

Normative explanation unchained

Pekka Väyrynen 

University of Leeds

Correspondence

Pekka Väyrynen, University of Leeds.

Email: p.vayrynen@leeds.ac.uk

Abstract

Normative theories aim to explain why things have the normative features they have. This paper argues that, contrary to some plausible existing views, one important kind of normative explanations which first-order normative theories aim to formulate and defend can fail to transmit downward along chains of metaphysical determination of normative facts by non-normative facts. Normative explanation is plausibly subject to a kind of a justification condition whose satisfaction may fail to transmit along the relevant kind of metaphysical hierarchy. A broader aim of the paper is to contribute to systematic theorizing about normative explanation: whereas there has been a great deal of work on scientific explanation, there has been little by way of systematic exploration of what sort of explanations normative theories aim to formulate and defend.

1 | INTRODUCTION

Gender-based wage discrimination is morally wrong because it is arbitrary. Eating dark chocolate is good for us in virtue of how it is pleasurable and healthy. My hitting my brother was wrong because it was done with the sole intention of causing harm. These are examples of normative explanation: a kind of non-causal explanation of why something is wrong, good, what we ought or have reason to do, or the like.¹ Some kind of further explanation could often go on. For instance, eating dark chocolate is healthy because dark chocolate is a powerful source of antioxidants. Here a question arises: would it *follow* that eating dark chocolate is good because dark chocolate is a powerful source of antioxidants? The question I mean to ask here is whether the provision of normative explanation transmits downward along such chains of metaphysical determination.

¹I'll use 'normative' as shorthand for 'deontic or evaluative'.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Philosophy and Phenomenological Research* published by Wiley Periodicals LLC on behalf of Philosophy and Phenomenological Research Inc

Whether it does so has direct implications for the explanatory structure of first-order normative theories, since such theories aim to tell us not only *which* things are right and wrong, good and bad, and so on, but also *why* they are so. The question bears directly also on how normative explanation relates to other kinds of explanation and how it fits into our overall explanatory picture of the world.

This paper argues that normative explanation can fail to transmit downward along chains of metaphysical determination of the particular normative fact that is being explained. More precisely, I'll argue that this is so for an important kind of explanations of normative facts which first-order normative theories aim to formulate and defend. I'll first say more about the transmission question and its significance and describe an impasse on the issue (§2). The impasse can be extracted from discussions of normative explanation in the literature even as the transmission question itself is rarely discussed in its own right. Next I explain why some possible ways out of the impasse are insufficient and formulate a new proposal: normative explanation is plausibly subject to a kind of a justification condition whose satisfaction may fail to transmit along the relevant kind of metaphysical hierarchy (§3). I then further defend this proposal by addressing various potential complications (§4).

A broader aim of the paper is to contribute to systematic theorizing about normative explanation. The question of what constitutes a normative explanation is as central to normative inquiry as the question of what constitutes a scientific explanation is to scientific inquiry. But whereas there has been a great deal of work on scientific explanation, there has been comparatively little on normative explanation, and especially little on what sort of explanations first-order normative theories aim to formulate and defend. Those who already think that normative explanation isn't transmissive in the relevant way may take my discussion as an attempt to spell out the view in some detail (not a straightforward task), while those with opposite sympathies will, I hope, find reason to reconsider.

2 | AN IMPASSE ON NORMATIVE TRANSMISSION

Consider chains of non-causal explanations with the following kind of structure:

Pleasure

- (E1) Experience E is good because E is an experience of pleasure.
- (E2) E is an experience of pleasure because the subject of E is in brain state B (dopamine activation, or whatever).
- (E3) E is good because the subject of E is in brain state B.

