are supposed to be authoritative in a robust, non-arbitrary, way (Väyrynen 2021: 211). Whatever normative authority these principles have, it is a stronger kind of authority compared to the authority that principles of (say) etiquette involve. However, it seems that moral contingentist accounts are unable to meet that feature of moral principles. To illustrate, if deontology is true at the actual world, but there is a possible world in which act-utilitarianism is true, then why should

---

[22] For example, Roberts (2018) argues that there are conceptually coherent possible worlds where there is normativity "all the way down" (cf. Bohn 2018).

deontology be normatively authoritative instead of act-utilitarianism? (Väyrynen 2021: 210-1)

It should be noted that Väyrynen has a *particular* version of moral contigentism is mind when he raises this worry. Rosen (2020; 2021), building on Fine's (2002) notion of normative necessity, proposes a contingentist theory where moral principles are metaphysically contingent and *fact-independent.* He takes a fact p to be fact-independent when "p is the case and would have been the case no matter how things had been in wholly nonnormative respects" (Rosen 2020: 219). In this sense, the putative fact-independence of moral principles is one of the ways one *could* explain their metaphysical contingency.[23] If moral principles hold independently of the actual non-normative facts, then there is no reason why they wouldn't also occur in worlds with different non-moral facts. Conversely, there is also no reason why there couldn't be a difference in moral principles between two worlds which are, nevertheless, identical in their wholly non-normative respects.

Let two possible worlds w_1 and w_2 be identical in wholly non-normative respects but different in terms of their underlying moral principles. This means that it is a brute fact whether some moral principle P holds at w_1.[24] But if P exerts normative authority at w_1, and P is a brute fact, then the explanation behind P's normative authority at w_1

---

[23] I say "could" because, as it will become apparent, fact-independence and metaphysical contingency can come apart.

[24] It could be objected that this move is too hasty. Assuming fact-independence, P is independent from the way w_1 is organized in terms of its wholly non-normative features. But could there be *other* kinds of features that plausibly explain P? I don't think so. Note that we are dealing with fundamental (i.e. pure) moral principles. In this sense, every other moral fact that obtains in w_1 would need to follow, ex hypothesi, from P (otherwise, P wouldn't be a *fundamental* moral principle). So, there is nothing non-normative or distinctively moral that could explain P. I also can't think of any non-moral, but normative, feature that could serve as a plausible explanans.

needs to appeal to the fact that P holds in that world. But this is not a good explanation.[25]

Thankfully there is a way for NHM to escape this result. Specifically, NHM is a contingentist account which doesn't take moral principles to be fact-independent. This should be clear by now given that NHM can provide at least two plausible explanations of fundamental moral principles. According to one of them, moral principles are explained by their instances which, in turn, ultimately depend on the distribution of base-properties holding at that world. So, if there's variation between two worlds in terms of their moral principles, then that difference is reflected in a difference in their base-properties.

For this reason, P's normative authority in w_1 is not an arbitrary fact (cf. Väyrynen 2021: 211). There are facts holding in w_1 in virtue of which moral agents at w_1 have a concern to act according to P. The w_2-counterparts of these agents will have a concern to act according to some other moral principle (P*). But, again, if P* is different from P, there will be a significant difference in the base-properties of w_2, thus making P*'s normative authority at w_2 equally non-arbitrary.

It could be objected that without a full account of normative authority and its metaphysical ground it is difficult to assess whether NHM has definitively dealt with the challenge. I disagree. Building a particular theory of normative authority *into* NHM is problematic since it is a substantive issue whether some theory of normative authority

---

[25] I say 'good' explanation because P's moral authority could, trivially, be explained in virtue of something like P's intrinsic nature (for discussion on this point see Väyrynen [2021: 211, 213]). It could be said that Väyrynen is too quick in dismissing this option since there are analogous and promising strategies being employed in the literature on scientific laws. For example, primitivist accounts of laws of nature similarly try to explain the governing feature of laws by appealing to their primitive nature (Schaffer 2016).

is the correct one. Instead, it is sufficient to say that *whatever* normative authority turns out to be (assuming it is not a brute fact) it will hold in virtue of base-facts. And given that the relevant moral principles would hold in virtue of the *same* base-facts, it follows that when a moral principle holds at some world, the normative authority of *that* principle will be traced at the base-facts of *that* world.

So NHM can accommodate the datum that moral principles exert normative authority. But what about the datum that moral principles are epistemically detectable? Dreier (2019: 17) has recently suggested that if there are two possible worlds which are identical in their non-moral facts but different in terms of their moral principles, then whether some action is good is ultimately a question of moral luck. If not being an egoist is morally good at w_1 but morally bad at w_2, and if w_1 is identical to w_2 in terms of non-moral facts, then it is epistemically difficult to tell (from the perspective of someone at, say, w_1) whether they should be an egoist or not.

But, again, Dreier's worry is based on a particular understanding of moral contingentism according to which moral principles are fact-independent. As noted, NHM holds that moral principles are ultimately traced at the base-level. Assuming that base-level properties (or properties that supervene upon them) can be epistemically accessed, it follows that moral principles can be epistemically accessed. In this sense, whether some moral principles hold at w_1 or w_2 is, at least in principle, detectable.

One could insist that moral principles *are* epistemically detectable even if they're fact-independent. Perhaps, there are *directly* accessed through some kind of Moorean faculty of intuition. Relatedly, Rosen (2020: 231) suggest that "[i]f some theory tells me that I can't know that it is wrong to cause intense pain in a non-consenting victim for fun (because my belief in this fact is not sensitive to the fact that makes it true), I am inclined to say: so much the worse for your philosophical theory of knowledge, since I certainly do know/justifiably believe this moral fact.".

This might be true, but it is definitely not the perceived view on moral epistemology. To compare, there are plausible coherentist accounts of moral justification which take moral principles to be justified via wide reflective equilibrium which are, nonetheless, neutral towards many different metaethical views (e.g. concerning the naturalist/non-naturalist distinction).[26] Also, note that Rosen's example involves a quite distant possible world. But what about more remote worlds where the moral principles are slightly different? I am inclined to agree with Rosen that this case is less worrisome than the case presented by Dreier (Rosen 2021: 276-7 fn. 23). But it would still be nice if even these small variations between worlds were epistemically detectable. NHM can deliver that result.[27]

6. The Modal-Miracle Problem

I have argued that the normative authority problem and the epistemic access problem can be dealt with by taking moral principles to be fact-dependent. But there is a more substantive challenge which directly targets the contingentist aspect of NHM.

[26] Rosen notes that the difficulties of tracking moral principles "retain whatever force they have regardless of how we classify the worlds in which the moral [principles] are otherwise. There is an analogous puzzle about how we can know whether God exists and what she's like given that we cannot rule out many of the relevant possibilities in the lab or in the analytic armchair. But this problem is not made easier by insisting that the theological facts, whatever they may be, are metaphysically *non-contingent.*" (2021: 277; my emphasis) I agree. But it *does* make a difference if the moral principles are *fact-independent.*

[27] Perhaps, there *should* be cases where the moral principles are epistemically inaccessible. As Rosen notes, there could be worlds in which our beliefs are systematically mistaken (2021: 275). I agree. My view can accommodate this possibility since, as noted, moral principles according to NHM can be epistemically accessed *at least in principle.* The only caveat I would add is that in worlds where the principles are very difficult to access, there would be an *explanation* for this (ultimately traced at the Humean mosaic). Presumably, there is a *reason* why it is very difficult to track the relevant principles (perhaps, there's an evil scientist who intercepts our ability to form true moral beliefs).

Fogal & Riseberg (2020: 179-81) have recently noted that a view like NHM involves a modal miracle.[28] According to NHM, moral principles are dependent on the distribution of the base-properties of a given world. But base-properties are inherently non-modal. So there are worlds with radically different distributions of base-properties, and hence, radically different principles. More importantly, there are also worlds where the distribution of base-properties is as such where there are *no* moral principles. For this reason, NHM seems to be committed to the following thesis:

> **(Miracle)** If there are moral principles holding in the actual world, then this constitutes a modal miracle.

There are three things to note in order to make the challenge more salient. First, the underlying assumption is that that there are *incredibly many* ways in which the base-properties can be recombined. In this sense, given the sheer number of possible worlds in which there are *no* moral principles, it is surely a miracle that our world is not one of them. For this to work, each possible distribution of base-properties must have the same degree of probability.[29] Also, I take it that in order for the objection to be forceful, another assumption is that our world *does in fact* contain moral principles.

---

[28] Two caveats. First, Fogal & Riseberg focus on a view they call "Hyperintensional Humeanism" which has been recently defended by Berker (2018: sec. 9). Berker's view is significantly different from mine but in a non-related sense (roughly, he takes moral principles to not be explanatory vis-à-vis their instances, and he explicitly countenances a robust notion of grounding). Secondly, they frame their problem as a challenge to explain *moral supervenience* but, in essence, it seems to me that their deeper point concerns the distribution of the Humean mosaic itself (cf. Fogal & Riseberg 2020: 180).

[29] Equiprobability presumably follows from a principle of indifference: since we don't know what the distribution of probabilities is over the possible worlds, we should assign the same probability across the board. This assumption can be contested but I will grant it for dialectical purposes. Perhaps, the right attitude in such cases of epistemic uncertainty is to suspend belief instead of assigning equal probabilities (cf. Filomeno 2019). Also, Loewer (MS: 20) has recently argued that some distributions of base-

Finally, the modal miracle worry shouldn't be understood as a particularist challenge against moral generalism. Presumably, even if moral particularism was the case, an analogous worry could be raised (i.e. it would also be a modal miracle that there are particular moral truths at the actual world). This is because even moral particularism plausibly requires some level of organization at the level of base-properties (i.e. at the level of properties that give rise to *resultant* moral properties) (Dancy 1981). After all, moral particularism is compatible with certain supervenience theses.[30] Also, there are versions of moral particularism which are compatible with the existence of moral principles (Väyrynen 2018: sec. 4).