This example is simple, and perhaps banal. But simple examples can be useful for exploring the structure of a phenomenon, which is my aim. To generate nearby variants, we might replace the brain state B with a microphysical state M such that the subject of E is in B because she is in M. Or we might replace 'because' with other explanatory idioms, such as 'in virtue of'. For the present purposes, I'll also keep my interpretation of *Pleasure* reasonably neutral on the nature of pleasure and what is good-making about it.² Examples like *Pleasure* also multiply. For instance, a societal arrangement might be unjust

²A reviewer for *PPR* worried that (E1) and (E2) won't be true at the same time in the way that my use of *Pleasure* requires. Perhaps (E1) is plausible when we focus on what it is like to feel pleasure. But broadly physicalist accounts of pleasure required for (E2) might have a hard time capturing the phenomenological aspects of pleasure and raise the question whether pleasure is, in itself, good-making. Fortunately, my purposes require fairly minimal assumptions here. First, while pleasure might not be, in itself, good-making if it were nothing but being in a certain brain state, we may worry less if pleasure is, instead, realizable in different physical states in different kinds of physical structures. (E2) allows this. Second, (E1) takes no stand on whether what is good-making about pleasure are (only) its qualitative character or (also) something else, such as perhaps its functional or dispositional aspects. Also note that in debates about the nature of pleasure, one common objection to "felt-quality" theories (on which what makes an experience one of pleasure is its qualitative character) is that these theories in fact cannot adequately explain the normative status of pleasure (for a summary and references, see Pallies, 2020, §2.3).

because it involves unequal access to educational opportunity, where this holds in virtue of some more specific configuration of social and economic conditions (perhaps different ones in different cases).

Suppose (E1) and (E2) are true in *Pleasure*. What should we think about (E3)? One interpretation of this question is that it asks whether the following transmission principle regarding normative explanation is true (where ‘because’ stands for a non-causal explanatory relation):

Transmission Principle: For any normative fact N and any non-normative facts A and B, if N obtains because B obtains and B obtains because A obtains, then N obtains because A obtains.

The Transmission Principle is a “mixed” principle: it tells us that explanations of normative facts and explanations of non-normative facts chain in a certain way. The principle is false if there are cases where the antecedent is true but the consequent is false. If *Pleasure* were such a case, the explanation of the normative fact [E is good] in (E1) wouldn’t transmit downward, via the explanation of the non-normative fact [E is an experience of pleasure] in (E2), so as to make (E3) a true explanation of [E is good].³

Things are unfortunately not so simple that we can simply proceed to ask whether the Transmission Principle is true. Talk of explanation is notoriously messy and diverse.⁴ By ‘explanation’ we may mean acts of explaining, products of such acts, or informational structures that can exist without being invoked in acts of explaining. We use ‘explanation’ to relate sentences, propositions, facts, or even objects. We distinguish partial and complete explanations.⁵ We not only contrast true and false explanations, but also rank explanations as better or worse. We may also be the kind of explanatory pluralists who think that ‘because’ may express different kinds of explanations even among non-causal explanations. So the Transmission Principle is but a schema that can be filled out in many different ways. Without further specification, it will be all too easy for different parties to talk past one another in applying the Transmission Principle to cases like *Pleasure*.

My interest concerns one important kind of explanations of normative facts which first-order normative theories (theories of personal good, normative ethics, distributive justice, political legitimacy, and the like) aim to formulate and defend. My initial examples and examples like *Pleasure* illustrate the kind I have in mind. I ask whether the truth or correctness of such explanations transmits downward along chains of metaphysical determination. So the transmission principle which I am interested in assessing is this:

Normative Transmission Principle: For any normative fact N and any non-normative facts A and B, if N obtains (in part) because_N B obtains and B obtains (in part) because_M A obtains, then N obtains (in part) because_N A obtains.⁶

³I use square brackets to denote facts. Those who are so inclined may read ‘normative fact’ in a deflationary way or replace it with a preferred alternative.

⁴For discussion, see, e.g., Jenkins (2008) and Ruben (2012, Chapter 1).

⁵(E1)–(E3) look more like partial than complete explanations. I’ll work with partial explanations, because it would be hard to make progress without bracketing the controversial general issue of what constitutes a complete explanation of a particular fact in the first place. For instance, what are the roles of factors like laws or principles and enablers? Are they parts of complete explanations? Are they some kind of background conditions for complete explanations? Or do they play no role at all in singular explanation proper? And what might their explanations look like, in case they have any? (Some relevant discussions of the explanatory role of moral principles include Dancy, 2004, Chapter 3; Rosen, 2017; Enoch, 2019; Berker, 2019; Fogal & Risberg, 2020; and Baker, 2021. On enablers in explanation, see, e.g., Dancy, 2004, pp. 45–9, and Skow, 2016, pp. 85, 109.) Bracketing such questions here isn’t ad hoc. Many think that the explanation-backing relations which figure in the relevant kind of partial explanations transmit in the relevant way. Examples like *Pleasure* should then give us guidance on what overall patterns hold.

⁶I would expect those who accept this principle to think that it holds by necessity.

Here ‘because_M’ picks out a non-causal relation of metaphysical explanation and ‘because_N’ picks out the kind of explanation of normative facts (whatever that turns out to be) which is of interest to first-order normative theories. For now, it is an open question whether the two are distinct and how they are related. For instance, ‘because_N’ and ‘because_M’ might merely have different kinds of explananda, without thereby expressing different explanatory relations. Either way, the Normative Transmission Principle might still be false.

Above I speak of “the kind” of explanations which first-order normative theories seek only for the sake of simplicity. Whether the Normative Transmission Principle is true is independent of whether normative theories may offer also other kinds of explanations of normative facts, in addition to the kind of bottom-up explanations which the Normative Transmission Principle models.⁷ For present purposes, I can be open to an explanatory pluralism on which multiple distinct kinds of normative explanation can complement each other, not compete to replace each other.

To highlight the coming attractions, I’ll argue that normative facts have features which constrain properly explanatory first-order normative theories in a way that doesn’t apply to other sorts of explanations. If that is right, then there will be principled grounds to deny that explanations like (E1) and (E2) can always be chained to yield (E3) as a true normative explanation of a sort that is of interest to first-order normative theories. We can fix some ideas about normative explanation without stacking this deck either way. I’ll understand explanations as contents of correct answers to why-questions. Since such contents can be true (or false) irrespective of whether they are invoked by speech acts of explanation, this allows us to think of correct explanations of normative facts as something to be discovered. Here I’ll also take the Normative Transmission Principle to concern specifically the truth or correctness of explanations which relate particular facts (such as that it was wrong of Janet to break her promise to Mehmet) or patterns among them (such as that promise-breaking is at least normally wrong). In the general theory of explanation, the explanation of laws or principles is widely taken to raise distinct issues.

There are good reasons to interpret the Normative Transmission Principle as concerning a notion of explanation which is objective in two further respects.⁸ First, correct explanation isn’t relative to the interests or background beliefs of an audience in ways that would rule out understanding it as an objective relation between its relata.⁹ Second, correct explanation needn’t be such as to induce understanding.¹⁰

One reason to adopt a notion of explanation which is objective in these respects is that it provides for a significant debate about whether the Normative Transmission Principle is true. It seems clear that the epistemic or pragmatic features which less objectivist views take to constrain correct explanation can fail to transmit in accordance with the principle. An audience’s interests or background beliefs can easily be such that hedonic facts relate to goodness in a way which such views require for correct explanation, but brain-state facts don’t. It seems equally clear that factors thanks to which an explanation induces understanding (when it does) may similarly fail to transmit along

⁷Walden (2016, p. 206) argues that ethical theorizing has multiple distinct explanatory goals. Baker (2021) gives an account of normative explanation as a form of generalizing explanation.

⁸This characterization follows Väyrynen (2019).

⁹Some reasons behind this slightly cumbersome formulation are discussed in Lipton (2004, Chapter 3).