I have two responses against this challenge. First, (Miracle) overgeneralizes. It might be that it is a modal miracle that our world has moral principles. But it will also be a modal miracle that the relevant base-properties are distributed in a way that give rise to these principles. In this sense, there is nothing *particularly* problematic about NHM. Rather, it is a general challenge against neo-Humeanism (Bhogal 2020: 2.3.). The laws of physics, and the principles of biology would also be modally miraculous. To see this, consider an amended version of (Miracle) that specifically targets NHM:

> **(Miracle*)** Given the number of possible worlds that are apt to host moral facts, if there are moral principles holding in the actual world, then this constitutes a modal miracle.

(Miracle*) is most likely false. The number of worlds which fulfil the condition of being apt to host moral facts is significantly lower than the number of every possible world. Given plausible assumptions about what moral facts *could be,* there is only an elite

---

properties have inherently higher probability than other distributions (although this move signals an important departure from the neo-Humean framework).

[30] The particularist can accept that moral properties supervene on *some* collection of non-moral properties (which, given reasons holism will be extremely complex) (Strandberg 2008).

subset of possible worlds where the question of whether there are moral principles could even be raised. For example, assume that sentience is a necessary condition for moral consideration. This minimal requirement already significantly cuts down the relevant modal space: the worlds that give rise to sentient beings are worlds where specific types of biological and psychological facts hold. Some of these worlds will be worlds where moral particularism is true, whereas others will include true moral principles.[31] But it is hardly a distinctive miracle that the actual world is one of *these* worlds.

Secondly, (Miracle) overgeneralizes in yet *another* way. It could be argued that even if it overgeneralizes, (Miracle) is problematic for NHM since NHM is a part of the neo-Humean program. But even in this mitigated form, (Miracle) is a problem for *every* metaphysical theory of moral principles. To illustrate, consider a primitivist account of moral principles according to which the distribution of moral properties at a given world is governed by a primitive moral law.

According to the primitivist, the existence of moral principles in the actual world is not a lucky occurrence. Such moral principles hold because there is a moral law which *ensures* that the distribution of moral properties will be regular for past, present, and future instances. Note that this move doesn't require for the primitive moral laws to be metaphysical necessary. Rather, it appeals to a presumed feature of moral laws: that they are metaphysically uniform.[32]

Under closer inspection, the primitivist move is also subject to a modal miracle type worry. Simply put, we can resist the claim that the primitive moral laws that

---

[31] This is one of the reasons why, contra Rosen (2017a: 858), I take the particularist/generalist debate to be orthogonal to the Humean/non-Humean debate about the nature of moral principles.

[32] The primitivist account works nicely for illustration purposes since other more sophisticated substantivist accounts share the same structure (Rosen 2017a; Robinson 2011). For reasons of space, I won't run the same argument separately for every substantivist view.

govern moral regularities are necessarily metaphysically uniform. Why couldn't it be the case that the moral laws stay metaphysically uniform *up to a certain point in time?* (Beebee 2011). To echo Psillos (2017), even if moral laws are primitive, their uniformity could still have an "expiration date".[33]

One could worry that if a principle has an expiration date, then it cannot be a *moral* principle since moral principles are essentially atemporal. How could it be the case that a moral principle fails to apply to instances that occur after a certain time?[34] At this point one should distinguish between the application conditions and the existence conditions of a moral principle. The former concern the conditions under which a moral principle can be applied (e.g. assuming that P universalizable, if P applies to some particular case C then P should also apply to any other case that is qualitatively the same as C). The latter concern the conditions under which P exists. For example, according to NHM, moral principles hold at a particular world insofar as the base-properties of that world are appropriately organized.

When I note that a primitive moral law could have an expiration date, this has implications for the *existence* conditions of the relevant moral principle. To put things

---

[33] This point is certainly not novel although this is the first time it is applied in the context of moral principles. In response, Hildebrand (2016: 15-6, 20) suggests that the substantivist can appeal to a naturalness constraint to rule out primitive laws with temporally restricted contents. But whatever the merit of this move, it is also available to the neo-Humean. Following Lewis (1983), the proponent of NHM can argue that laws need to involve natural properties (the picture is this: base-properties are perfectly intrinsic natural properties, although they are modally recombinable) (cf. Dunaway & McPherson 2016: sec. 3.1.).

[34] I take it that this is not a worry about the *generality* of that moral principle but about whether it is *universal.* To illustrate, a moral principle could have a narrow scope (i.e. by applying to very specific circumstances) while also being universal (in the sense that it is not restricted to specific persons, times, or places) (Väyrynen 2018: sec. 2).

differently, when the primitive moral law 'expires', the relevant moral principle seizes to exist. But the principle's application conditions remain unchanged. In this sense, even if we grant that moral principles are essentially atemporal, its underlying primitive moral principle might not be. So, the fact that we live in a world with atemporal moral laws constitutes a modal miracle (analogous to the one NHM faces).

Perhaps the primitivist can argue that even if their view falls victim to a similar modal-miracle worry, it is nevertheless explanatorily richer. NHM proposes the following the explanatory story: a moral principle is explained in terms of its instances, which are in turn explained in terms of base-properties. Primitivism accepts the same story but posits an extra step: the distribution of the base-properties is explained in terms of a primitive moral law.

But even if we grant that this additional step is explanatorily valuable, there is a cost in terms of ontological parsimony. NHM has many virtues while staying metaphysically minimal. Also note that the explanation of moral principles in terms of their instances, is not hostage to whether base-properties are explained or not. After, it is well known in the literature that explanations are not (typically) transitive.[35] Explanations can "chain" only when certain specific conditions are met (e.g. the relevant contrast classes are properly aligned). So, even in a gunky world, moral principles would be explained in terms of their constituents. Under these considerations, I conclude that NHM is not particularly prone to (Miracle). Other metaphysical accounts face the same problem, and it is not clear that their proposed solutions are particularly successful.

---

[35] See, e.g., Hicks & van Elswyk (2015). Also, classic multiple realizability arguments for the explanatory autonomy of special science entities (and macro phenomena overall) also apply here (Fodor 1974).

7. Conclusion

In this paper I motivated and defended NHM. NHM provides at least two different explanations of moral principles, while also providing an explanation of moral supervenience which is dialectically superior to other similar proposals. I also noted that the contingentist aspect of NHM is a defensible feature. First, I noted that what I called (Necessitarian) is problematic on its own. Secondly, I showed that NHM can successfully deal with three recent challenges against moral contingentism. The normative authority worry and the epistemological worry can be dealt with by noting that NHM rejects the fact-independence of moral principles. Finally, I noted that the modal miracle worry overgeneralizes in at least two ways. First, what I called (Miracle), is not a *distinctive* challenge against NHM but, rather, a challenge against neo-Humeanism. Secondly, the same type of worry can also be raised against rival accounts about the nature of moral principles.

References

Bader, Ralf M. (2017). The grounding argument against non-reductive moral realism. *Oxford Studies in Metaethics* 1.

Beebee, Helen (2011). Necessary Connections and the Problem of Induction. *Noûs* 45 (3):504-527.

Berker, Selim (2018). The Explanatory Ambitions of Moral Principles. *Noûs* 53 (4):904-936.

Bhogal, Harjit & Perry, Zee (2017). What the Humean Should Say About Entanglement. *Noûs* 51 (1):74-94.

Bhogal, Harjit (2020). Humeanism about laws of nature. *Philosophy Compass* 15 (8):1-10.

Bloomfield, Paul (2018). Tracking Eudaimonia. *Philosophy, Theory, and Practice in Biology* 10 (2).

Bohn, Einar (2018). Normativity all the way down: from normative realism to pannormism. *Synthese* 195 (9):4107-4124.

Brink, David O. (2013). *Mill's Progressive Principles.* Oxford University Press UK.

Brown, Campbell (2011). Consequentialize This. *Ethics* 121 (4):749-771.

Copp, David (1995). *Morality, Normativity, and Society.* Oxford University Press.

Dancy, Jonathan (1981). On moral properties. *Mind* 90 (359):367-385.

Divers, John & Elstein, Daniel Y. (2012). Manifesting belief in absolute necessity. *Philosophical Studies* 158 (1):109-130.

Dreier, Jamie (2019). Is there a supervenience problem for robust moral realism? *Philosophical Studies* 176 (6):1391-1408.

Dretske, Fred I. (1977). Laws of nature. *Philosophy of Science* 44 (2):248-268.

Dunaway, Billy & McPherson, Tristram (2016). Reference Magnetism as a Solution to the Moral Twin Earth Problem. *Ergo: An Open Access Journal of Philosophy* 3.

Filomeno, Aldo (2019). Are non-accidental regularities a cosmic coincidence? Revisiting a central threat to Humean laws. *Synthese* 1:1-23.

Fine, Kit (2002). Varieties of Necessity. In Tamar Szabo Gendler & John Hawthorne (eds.), *Conceivability and Possibility.* Oxford Up. pp. 253-281.

Fine, Kit (2012). Guide to Ground. In Fabrice Correia & Benjamin Schnieder (eds.), *Metaphysical Grounding.* Cambridge University Press. pp. 37--80.

Fodor, J. A. (1974). Special sciences (or: The disunity of science as a working hypothesis). *Synthese* 28 (2):97-115.

Fogal, Daniel & Risberg, Olle (2020). The Metaphysics of Moral Explanations. *Oxford Studies in Metaethics* 15.

Hare, R. M. (1963). *Freedom and Reason.* Oxford, Clarendon Press.

Hattiangadi, Anandi (2018). Moral supervenience. *Canadian Journal of Philosophy* 48 (3-4):592-615.

Hicks, Michael Townsen & van Elswyk, Peter (2015). Humean laws and circular explanation. *Philosophical Studies* 172 (2):433-443.

Hildebrand, Tyler (2016). Natural Properties, Necessary Connections, and the Problem of Induction. *Philosophy and Phenomenological Research*:668-689.

Hill, Christopher (2016). Conceivability and Possibility. In Herman Cappelen, Tamar Szabó Gendler & John Hawthorne (eds.), *The Oxford Handbook of Philosophical Methodology.* Oxford University Press.

Hills, Alison (2009). Supervenience and moral realism. In Hieke Alexander & Leitgeb Hannes (eds.), *Reduction, Abstraction, Analysis.* Ontos Verlag. pp. 11--163.

Horgan, Terence E. (1978). Supervenient bridge laws. *Philosophy of Science* 45 (2):227-249.

Jaag, Siegfried & Loew, Christian (2018). Making best systems best for us. *Synthese*:1-26.

Jaag, Siegfried & Loew, Christian (2020). Why Defend Humean Supervenience? *Journal of Philosophy* 117 (7):387-406.

Ladyman, James & Ross, Don (2007). *Every Thing Must Go: Metaphysics Naturalized.* Oxford University Press.

Lewis, David (1973). *Counterfactuals.* Blackwell.

Lewis, David (1983). New work for a theory of universals. *Australasian Journal of Philosophy* 61 (4):343-377.

Litland, Jon (2017). Grounding Grounding. *Oxford Studies in Metaphysics* 10.

Little, Margaret & Lance, Mark (2004). Defeasibility And The Normative Grasp Of Context. *Erkenntnis* 61 (2-3):435-455.

Loewer (MS), Barry, Are Humean Laws Flukes?

Loewer, Barry (2007). Laws and Natural Properties. *Philosophical Topics* 35 (1/2):313-328.

Loewer, Barry (2020). The package deal account of laws and properties. *Synthese*:1-25.

Maudlin, Tim (2007). *The Metaphysics Within Physics.* Oxford University Press.

McKeever, Sean & Ridge, Michael (2006). *Principled Ethics: Generalism as a Regulative Ideal.* Oxford University Press.

McPherson, Tristram (2012). Ethical Non-Naturalism and the Metaphysics of Supervenience. In *Oxford Studies in Metaethics Vol 7*. pp. 205.

McPherson, Tristram (2013). What is at Stake in Debates among Normative Realists? *Noûs* 49 (1):123-146.

McPherson, Tristram (2021). "Supervenience in Ethics" in *The Stanford Encyclopedia of Philosophy* (Summer 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2021/entries/supervenience-ethics/>

Priest, Graham (2018). Metaphysical necessity: a skeptical perspective. *Synthese* 198 (Suppl 8):1873-1885.

Psillos, Stathis (2014). Regularities, Natural Patterns and Laws of Nature', *Theoria* 79: 9-27.

Psillos, Stathis (2017). Induction and Natural Necessities. *Journal for General Philosophy of Science / Zeitschrift für Allgemeine Wissenschaftstheorie* 48 (3):327-340.

Raven, Michael J. (2014). Ground. *Philosophy Compass* 10 (5):322-333.

Roberts, Debbie (2017). Depending on the Thick. *Aristotelian Society Supplementary Volume* 91 (1):197-220.

Roberts, Debbie (2018). Why Believe in Normative Supervenience? *Oxford Studies in Metaethics* 13.

Robinson, Luke (2011). Moral Principles As Moral Dispositions. *Philosophical Studies* 156 (2):289-309.

Rosen, Gideon (2010). Metaphysical Dependence: Grounding and Reduction. In Bob Hale & Aviv Hoffmann (eds.), *Modality: Metaphysics, Logic, and Epistemology*. Oxford University Press. pp. 109-36.

Rosen, Gideon (2017a). What is a Moral Law? *Oxford Studies in Metaethics* 12.

Rosen, Gideon (2017b). Ground by Law. *Philosophical Issues* 27 (1):279-301.

Rosen, Gideon. (2020). What is normative necessity? In M. Dumitru (Ed.), Metaphysics, Meaning and Modality: Themes from Kit Fine (pp. 205-233). Oxford: Oxford University Press.

Rosen, Gideon. (2021). The modal status of moral principles, Oxford Studies in Metaethics, 16, 257-279. doi: 10.1093/oso/9780192897466.003.0011

Ross, David (1939). *Foundations of Ethics.* Oxford University Press.

Scanlon, T. M. (2014). *Being Realistic About Reasons.* Oxford University Press.

Scanlon, T. M. (2017). Normative realism and ontology: reply to Clarke-Doane, Rosen, and Enoch and McPherson. *Canadian Journal of Philosophy* 47 (6):877-897.

Schaffer, Jonathan (2016). It is the Business of Laws to Govern. *Dialectica* 70 (4):577-588.

Schaffer, Jonathan (2018). Laws for Metaphysical Explanation. *Royal Institute of Philosophy Supplement* 82:1-22.

Schroeder, Mark (2005). Cudworth and Normative Explanations. *Journal of Ethics and Social Philosophy* 1 (3):1-28.

Sider, Theodore (2018). Ground grounded. *Philosophical Studies* 177 (3):747-767.

Strandberg, Caj (2008). Particularism and Supervenience. In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics.* Oxford University Press.

Sturgeon, Nicholas L. (2009). Doubts about the Supervenience of the Evaluative. In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics.* pp. 53-92.

Väyrynen, Pekka (2009). A Theory of Hedged Moral Principles. *Oxford Studies in Metaethics* 4:91-132.

Väyrynen, Pekka (2018). Reasons and Moral Principles. In Daniel Star (ed.), *The Oxford Handbook of Reasons and Normativity*. pp. 839-61.

Väyrynen, Pekka (2021). Against Moral Contingentism. *Thought: A Journal of Philosophy*.

Wallner, Michael (2018). The ground of ground, essence, and explanation. *Synthese* 198 (Suppl 6):1257-1277.

Walton, Kendall Lewis (2006). On the (so-called) puzzle of imaginative resistance. In Shaun Nichols (ed.), *The Architecture of the Imagination*. Oxford University Press. pp. 137-148.

Walton, Kendall Lewis (1994/2015). Morals in Fiction and Fictional Morality (I). *Proceedings of the Aristotelian Society* 68:27-50.

Wayne, Andrew (2018). Explanatory integration. *European Journal for Philosophy of Science* 8 (3):347-365.

Wilson, Jessica M. (2014). No Work for a Theory of Grounding. *Inquiry: An Interdisciplinary Journal of Philosophy* 57 (5-6):535-579.

Wilson, Jessica M. (2015). Hume's Dictum and metaphysical modality: Lewis's combinatorialism. In Barry Loewer & Jonathan Schaffer (eds.), *The Blackwell Companion to David Lewis*. Blackwell. pp. 138-158.

Wong, David B. (2006). *Natural Moralities: A Defense of Pluralistic Relativism*. Oxford University Press.

Yablo, Stephen (1993). Is conceivability a guide to possibility? *Philosophy and Phenomenological Research* 53 (1):1-42.

Chapter 5 | How Moral Generalizations Explain

Abstract: I argue that moral principles, construed as moral generalizations, can be genuinely explanatory. First, I clarify the metaphysical status of moral generalizations. Moral generalizations are summaries of particular moral phenomena, and they are non-substantive in nature. Then, I present a recent and important challenge by Berker (2018) according to which moral generalizations are explanatorily redundant. I argue that Berker's argument rests on the assumption that the non-idleness of the relevant explanantia is a *requirement* on explanation. In response, I present and defend a different explanatory relevance constraint that is based on the idea of unification. I also show that this constraint follows from the independently plausible thesis of moral explanatory pluralism. Finally, I appeal to the same theoretical resources to tackle a revenge problem in the form of Rosen's (2017) circularity challenge.

1. Introduction

In this paper I defend a novel account of moral principles as standards.[1] I argue that moral principles, construed as moral generalizations, can be genuinely explanatory. Moral generalizations are *lists* of moral phenomena (I will talk about the metaphysics

---

[1] Moral principles can be construed as either guides or standards (or both). A moral principle can be understood as a guide once it is individuated in terms of its practical function to guide action. Moral principles as standards are said to provide an explanation for why particular moral facts obtain (McKeever & Ridge 2006; Väyrynen 2018).

and logical form of moral generalizations in the next section). The core principle of consequentialism, so construed, would consist in all and only those acts that are wrong in virtue of their non-moral constituents (which presumably contribute towards some specified consequence). In this sense, moral principles *just are* the sum of their particular instances.

To say that moral principles are mere generalizations is to say that they are summaries of moral phenomena. But if a moral principle just is a sum of its actual and possible instances, how can it be *explanatory* of those instances? Selim Berker (2018) has recently argued that moral generalizations are, by their nature, explanatorily idle.[2] Based on plausible assumptions about the form of moral principles, moral generalizations specify the full grounds of particular moral facts. But then it seems that particular moral facts are fully explained in terms of their full grounds, thus rendering the relevant moral principle explanatorily obsolete. Secondly, even if somehow moral generalizations can actively explain particular moral facts, there is a revenge problem lurking in the background (Rosen 2017). It is plausible that moral generalizations are explained by their instances. So, on pain of circularity, we cannot appeal to those generalizations to explain those very instances.