¹⁰While providing understanding may be something that explanations are particularly well suited to do, it doesn’t follow that this is a constraint on the correctness of explanations (Strevens, 2013; Skow, 2017).

the kind of hierarchy we find in *Pleasure*. An epistemic or pragmatic notion of explanation would thus stack the deck unduly against the Normative Transmission Principle.¹¹

Another reason to adopt a more objective notion of explanation is that this fits better with the explanatory ambitions of first-order normative theories. A simple illustration is act-utilitarianism, understood as an explanatory theory according to which any action is right if and only if, and because, it maximizes general happiness. Act-utilitarianism doesn't constrain its explanatory claims about what makes actions right by such epistemic or pragmatic factors as what a given audience's background beliefs may be (at least, not beyond any that matter to which actions are right to begin with) or whether the audience's interests include impartial benevolence or a concern to act morally.¹² The fundamental reason why wrong actions are wrong is that they fail to maximize general happiness, irrespective of our interests. Or consider that if Rossian pluralism is the correct moral theory, then any act is a pro tanto duty insofar as, and because, it involves fidelity to promises. The truth of this explanatory claim doesn't depend on whether the given audience understands how fidelity to promises contributes to moral obligation. As I'll explain in §4, my claims allow that which question a why-interrogative such as 'Why was it wrong to tell him?' is used to ask depends on the context. Once the question is set, something's being a correct answer to it can be independent of interests and background beliefs. My claims here are compatible also with the thought that moral theories aim to guide action. This practical aim of moral theories is more plausibly understood as distinct from their theoretical aim of explaining why the right actions are right.

Returning to *Pleasure*, suppose still that (E1) and (E2) are true. Is (E3) then a correct explanation of why E is good, in the kind of objective sense with which explanatory normative theorizing seems to be concerned? The issue remains controversial in a way that doesn't seem to be resolvable by batting around intuitions about examples.

Some philosophers claim that normative explanation doesn't transmit downward in the relevant kind of way. For instance, Zangwill writes: "Epistemic properties depend (wholly or partly) on mental states, and it is not implausible that mental states depend (wholly or partly) on neural states, but it is not plausible that epistemic properties depend (wholly or partly) on neural states" (Zangwill, 2018, p. 842). Zangwill takes epistemic properties to be normative and dependence to be an explanatory relation that cannot be cashed out in purely modal terms. He suggests that when explanatory chains go across categories of things, not even the generic Transmission Principle holds – or at least that this is the most plausible default view (Zangwill, 2018, p. 842). *Pleasure* is such a case. So Zangwill would regard the failure of the Normative Transmission Principle also as the most plausible default view.

However, many other philosophers at least implicitly commit themselves to the Normative Transmission Principle through their claims about what kind of relation 'because' or 'in virtue of' picks out in normative explanations like (E1). For instance, Shafer-Landau claims that particular moral facts hold in virtue of particular non-moral facts in the sense that they are constituted or realized by those non-moral facts (Shafer-Landau, 2003, pp. 75-77, 92). If E is good because it is an experience

¹¹An objective notion of explanation can accommodate certain epistemic considerations by allowing that (E3) can be true in an evidential sense of 'because'. For those with advanced neurophysical knowledge, the occurrence of brain state B might be sufficient evidence that E is good. Other senses in which the truth of (E3) wouldn't support the Normative Transmission Principle will emerge in the course of my discussion.

¹²My claim here doesn't stack the deck against expectational versions of moral theories. The epistemic factors built into such theories entail different verdicts already about what actions are right. Once those facts are fixed, explanation of why an act is right needn't make further reference to background beliefs or credences. For parallel reasons, the claim doesn't conflict with moral relativism or subjectivism or with the principle that 'ought' implies 'can' in a sense that entails epistemic accessibility. Nor does it conflict with internalism about normative reasons. The view that an agent has reason to perform an act that is right only when it is suitably connected to her desires is silent regarding that in virtue of which the act is right.

of pleasure, there seems to be no principled reason for Shafer-Landau not to allow that [E is an experience of pleasure] could in turn be constituted or realized by some such lower-level fact as [The subject of E is in brain state B]. Since constitution and realization are metaphysical determination relations, it seems to follow that E is good because its subject is in B. So if normative explanations like (E1) work on the back of constitution or realization relations, one would expect the kind of explanation that goes on in (E1) to transmit downward along chains of such relations. The same point goes through if normative explanation works instead on the back of minimally metaphysically sufficient determination (Strandberg, 2008) or certain kinds of “grounding”.