The view that moral principles are generalizations in the above sense might seem striking to some. Still, it is worth defending for at least two reasons. First, it would follow from a broadly neo-Humean metaphysical perspective. A neo-Humean account of moral principles would be the view that moral principles are nothing over and above particular entities figuring at the Humean mosaic (roughly, a base of highly specific, non-modal, categorical properties). Neo-Humeanism is a strong contender in the literature on scientific laws and a powerful metaphysical framework overall. Its successes can be found in multiple areas in philosophy, ranging from metaphysics (Lewis 1983)

---

[2] A version of the same type of worry can be also found in Robinson (2007: sec. 3; 2011: 295-6).

to moral metasemantics (Dunaway & McPherson 2016). At the very least, neo-Humeanism about moral principles should be understood as a live conceptual possibility that is worth exploring.[3] Secondly, the view that moral principles just are the sum of their instances is superior to rival views in terms of ontological parsimony. Views which understand moral principles as something *distinct* from their instances need to further specify the nature of that separate entity.

In fact, I should note that the view that moral principles are generalizations shouldn't be considered as *particularly* controversial. Understanding moral principles as generalizations is a metaphysical theory about the status of moral principles. But *any* metaphysical theory about the status of moral principles is going to involve some level of controversy since metaphysical views *in general* do not immediately follow from first-order discourse. To illustrate, Robinson's (2011) view about moral principles appeals to essentialist facts about moral dispositions. But no facts about essences can be directly read off from the way moral discourse usually operates. In this sense, first-order normative discourse, at least at a surface level, is metaphysically neutral.[4]

Why would one want to defend the genuine *explanatoriness* of moral generalizations? There are two main reasons. First, there is a consensus in first-order normative ethics that moral principles are explanatorily important.[5] Explanations of

---

[3] To my knowledge, a full defense of this sort of view hasn't appeared in print yet. Robinson (2007; 2011) and Fogal & Risberg (2020) briefly consider it. Berker's (2018: sec. 9) hyperintensionalist Humeanism is significantly different from the view I have in mind mainly because he accepts that moral generalizations are not explanatory (i.e. not *explanation-serving* in his own terminology) (2018: 2).

[4] The same applies concerning the metaphysics of scientific laws (Carroll 2020: sec. 1).

[5] For references, see Robinson (2011: 290; 294 fn. 7) and Walden (2016). Fogal & Risberg (2020) take the genuine explanatoriness of moral principles to be a key datum that every theory of moral explanation should accommodate. See also what Schroeder calls the 'Standard Model' of normative explanation

particular moral facts somehow involve an appeal to some general moral fact. A particular act is either good or bad partially because a specific moral principle holds. For example, a particular killing is morally wrong partially because this act minimizes utility and, in general, wrongness consists in the lowering of utility. If moral generalizations are a genuine kind of moral principles, then they should be able to accommodate this datum.[6] Secondly, *scientific* laws are also routinely used to explain particular scientific facts. Many theories of scientific explanation take this to be uncontroversial.[7] Insofar as the analogy between scientific and moral laws is warranted, we should expect moral principles to work in a similar way.

After presenting the metaphysics and logical form of moral generalizations (section 2), I will present Berker's challenge against their putative explanatoriness (sections 3 and 4). One of Berker's key assumptions is that a requirement on explanation is that a putative explanans should not involve any idle parts (i.e. parts which, on their own, do not contribute towards making the explanandum-fact the case). In response, I propose an alternative constraint according to which P can explain Q via subsumption (i.e. Q is explained in terms of its inclusion in a larger whole) (section 5). This constraint is further vindicated by showing that it is a part of larger

---

(2005: 24). I also take it that the genuine explanatoriness of moral principles is at least *one* of the things that are at stake in the debate between moral generalists and moral particularists.

[6] The claim that moral principles are explanatory gets indirect support from a plausible assumption about metaphysical grounding. Many accounts of grounding take general facts (typically construed as laws) to figure in the explanation of particular facts in some way or another: either as grounds themselves (e.g. Bader 2017) or as meta-facts (A. Wilson 2020; cf. Berker 2018: sec. 6).

[7] For example, Railton (1978), Kitcher (1989), Salmon (2002), Woodward & Hitchcock (2003). Of course, there are exceptions (Scriven 1962). Also, appeals to nomic regularities can be found in modern metaphysics. For example, see Schaffer (2016; 2017) for a tripartite structure of explanation according to which both scientific and metaphysical laws can play the role of *linking* the explanantia-facts with the explanandum-fact.

theory on explanation: moral explanatory pluralism (defended in section 6). I will also argue that this theory manages to capture the main intuition that made Berker's challenge plausible to begin with. Finally, in section 7, I will appeal to moral explanatory pluralism to address Rosen's revenge problem.

2. On Moral Generalizations

I understand moral generalizations as lists of particular moral phenomena.[8] In this sense, views that take moral principles to be mere generalizations are *non-substantivist* in nature: they take moral principles to be nothing over and above the instances that figure in them. As mentioned, the consequentialist principle, so construed, would consist in all and only those acts that are morally good in virtue of their non-moral constituents.[9]

Consider moral principles in the abstract. As any kind of lawlike generalization, moral principles have the following form:

(*) (Necessarily) (For every x) (Fx fully because Gx)[10]

---

[8] Moral generalizations are not *generic* generalizations (cf. Carlson 1977). Generic generalizations stay silent concerning the number of instances that are involved in the relevant phenomenon (for this reason, generics usually have exceptions). However, moral generalizations *can* accommodate the view that moral principles involve a ceteris paribus clause (see fn. 10).

[9] Of course, moral principles are not *literally* nothing over and above their instances. Rather, the *referent* of a moral principle is nothing over and above the proper parts of that entity (this is what Robinson [2011] calls the *genetic* sense of moral principles). For simplicity, I will sometimes gloss over this distinction.

[10] This is a simplification. As noted, it is plausible that the form of moral principles involves separate clauses for enablers/disablers, right/bad-makers, and, perhaps, ceteris paribus clauses (as per Väyrynen's [2009] 'hedged' view) (cf. Ridge & McKeever 2006). It could also turn out that moral principles are not

If moral principles are generalizations, then (*) fully reduces to the following formulation:

> **(List)** (Necessarily) [(Fx_1 fully because Gx_1) and (Fx_2 fully because Gx_2) and … (Fx_n fully because Gx_n)][11]

This would not be the case if moral principles referred to entities that are wholly distinct from their instances. A non-moral substantivist analogue might help to illustrate. Maudlin's 2007 account takes scientific laws to be metaphysically primitive and irreducible to the instances they govern.[12] Similarly, Rosen (2017) has recently argued for an account according to which moral principles are relations between universal-like entities.

I take it that if there's anything philosophically interesting about moral generalizations it's their non-substantivity which entails that (*) fully reduces to (List). Note that, for someone who accepts the existence of moral principles, there is nothing particularly controversial in (*). The existence of the necessity-operator is trivial.

---

universally quantified conditionals to begin with (e.g. they could be 'defeasible generalizations' as per Lance & Little [2004]) (cf. Scanlon 2014). But such complications are orthogonal to the point I am trying to illustrate.

[11] It might be objected that, regardless of one's view concerning the metaphysics of moral principles, (*) *trivially* entails (List) (cf. Schaffer 2017: sec. 4.1.). In response, I would say that even if it is true that (*) can be *translated* into (List), it does not follow that (List) would fully express the content of the relevant view. If, following Rosen (2017), one takes moral principles to be relations between universal-like entities, then it would be false to claim that such a view can be adequately expressed in terms of (List). In this sense, when I claim that, according to the moral-generalizations view, (*) fully reduces to (List), I mean that (List) fully captures the *content* of that view.

[12] I want to resist cashing out the difference between substantivist and non-substantivist views in terms of 'governing' given that the term is notoriously obscure and hasn't received a lot of systematic treatment in the literature (although see Wilsch [2021]).

Everyone accepts that, if there are moral principles, they hold with some kind of necessity. Summaries of moral phenomena can hold with metaphysical, normative (Fine 2002; Rosen 2017) or even nomological necessity (Dretske 1977). Nor is the because-operator particularly suspect. Intensionalists can interpret the because-operator as a material conditional or a biconditional, whereas hyperintensionalists can take "fully because" to generate a context where the substitution of intensionally equivalent entities doesn't necessarily preserve its overall truth-value.[13]

Even though Berker adopts the latter option (I will say more about this in the next section), what is important at this point is that the differentiating feature of moral generalizations is their non-substantivity, not the way the operator figuring in (*) is understood.[14] As mentioned, there is a sense in which, *everyone* should accept that moral principles are generalizations. Non-substantivists claim that a moral principle is a moral generalization, and *that's it*.[15] The substantivist, rather, posits an entity *on top* of that moral generalization (e.g. essences, relations between universals, etc.). This concludes my presentation of the non-substantivist view. In the next section I will focus on Berker's redundancy challenge against it.

---

[13] Concerning the intensionalist framework, I take it that fundamental or 'pure' moral principles would involve a biconditional, whereas a derivative, or 'mixed', principle would involve the material conditional. Value pluralists can either accept multiple principles involving a conditional, or one biconditional involving separate disjuncts for every type of value. Hyperintensionalists have at their disposal a variety of metaphysical determination relations (Rosen 2010; Wilson 2014).

[14] It should not come as a surprise that the logical form of moral principles is significantly detached from their underlying metaphysics. A similar point is also plausibly true in the context of scientific laws (see, e.g., Lloyd [1955]; Dretske [1977]; Friend [2016]).

[15] As Rosen (2010: 120-1) notes, the explanation of such generalizations would plausibly also involve a *totality* fact (i.e. a fact stating that the relevant instances involve such-and-such features and *that's it*). I will ignore this complication for presentation purposes.

3. Berker's Redundancy Challenge

Berker (2018) has recently raised important concerns about the many different ways in which moral principles could be explanatory. I will focus on his challenge to the explanatoriness of moral generalizations (2018: sec. 3). Consider the following schema:

(Simple) M_1 (a particular moral fact) is partially explained by N_1 (a particular non-moral fact) and partially explained by G (moral generalization).

Berker claims that (Simple) is false. This is because moral generalizations (G) are explanatorily redundant. First, he argues that the only plausible conception of G is hyperintensional: moral principles generate hyperintensional contexts. Assume that a moral principle states a connection between a general moral fact (M) and a general non-moral fact (N) (i.e. a fact whose particular instances exemplify it). For example, take the principle according to which the maximization of goodness consists in the maximization of pleasure. This relation shouldn't be understood in purely modal terms: it is not simply the case that goodness (M) modally covaries with pleasure (N). Rather, M obtains fully *in virtue* of N.[16] The 'in virtue of' locution could be plausibly underwritten by a metaphysical determination relation like grounding (Rosen 2010) or some other, similar, relation (Wilson 2014). For the purposes of illustration, the exact details of that relation are not important (I will use 'grounds' in a neutral sense unless stated otherwise).

What is important is Berker's follow-up assumption that full grounding claims correspond to full explanations. If M-facts are fully grounded by N-facts, then the same

---

[16] It is not so clear why the relevant 'in-virtue-of' claim needs to be 'full' in order for the hyperintensionality of G to be accommodated (cf. Duncan et al. 2017). Still, for dialectical purposes I will grant this move.

holds about particular instances of those facts. For example, if the wrongness of killing holds fully in virtue of the descriptive features of such acts, then the wrongness of particular killings also holds fully in virtue of *their* descriptive features. In this sense, the explanation of the wrongness of a particular killing in terms of its descriptive features corresponds to a full explanation of the form "The wrongness of this particular killing is fully explained by its descriptive features".

If this is the case, then (Simple) is false. According to (Simple), particular moral facts (like M_1) are partially explained by the appropriate moral generalization (G). But if G partly explains M_1 then M_1 is not *fully* explained by its non-moral features (N_1). For this reason, Berker concludes that G is explanatorily redundant.

4. On Redundancy

What is the notion of redundancy at play? The notion of redundancy seems to be antithetical to the notion of relevance: a fact is explanatorily relevant iff it is not explanatorily redundant. Grounding theorists typically appeal to a non-monotonicity clause: a clause that makes sure that a fact counts as a proper ground for another fact only if it is appropriately relevant (Litland 2013).

> **(Non-Monotonicity)** If A explains B then it is not necessarily the case that the conjunction of A and C explains B.

(Non-Monotonicity), in this form, is not very illuminating. It simply states that a ground must be explanatorily relevant to the fact that it grounds. But what, according to Berker, is this criterion? Here is what he has to say (2018: 8; modified):

> [...] [G] can only serve as a partial [explanation] of [M_1] if it is a redundant partial [explanation] of [M_1] [...] [I]t is implausible that moral principles play

their distinctive explanatory role in a redundant way: how can that role be distinctive if it is not needed?

Berker does not say more about this notion of redundancy except that something is not redundant when it is *needed.* The most direct way to make sense of this claim is in the context of a theory of explanation. But Berker does not explicitly provide one. Still, an important clue is given in the form of how the relation between grounding and explanation is cashed out. Recall that full grounding facts are supposed to correspond to full explanations. This means that explanatory redundancy collapses into grounding redundancy.

What does it mean to say that a ground is redundant? Again, a ground is redundant, according to Berker, if it not *needed* for the target-fact to obtain. A plausible way to read this claim is to say that a ground P is redundant for a target-fact Q when Q obtains *regardless* of whether P obtains. What is the notion of ground at play here? Even under minimal assumptions about grounding, everyone accepts that grounds contribute towards making their target-facts obtain (e.g. Litland 2013). In this sense, a full ground is *sufficient* for its target-fact.

There are many different ways in which a ground can have a contributory role. On the one end of the spectrum, a ground can be *essentially* contributory in the sense that its contribution is particularly important (in some sense to be specified) for the fact it grounds. For example, one could claim that many factors contributed towards some organism having phenotype P, but the most important factor for having P is having gene Q. On the other end, a ground can be *non-essentially* contributory in the sense that it is not *needed* for its target-fact to obtain; many other facts could figure in its place without altering the result. For example, my injury was essentially caused by the fact that I was hit by an assailant, but it was *non-essentially* caused by the fact that I was hit

by a *blonde* assailant.[17] Or, there might be cases where a factor is not needed in the sense that it is *trivial*: the fact that I exist can be non-essentially explained by the fact that the world exists, and I am a component of that world.

How do these distinctions apply to Berker's argument? A very plausible way to read Berker's redundancy-claim is to adopt the distinction between essential and non-essential grounds. Redundant grounds are non-essential grounds and non-redundant grounds are essential grounds.[18]. Berker is correct in saying that, in the context of (Simple), G is less essential than N_1.[19] N_1 is more important than G in the following sense: even though both facts, strictly speaking, fully contribute towards making M_1 the case, G has many *idle parts*.

To see this, consider G's form. As mentioned in section 2, G fully reduces to (List) (i.e. the list of every instance of G). (List), in turn, refers to a composite object: the fusion of every instance of M holding in virtue of a corresponding instance of N:

> **(Fusion)** [(N_1 fully grounds M_1) and (N_2 fully grounds M_2) and … (Nn fully grounds Mn)]

Appealing to (Fusion) makes salient the way in which not every instance of G actively grounds M_1. A moral generalization about the wrongness of killing is simply a summary

---

[17] What I have in mind in this example is something like Woods' notion of *vacuous* grounding (2018).

[18] Naturally, there are many possible ways to understand essential grounds which I am not going to rehearse here. I take the distinction between essential and non-essential grounding to be intuitive and independently plausible. My argument relies only on a general characterization of this distinction. More fleshed out accounts have been offered by Krämer & Roski (2017), Woods (2018), and Makin (2019). For general discussion on the varieties of 'non-causally making it the case' see Wygoda Cohen (2020: sec. 2).

[19] Note that Berker never explicitly claims that G has no grounding role. Rather, he claims that its role is *trivial* (2018: 6).

of every particular killing (actual or possible) alongside the non-moral features that brought them about. In that respect, G involves N_1 which, by itself, is a full ground for M_1, while also involving the grounds of *other* instances.

Even though G, on its own, fully contributes towards M_1, it does so in a non-essential way: other facts having N_1 as their part would also bring about M_1. In that respect, there is nothing particularly *special* about G.[20] The explanatory intuition that Berker wants to capture is clear:

> **(Non-Idleness)** At least some explanations involve identifying the non-idle entities that brough the explanandum-fact into existence.

Berker takes (Non-Idleness) to be a *requirement* on explanations. This means that explanations should only cite those factors that contributed, in a more-or-less non-vacuous way, towards bringing the relevant explanandum-fact into existence. In other words, explanations must be specific and should avoid involving idle factors (such as many of the parts of Fusion).

5. Subsumption Explanations

There is certainly something true about (Non-Idleness). So-called 'mechanistic' explanations focus on the fine-grained details that lead from a mechanism's set-up to its termination conditions (Machamer et al. 2000). For example, the mechanistic

---

[20] There are other ways to argue against G being an explanation for M_1 which I do not have the space to fully address here. One could argue that G is not even a ground (whether full or partial) for M_1 because grounding is irreflexive (and G has M_1 as its part). Analogously, causation is said to only relate logically distinct entities. Still, it should be noted that both constraints have been challenged in the literature (some argue that causation can be immanent [Mellor 1995], and grounding can be reflexive [Jenkins 2011; Woods 2018]).

explanation for a given event of protein synthesis involves a pre-initiation sequence (roughly, an mRNA and a tRNA coming together) and a termination phase (roughly, the mRNA degrades so the relevant nucleotides can be used elsewhere for other processes). It is not a coincidence that many accounts of mechanistic explanation do not include laws in the relevant set of explanantia (cf. Cartwright et al. 2020). Such explanations seem to get their explanatory force from the fact that they involve the components, and only those components, that actively brought about the relevant phenomenon. Berker's notion of redundancy seems importantly analogous.

In response, I want to challenge the inference from the fact that G non-essentially grounds M_1 to the claim that G cannot be *explanatory* towards M_1. Specifically, even though "G explains M_1" is not an explanation in terms of (Non-Idleness), it does not follow that it cannot be an explanation of another kind. Indeed, as I will argue, it is independently plausible that (Non-idleness) is only *one* of the virtues that an explanation can exhibit.

Consider what I will call a *subsumption* explanation. Such explanations explain a particular phenomenon by showing that it is a part of a larger whole. For example, it is plausible that one could explain the wrongness of a particular killing in virtue of the fact that all killings have such and such features, thus showing that the explanandum-fact is a part of a larger pattern.[21] These explanations aim for unification: they get their explanatory force by showing how the relevant phenomena fit together.

> **(Unification)** At least some explanations proceed by unifying. Such explanations involve the explanation of a particular phenomenon by showing how that phenomenon is a part of a larger pattern.

---

[21] In the philosophy of science literature, the two classics are Friedman (1974) and Kitcher (1989).