In fact, many ground-theoretic frameworks support a simple argument that normative explanation in (E1) transmits via (E2) down to (E3). Suppose (G1)-(G4) are all true:

(G1)	[The subject of E is in brain state B] (partly) grounds [E is an experience of pleasure].
(G2)	[E is an experience of pleasure] (partly) grounds [E is good].
(G3)	<i>Grounding-Explanation Link</i> : For any facts A and B, if A (partly) grounds B, then B obtains (in part) because A obtains.
(G4)	Grounding (partial or otherwise) is transitive.

(G4) is widely accepted.¹³ And a guiding idea about how grounding differs from purely modal determination relations like supervenience is that grounding is explanatory in a way that these aren't.¹⁴ This idea might be cashed out by saying that grounding is itself an explanatory relation or by saying that it “backs” explanation. (G3) is neutral between these views. But (G1)-(G4) jointly entail that the normative explanation in (E1) is true and transmits via (E2) in accordance with the Normative Transmission Principle in such a way that (E3) is true under the relevant interpretation.¹⁵

Precisely which transmission principle this reasoning supports depends on various internal disputes within the ground-theoretic framework. For instance, grounding pluralists think that there are fundamentally distinct kinds of relations of metaphysical and normative grounding (Fine, 2012; Rosen, 2017) whereas grounding monists think that these are of the same fundamental, generic kind of grounding relation (Berker, 2018). These views differ on whether various mixed grounding transmission principles would be fundamentally distinct. I won't take the space to make them explicit. For my purposes, three points suffice.

First, those who have worries about the explanation transmission principles set out above would presumably have parallel worries about grounding transmission principles.

Second, the view that there are more than one fundamentally distinct kinds of grounding relation is compatible with the claim that (E1) and (E2) can be combined to yield a true hybrid explanation in (E3).¹⁶ But such “grounding pluralism” doesn't by itself entail this claim. Whether such an entailment holds depends on how normative grounding relates to the kind of normative explanations which first-order normative theories aim to formulate and defend and whether the relevant sort of hybrid explanations would be of this kind.

¹³Agreement on (G4) isn't universal. Schaffer (2012) argues that metaphysical grounding isn't transitive. For replies, see, e.g., Litland (2013; 2018), Krämer and Roski (2017), and Wygoda Cohen (2020).

¹⁴Critical discussions of the allegedly close connection between grounding and explanation include Kovacs (2017) and Maurin (2019).

¹⁵Parallel transmission arguments for (E3) can be generated by substituting grounding with certain other metaphysical determination relations in (G1)-(G4).

¹⁶Compare Lange (2018) on mixed transitivity, as well as Litland (2018). Securing such an entailment might require modifying the Grounding-Explanation Link in (G3) in some appropriate way.

Third, some grounding theorists defend principles very similar to the Normative Transmission Principle. For instance, Berker argues that grounding pluralism cannot explain why the following mixed transitivity principle is “extremely plausible”:

($\text{Tran}_{\text{nor/met}}$) If [p] is normatively grounded in [q], and [q] is metaphysically grounded in [r], then [p] is grounded (in some non-rigged-up sense) in [r]. (Berker, 2018, p. 751.)¹⁷

Here a “non-rigged-up” sense of ‘because’ would be one that is sufficiently generic to support the grounding monist view that there is a single fundamental grounding relation. Berker offers the following example in support of ($\text{Tran}_{\text{nor/met}}$). Let W = [She acted wrongly in telling him] (a normative fact) and S = [She could have done something else instead of telling him that would have brought about more overall happiness] (a non-normative fact). And suppose that W normatively obtains in virtue of S . Berker then argues as follows:

[S]uppose that, in addition, S obtains metaphysically in virtue of L = [She could have lied instead of telling him, her lying would have brought about 100 overall units of happiness, and her telling him brought about 20 overall units of happiness]. Then it is very natural to hold that W also obtains in virtue of L , in some non-rigged-up sense of ‘in virtue of’. This offers partial confirmation of ($\text{Tran}_{\text{nor/met}}$). (Berker, 2018, p. 752.)