In this sense, (Simple) seems to be compatible with (Unification) and should thus be regarded as a legitimate form of explanation.

Are there any plausible objections against (Unification)? First, there might be some suspicion concerning the *metaphysics* of subsumption-explanations. It is usually suggested that explanations are underwritten by determination relations. If P explains Q, then there must be some determination relation linking the entities that stand for P and Q. But, on the face of it, there is no determination relation that is directed from a moral principle towards its instance.[22]

I have two responses. First, we can challenge the assumption that a general entity (in this case, a moral principle) cannot determine a particular instance. 'Top-down' grounding, though metaphysically extravagant, is not unheard-of (e.g. Schaffer 2010). There is also the option of countenancing a distinct relation of *subsumption* (cf. Kim 1973; Strevens 2008: ch. 1).

Secondly, it is not clear what motivates the requirement that *every* explanation needs to be underwritten by a determination relation to begin with.[23] Presumably, this requirement is based on the intuition that explanation needs to be *objective* as per the

---

[22] Baker's (2021) recent suggestion builds on a similar notion of subsumption. But his account is significantly different from the one presented in this paper. For one, Baker is appealing to a *pragmatic* theory of explanation in order to provide an expressivist-friendly construal of the 'because' locution (cf. 2021: 56). Specifically, he adopts an understanding of moral principles as *guides* (instead of *standards,* as I do) to suggest that explanations via subsumption facilitate the function of social coordination (Baker 2021: 64, 70) (although see fn. 25). Also, he takes the relata of explanation to be *imperatives*, rather representational entities standing for objective features of reality (2021: 59). For this reason, Baker can stay agnostic about whether there is a determination relation underwriting these explanations (2021: 58, 76) (he also doesn't have to deal with the circularity objection as per sec. 7).

[23] Note that this formulation is neutral towards the unionist/separatist debate concerning the relation between explanation and metaphysical determination (cf. Raven 2015).

ontic conception of explanation. According to this conception, genuine explanations are objective by involving entities which exist in a stance-independent way (Craver 2009). Consider a type of explanation which does not meet that requirement. Take, for example, a set of explanantia which involve abstractions in the form of idealizations (e.g. Strevens 2008: ch. 8). Such abstractions might be epistemically useful but, according to the ontic conception, they are not genuinely explanatory since they are not objective features of the world.

But even if fictional explanantia are objectionable, moral generalizations like G are not like fictions. As noted, entities like (Fusion) are concrete, fully objective entities that exist stance-independently. According to the ontic conception, explanations need to involve nothing but the objective features of the world. But there are *many* ways in which this requirement can be met. One way is by appealing to an explanans with idle parts like G. Another is to only appeal to those components of G that are essential grounds for the relevant explanandum-fact. If the ontic conception is simply a constraint concerning the stance-independent existence of the explanantia-facts, then (Non-Idleness) is not the only explanatory virtue that can accommodate it.[24]

Finally, one could worry about the nature of the *explanandum* involved in subsumption explanations. An explanation constrained by (Non-Idleness) would explain a particular moral instance (Q) in terms of (say) its non-moral features (P). But it could be objected that a subsumption explanation in terms of a moral principle is not

---

[24] According to a stricter interpretation of the ontic conception, explanations *just are* worldly, stance-independent, facts (Bokulitch 2016; cf. Craver 2019). But there are serious problems with this account as it is not clear how any explanatory relevance constraints are to be implemented. As Wright & Dingmar (2018) convincingly argue, the ontic/epistemic debate shouldn't be understood as a debate concerning whether the success-conditions of an explanation have an epistemic component. Rather, the issue should concern which epistemic virtue should be *prioritized*.

an explanation of Q. Rather, it is an explanation of Q's *place* in a larger pattern (in this case, the appropriate moral principle). If this is the case, then subsumption explanations cannot accommodate (Simple).

I want to resist the suggestion that there is an explanandum shift. An explanation constrained by (Unification) explains (say) a particular moral instance *in terms* of its place in a moral principle. In this sense, (Unification) concerns the *way* a given explanandum-fact is explained. Explaining an instance's *place* under a larger pattern seems importantly different. For example, explaining the place of a particular moral instance in terms of the principle of utility would presumably involve identifying the features according to which that particular instance falls within the scope of that principle. But this is surely different from explaining that moral instance *in terms* of the principle of utility.

6. Moral Explanatory Pluralism

One can have it both ways by combining (Non-Idleness) with (Unification). This can be done by adopting a pluralistic theory of moral explanation which individuates the relevance constraints of every explanation in terms of what is decided by the appropriate research project. In this sense, an explanans is redundant only in the context of an explanatory project:

> **(Pluralism)** P explains Q, iff, P is an appropriate explanans for Q given the aims and goals of the research project in which the explanation "P explains Q" figures.

Berker is, indeed, correct in claiming that moral generalizations are explanatorily redundant if one assumes that (Non-Idleness) is a necessary condition for the success of an explanation: the explanandum-fact (a particular moral fact) is fully explained (in the

relevant sense), without residue, in virtue of those particular facts that essentially contributed towards making that fact the case.

What research aim could (Non-Idleness) plausibly serve? One central reason concerns the fact that (Non-Idleness)-explanations demonstrate a highly specific level of *grain*. Their function is to showcase in great detail what brought about a given explanandum-fact: the moral wrongness of a particular act was non-causally brought about by such-and-such particular facts. Analogously, an event involving the synthesis of a protein is explained by those fine-grained components that contributed towards it.

Still, as already argued, showing that some explanations are plausibly constrained by (Non-Idleness) does not entail that *every* explanation should be as such. What about (Unification)? One plausible aim which (Unification)-explanations can serve is predictive. Note that even though explanation and prediction are fundamentally different enterprises, some explanations are more useful than others in terms of facilitating some predictive goals. For example, having an actual-sequence explanation of a given moral fact (i.e. in terms of its exact causal history or in terms of its exact constitutive profile) does not provide enough information that would allow us to predict what would happen if similar circumstances were to obtain. On the other hand, knowing that, *in general*, wrongness is brought about by such-and-such non-moral circumstances, is much more useful.[25] Situating a particular grounding-fact ([N_1 grounds M_1]) into a larger pattern allows us to predict what would happen if similar but uninstantiated instances of that pattern were to obtain.

---

[25] The fact that moral principles have the function to predict seems to be more connected to their function as guides rather than standards. But this is not necessarily the case. Moral principles can be understood as standards which explain, and, in turn, one can understand prediction as a *type* of explanation. So, one can appeal to moral principles as standards to predict, but only *derivatively*, in virtue of explanation *itself* having a predictive function (in some of its instances).

Another aim that (Unification) can serve concerns our *understanding* of moral goodness itself. Consider monist attempts to bring together various moral principles under the same notion of goodness. Mill famously subsumes lower and higher-level types of pleasure under his utilitarian principle (Brink 2013). Ross initially lists seven prima facie duties, but later suggests that some of them (beneficence, self-improvement, and justice) can be understood as a *single* duty concerning intrinsic goodness (Ross 1939). These examples are controversial, but they work as illustrations. In both cases, regardless of the relevant consequences on first-order normative discourse, by subsuming less fundamental principles under more fundamental ones, it can be said that our understanding of moral goodness increases significantly. After all, we are not only interested in what we should morally do, but also in what moral goodness consists of. (Unification) works towards that direction.

What motivates (Pluralism) about the moral domain? The main reason involves evidence from the ways our moral explanatory practices operate.[26] The plurality of our explanatory practices is reflected in the different ways moral explanation is presented in the relevant literature. As mentioned in the introduction, a paradigmatic type of moral explanation is exemplified by what I have called (Simple): a moral fact is partially explained by a general moral principle. Others claim that there are cases where moral

---

[26] Another type of reason in favor of (Pluralism) concerns the plausibility of explanatory pluralism proper. First, explanatory pluralism is a relatively widespread thesis when it comes to the special sciences. Potochnik (2015), for example, argues that there can be multiple different explanations for a given biological phenomenon. It is a part of the relevant researcher's agenda to select a segment of the totality of contributing factors that brought about the relevant explanandum-fact. Secondly, explanatory pluralism can be traced back to an uncontroversial thesis about perception. According to Longino (2006) "all representations are partial in that any representation must select a limited number of aspects of a phenomenon (else it would not represent, but duplicate)." Given that there are many ways in which a phenomenon can be represented, it is plausible that there are many different kinds of explanation.

principles have no role whatsoever. Berker's view, as mentioned in the previous section, is one example. Another might be a particularist account of moral explanation. Particularism is generally understood as a global thesis about the moral domain (roughly, the thesis that there are no true moral principles). But we can imagine a more localized version where there are ethical truths about a particular area of moral discourse, without there being any general principles (for that area). In that scenario, particular moral instances wouldn't be explained by moral principles (since, ex hypothesi, there are none). Finally, there is historic precedent involving moral principles explaining other principles: Mill's principle pertaining to liberal rights (as, plausibly, every other right he proposes) can be traced back to his principle of harm (for discussion see Brink [2013]) (cf. Baker 2021: 76).

Some of these modes of explanation are, perhaps, controversial. But they all share a degree of initial plausibility. (Pluralism) has a clear explanation for this: they are *all* legitimate modes of explanation depending on the agenda of the research program in which they figure.

7. The Circularity Problem

(Pluralism) can successfully accommodate at least two different kinds of explanations: one falling under (Non-Idleness) and one falling under (Unification). There could be others. Still, in this section I want to focus on a problem that potentially arises once we accept these two different types of explanation: the circularity problem (Rosen 2017).

Consider again the moral generalization that killing is wrong in virtue of its non-moral features. This generalization has certain parts: its instances. In this respect, in

a certain context, we can explain this generalization, *constitutively*, in terms of its instances:

> **(GEN-IN)** A generalization P is partially explained by its instance Q.

On the other hand, in another context, we might explain a particular instance of that generalization (for example, the wrongness of an act) in terms of the whole under which it is subsumed:

> **(IN-GEN)** An instance Q is partially explained by generalization P.

Also, it is plausible that explanations are constrained by the following formal features:

> **(Transitivity)** If A explains B, and B explains C, then A explains C.

> **(Irreflexivity)** P cannot explain itself.

The circularity problem builds on the fact that these theses cannot be all true at the same time. If the generalizations are explained by their instances, then one cannot appeal to these generalizations to explain those very instances. Doing so would violate either (Irreflexivity), or (Transitivity), or both. Also, denying either (GEN-IN) or (IN-GEN) would go against (Pluralism).

7.1. Against (Transitivity)

Is (Transitivity) true? Many philosophers think so. But, as noted in the previous section, explanantia must be explanatorily relevant for the explanandum, where explanatory relevance can be cashed out in many different ways depending on the explanatory project under which we operate. Still, there seems to be something true about the idea that certain explanations can *chain* in certain contexts. Is there a way to reconcile something in the vicinity of (Transitivity) with the pluralist picture I have been proposing?

The natural way to proceed is to make (Transitivity) more fine-grained. As already noted, there are many ways to cash out the notion of explanatory relevance.

(IN-GEN) can be understood as employing a constraint like (Unification). This constraint makes sure that the information that is eventually included in the explanation contributes towards the epistemic goal of unifying the relevant phenomena. It is obvious that this constraint does not exclude transitive explanations. For example, a more fleshed-out moral version of (IN-GEN) might be the following:

> **(IN-GEN*)** The fact that a particular killing is morally wrong is partially explained by the fact that the principle "you ought not to kill" holds. Now assume, for simplicity, that every killing minimizes utility. The fact that the principle "you ought not to kill" holds is, in turn, partially explained by the fact that the principle of utility is true. Thus, the fact that this particular killing is morally wrong is, transitively, explained by the fact that the principle of utility holds.

I take (IN-GEN*) to be uncontroversial insofar as we can make sense of an explanation that proceeds by unifying. In this sense, I take it that derivative principles can be explained and do, themselves, explanatory work. The wrongness of a particular killing is first incorporated under a larger regularity ("you ought not to kill"), which, in turn, is subsumed under an even more general regularity (the principle of utility).

> Both explanations figuring in (IN-GEN*) have the same *aim*. (IN-GEN*), in this sense, provides a hint towards the right view of transitivity.

> **(Transitivity*)** If A explains$_1$ B and B explains$_2$ C then A explains C, iff, both explanations accommodate the same explanatory-relevance constraint.

Now consider the two following explanations:

> **(M)** [Killing is morally wrong] partially explains (via subsumption) [This particular killing is morally wrong].

> **(M+)** [This particular killing is morally wrong] partially explains (constitutively) [Killing is morally wrong], and, [Killing is morally wrong] partially explains (via

subsumption) [This particular killing is morally wrong]. Thus, [This particular killing is morally wrong] partially explains itself.

(M) is unobjectionable whereas (M+) is the circular result that the circularity problem warns us about.

At this point we should immediately notice that (M+) is (M) plus an extra component. (M+) involves the explanation of a particular killing by subsuming it under a moral regularity, but then it proceeds by explaining that very principle constitutively. But this additional step can be resisted. Knowing the constituents of a principle does not help the unificationist goal of (M). In this sense, (M+) is inferior to (M) since it introduces irrelevant information. Thus, by (Transitivity*) we can avoid the circularity worry without giving up a sufficiently substantive version of the transitivity principle.

One assumption in the previous argument is that (M+) has the same aim as (M). But this is not necessarily so. One might claim that there are contexts where explanations like (M+) are *exactly* what we should be looking for. It is obvious that we can manufacture an explanatory project with aims S, where it is stipulated that S requires (IN-GEN) and (GEN-IN) to be linked. But to my mind these cases are atypical for they require an implausible and highly artificial explanatory project. Why would someone be interested in an explanation like (M+)? In this sense, I am happy to grant that there are atypical cases where explanations are circular.

Note how (Transitivity*) compares to other amended versions of the transitivity principle. It has been suggested that the transitivity of explanation breaks down when the relevant explanations are not backed by the same determination relation (Hicks & van Elswyk 2014). Such an account would also avoid the circularity objection. (GEN-IN) is a constitutive explanation: it explains a regularity in terms of its constituents. (IN-GEN), on the other hand, explains particular instances by subsuming them under a larger pattern. In this sense, (GEN-IN) and (IN-GEN) are underwritten by different relations: the former appeals to constitution, and the latter to subsumption.

But this cannot work. There are plausible examples of successful 'hybrid explanations' (Lange 2018) which appeal to many different explanatory relations. Here's an example: Nucleotides *compose* Genes and Genes *functionally realize* DNA (Trogdon 2018). But it is plausible that a certain arrangement of nucleotides explains (transitively) the occurrence of a DNA-molecule.

It has also been suggested that transitivity breaks down when the relevant explanations are employed by different *disciplines.* Consider the previous example. Presumably, nucleotides explain (transitively) the occurrence of a DNA-molecule because the constitutive explanation of genes in terms of nucleotides and the realization-explanation of DNA-occurrence in terms of genes both figure in the same *discipline* (presumably biology).[27]

It is unclear whether this strategy can be successfully applied to the moral case. For (GEN-IN) not to be linked with (IN-GEN) it must be the case that these explanations are employed by different disciplines. Presumably, (IN-GEN) is a matter of moral metaphysics (since it concerns the constitutive explanation of a moral principle), whereas (GEN-IN) is a matter of first-order normative ethics (since is involves the typical explanation of a particular instance in terms of a moral principle).

But even if we grant that (GEN-IN) and (IN-GEN) figure in different disciplines, it is not clear that the difference between these disciplines is substantially deep to begin with.[28] It seems that there are clear cases where these two domains are

---

[27] For discussion see Loewer (1996: 113); Miller (2016: 16); Jaag & Loew (2018: 10); Dorst (2018: 17); Bhogal (2020: 13).

[28] Indeed, it is hard to tell whether (GEN-IN) is a paradigmatic instance of moral metaphysics. It could be argued, for example, that it seems perfectly reasonable in the context of first-order discourse to explain a moral principle by citing instances that figure within its scope. For example, a utilitarian might take a particular case of utility maximization to be *evidence* for the fact that the principle of utility holds. Of

intertwined. For example, metaethical accounts that propose the reduction of moral phenomena in terms of their non-moral, naturalistic, constituents are arguably friendly to some form of consequentialism (e.g. Boyd 1988: 4.3.). Additionally, a commitment to error theory (roughly, the view that there are no moral truths) seems to lead (on certain readings) to the first-order view that nothing is morally required. Finally, Berker (2018: 3) correctly notes that many of the claims made by ethicists *and* metaethicists can be both captured by grounding-locutions. These interconnections seem to blur the line between the two domains.[29]

7.2. Against (Irreflexivity)

(Transitivity*) tackles the circularity problem since it blocks the chaining of (IN-GEN) and (GEN-IN) (bracketing exceptional and artificial circumstances). But it could be argued that (Irreflexivity) should get the same treatment as (Transitivity). In this sense, (Irreflexivity) becomes:

> **(Irreflexivity*)** A cannot explain itself, unless, given two explanations of the form "A explains$_1$ B" and "B explains$_2$ A", both explanations accommodate the same explanatory-relevance constraint.

Some might find (Irreflexivity*) hard to accept. Some would take (Irreflexivity) to be self-evident. (Irreflexivity*), on the other hand, like (Asymmetry*), allows for some cases where explanation is circular.

As far as I can tell, there are at least two reasons to be skeptical towards (Irreflexivity*). First, circular explanations, in general, seem bizarre. Arguably, this is because they are predicated upon the existence of determination-loops. But the existence of such loops is an empirical issue. It might be that there is no retroactive

---

course, this evidential characterization of explanation is controversial but it further illustrates that the ethics/metaethics distinction can get quite murky.

[29] For more examples of such interconnections see Caleb & Schroeder (2019) and Väyrynen (2019).

causation at the actual world but there is nothing conceptually incoherent about it. In this sense, reflexive explanation cannot be denied on conceptual grounds. More important, however, is the worry that if (Irreflexivity) is rejected then it might be the case that some facts are ultimately explained purely in virtue of their own nature. If (M+) is true, then it seems that the ultimate explanation of such particular moral facts lies in themselves (ditto for moral principles). In this sense, (M+) would be like the fact that God explains herself and other exotic self-explanatory facts.

I am sympathetic to the claim that if (M+) is circular in the same way the fact that God explains herself is circular, then this is a problematic case of circularity. But this result can be avoided. As previously noted, if we assume a pluralistic conception of explanation, there is a sense in which many phenomena can be explained in many different ways depending on the details of our explanatory project. Of course, entities like God might be, ultimately, explained only in virtue of themselves. But these facts are in the minority and it is plausible that moral generalizations do not belong in this category. Moral generalizations are concrete objects in spacetime, and, in this sense, they are facts which can be explained in many different ways, one of which is a highly artificial explanation like (M+). So, the amendment of (Irreflexivity) to (Irreflexivity*) is harmless.

8. Conclusion

I argued that moral generalizations, understood as lists of particular moral phenomena, can be genuine explanatory. First, I explored the metaphysical status of moral generalizations. In section 3, I tackled Berker's challenge by showing that (Non-Idleness) is not the only explanatory relevance constraint and that (Simple) can be plausibly understood as an explanation that proceeds via subsumption as per (Unification). Then, I defended (Unification) by noting that there is nothing particularly suspect in terms of its metaphysics. In sections 6, I motivated (Unification)

by showing that it follows from the independently plausible thesis of moral explanatory pluralism. Finally, in section 7, I applied the explanatory pluralist framework to motivate a transitivity principle that blocks Rosen's circularity objection.

References

Bader, Ralf M. (2017). The grounding argument against non-reductive moral realism. *Oxford Studies in Metaethics* 12.

Baker, Derek (2021). If You're Quasi-Explaining, You're Quasi-Losing. In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics 16*. Oxford University Press.

Berker, Selim (2018). The Explanatory Ambitions of Moral Principles. *Noûs* 53 (4):904-936.

Bhogal, Harjit (2020). Nomothetic Explanation and Humeanism about Laws of Nature. *Oxford Studies in Metaphysics*.

Bokulich, Alisa (2016). Fiction As a Vehicle for Truth: Moving Beyond the Ontic Conception. *The Monist* 99 (3):260-279.

Boyd, Richard (1988). How to be a Moral Realist. In G. Sayre-McCord (ed.), *Essays on Moral Realism*. Cornell University Press. pp. 181-228.

Brink, David O. (2013). *Mill's Progressive Principles*. Oxford University Press UK.

Carlson, G. (1977). *Reference to Kinds in English*, Ph.D. dissertation, University of Massachusetts, Amherst.

Cartwright, Nancy; Pemberton, John & Wieten, Sarah (2020). Mechanisms, laws and explanation. *European Journal for Philosophy of Science* 10 (3):1-19.

Craver, Carl F. (2009). *Explaining the Brain*. Oxford University Press.

Craver, Carl F. (2019). Idealization and the Ontic Conception: A Reply to Bokulich. *The Monist* 102 (4):525-530.

Dorst, Chris (2019). Towards a Best Predictive System Account of Laws of Nature. *British Journal for the Philosophy of Science* 70 (3):877-900.

Dretske, Fred I. (1977). Laws of nature. *Philosophy of Science* 44 (2):248-268.

Dunaway, Billy & McPherson, Tristram (2016). Reference Magnetism as a Solution to the Moral Twin Earth Problem. *Ergo: An Open Access Journal of Philosophy* 3.

Duncan, Michael; Miller, Kristie & Norton, James (2017). Is Grounding a Hyperintensional Phenomenon? *Analytic Philosophy* 58 (4):297-329.

Fine, Kit (2002). Varieties of Necessity. In Tamar Szabo Gendler & John Hawthorne (eds.), *Conceivability and Possibility*. Oxford Up. pp. 253-281.

Fogal, Daniel & Risberg, Olle (2020). The Metaphysics of Moral Explanations. *Oxford Studies in Metaethics* 15.

Friedman, Michael (1974). Explanation and scientific understanding. *Journal of Philosophy* 71 (1):5-19.

Friend, Toby (2016). Laws are conditionals. *European Journal for Philosophy of Science* 6 (1):123-144.

Hicks, Michael Townsen & van Elswyk, Peter (2015). Humean laws and circular explanation. *Philosophical Studies* 172 (2):433-443.

Jaag, Siegfried & Loew, Christian (2018). Making best systems best for us. *Synthese*:1-26.

Jenkins, Carrie (2011). Is Metaphysical Dependence Irreflexive? *The Monist* 94 (2):267-276.

Kellert, Stephen H.; Longino, Helen E. & Waters, C. Kenneth (eds.) (2006). *Scientific Pluralism*. Univ of Minnesota Press.

Kim, Jaegwon (1973). Causation, nomic subsumption, and the concept of event. *Journal of Philosophy* 70 (8):217-236.

Kitcher, Philip (1989). Explanatory unification and the causal structure of the world. In Philip Kitcher & Wesley Salmon (eds.), *Scientific Explanation*. Minneapolis: University of Minnesota Press. pp. 410-505.

Krämer, Stephan & Roski, Stefan (2017). Difference-making grounds. *Philosophical Studies* 174 (5):1191-1215.

Lange, Marc (2018). Transitivity, self-explanation, and the explanatory circularity argument against Humean accounts of natural law. *Synthese* 195 (3):1337-1353.

Lewis, David (1983). New work for a theory of universals. *Australasian Journal of Philosophy* 61 (4):343-377.

Litland, Jon Erling (2013). On Some Counterexamples to the Transitivity of Grounding. *Essays in Philosophy* 14 (1):3.

Little, Margaret & Lance, Mark (2004). Defeasibility And The Normative Grasp Of Context. *Erkenntnis* 61 (2-3):435-455.

Lloyd, A. C. (1955). The logical form of law statements. *Mind* 64 (255):312-318.

Loewer, Barry (2012). Two accounts of laws and time. *Philosophical Studies* 160 (1):115-137.

Machamer, Peter; Darden, Lindley & Craver, Carl F. (2000). Thinking about mechanisms. *Philosophy of Science* 67 (1):1-25.

Makin, Mark (2019). Rigid/non-rigid grounding and transitivity. *Inquiry: An Interdisciplinary Journal of Philosophy* 62 (2):136-150.

Maudlin, Tim (2007). *The Metaphysics Within Physics*. Oxford University Press.

McKeever, Sean & Ridge, Michael (2006). *Principled Ethics: Generalism as a Regulative Ideal*. Oxford University Press.

Mellor, D. H. (1995). *The Facts of Causation*. Routledge.

Miller, Elizabeth (2015). Humean scientific explanation. *Philosophical Studies* 172 (5):1311-1332.

Perl, Caleb & Schroeder, Mark (2019). Attributing error without taking a stand. *Philosophical Studies* 176 (6):1453-1471.

Potochnik, Angela (2015). Causal patterns and adequate explanations. *Philosophical Studies* 172 (5):1163-1182.

Railton, Peter (1978). A deductive-nomological model of probabilistic explanation. *Philosophy of Science* 45 (2):206-226.

Raven, Michael J. (2015). Ground. *Philosophy Compass* 10 (5):322-333.

Robinson, Luke (2007). Moral Principles Are Not Moral Laws. *Journal of Ethics and Social Philosophy* 2 (3):1-22.

Robinson, Luke (2011). Moral Principles As Moral Dispositions. *Philosophical Studies* 156 (2):289-309.

Rosen, Gideon (2010). Metaphysical Dependence: Grounding and Reduction. In Bob Hale & Aviv Hoffmann (eds.), *Modality: Metaphysics, Logic, and Epistemology*. Oxford University Press. pp. 109-36.

Rosen, Gideon (2017). What is a Moral Law? *Oxford Studies in Metaethics* 12.

Ross, David (1939). *Foundations of Ethics.* Oxford University Press.

Salmon, W. (2002). A realistic account of causation. In Michele Marsonet (ed.), *The Problem of Realism.* Ashgate. pp. 106--134.

Scanlon, T. M. (2014). *Being Realistic About Reasons.* Oxford University Press.

Schaffer, Jonathan (2010). Monism: The Priority of the Whole. *Philosophical Review* 119 (1):31-76.

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

Schaffer, Jonathan (2017). Laws for Metaphysical Explanation. *Royal Institute of Philosophy Supplement* 82:1-22.

Schroeder, Mark (2005). Cudworth and Normative Explanations. *Journal of Ethics and Social Philosophy* 1 (3):1-28.

Scriven, Michael. (1962). Explanations, predictions, and laws. University of Minnesota Press, Minneapolis. Retrieved from the University of Minnesota Digital Conservancy, https://hdl.handle.net/11299/184631.

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

Trogdon, Kelly (2018). Grounding-mechanical explanation. *Philosophical Studies* 175 (6):1289-1309.

Väyrynen, Pekka (2009). A Theory of Hedged Moral Principles. *Oxford Studies in Metaethics* 4:91-132.

Väyrynen, Pekka (2018). Reasons and Moral Principles. In Daniel Star (ed.), *The Oxford Handbook of Reasons and Normativity.* pp. 839-61.

Väyrynen, Pekka (2019). Normative Commitments in Metanormative Theory. In Jussi Suikkanen & Antti Kauppinen (eds.), *Methodology and Moral Philosophy*. Routledge. pp. 193-213.

Wilsch, Tobias (2021). The governance of laws of nature: guidance and production. *Philosophical Studies* 178 (3):909-933.

Wilson, Alastair (2020). Classifying Dependencies. In D. Glick, G. Darby and A. Marmodoro (eds.) *The Foundation of Reality: Fundamentality, Space and Time*. Oxford University Press.

Wilson, Jessica M. (2014). No Work for a Theory of Grounding. *Inquiry: An Interdisciplinary Journal of Philosophy* 57 (5-6):535-579.

Woods, Jack (2018). Emptying a Paradox of Ground. *Journal of Philosophical Logic* 47 (4):631-648.

Woodward, James & Hitchcock, Christopher (2003). Explanatory generalizations, part I: A counterfactual account. *Noûs* 37 (1):1–24.

Wright, Cory & van Eck, Dingmar (2018). Ontic Explanation Is either Ontic or Explanatory, but Not Both. *Ergo: An Open Access Journal of Philosophy* 5:997–1029.

Wygoda Cohen, Shlomit (2020). Not All Partial Grounds Partly Ground: Some Useful Distinctions in the Theory of Grounding. *Philosophy and Phenomenological Research* 100 (1):75-92.