It is plausible that L in turn obtains metaphysically in virtue of some still lower-level non-normative fact F . So it should be equally natural to hold that W also obtains in virtue of F .

Thus if ($\text{Tran}_{\text{nor/met}}$) and (G1)–(G3) hold, then (E3) should also hold, in some “non-rigged-up” sense of ‘because’. In this framework, the question I ask when I ask whether the Normative Transmission Principle holds is whether the non-rigged-up sense of ‘because’ in ($\text{Tran}_{\text{nor/met}}$) expresses a kind of explanation of normative facts which first-order normative theories aim to formulate and defend. Berker thinks it does. He takes theories of normative ethics to be also metaphysical theories about what grounds moral facts in a generic, non-rigged-up sense (Berker, 2018, p. 770). If that is correct, the Normative Transmission Principle seems secure. As we’ll see, I deny that the case has been made.

Time to sum up. Although principles like Normative Transmission are rarely discussed explicitly in their own right, existing discussions of normative explanation and adjacent issues generate an impasse regarding whether the Normative Transmission Principle holds. Several versions of the idea that normative facts hold at least partly in virtue of non-normative facts support the idea that the corresponding sort of normative explanation transmits downward along the chains of metaphysical determination. Other writers doubt this. Considerations on each side look reasonable in their own terms but questionable under the alternative terms, so batting around intuitions about examples won’t resolve the impasse. Spelling out the impasse shows that various theorists disagree as to whether the Normative Transmission Principle is true of the kind of explanations of normative facts which first-order normative theories aim to formulate and defend. The need to spell it out explicitly shows that the disagreement has received little attention. And yet any resolution will have immediate consequences for the explanatory structure of normative theories.

A final clarification is due before considering ways out of the impasse. The Normative Transmission Principle can be read as a mixed transitivity principle analogous to Berker’s ($\text{Tran}_{\text{nor/met}}$). I don’t frame the paper as asking whether normative explanation is transitive to avoid unnecessary risks of

¹⁷ Berker states that his argument for grounding pluralism doesn’t require that this principle hold in all cases.

confusion.¹⁸ Whether the Normative Transmission Principle holds is distinct from the truth of both intranormative transitivity principles (such as ‘Necessarily, for any A, B, and N, if A normatively explains B and B normatively explains N, then A normatively explains N’) and general transitivity principles (such as ‘Necessarily, for any A, B, and C, if C obtains because B obtains and B obtains because A obtains, then C obtains because A obtains’). The former principles don’t apply in my cases. The latter don’t require that the relation linking A-B and the relation linking B-C be the same relation, but only that the relations can be linked to yield a true hybrid explanation in A-C (cf. Lange, 2018, p. 1345). Just which mixtures permit this is a case-by-case matter. My interest is restricted to the type of mixture that is found in the Normative Transmission Principle.

3 | NORMATIVE JUSTIFICATION AND TRANSMISSION

In this section I’ll offer a way out of the impasse. My proposed solution gives us a principled reason to think that the Normative Transmission Principle has instances where the antecedent clauses, such as (E1)–(E2) in *Pleasure*, cannot always be combined to yield a further true explanation such as (E3). If that is right, the Normative Transmission Principle is false.

To see what is distinctive about my solution, I’ll first briefly dismiss two alternative ways out of the impasse. (I won’t claim to exhaust the alternatives.) One possible way out of the impasse is to try to explain away either the intuitions that support the Normative Transmission Principle or the intuitions that challenge it. Intuitions which conflict with the principle might be explained away by interpreting such intuitions as reflecting epistemic gaps with respect to either the lower-level facts or their relations to the higher-level facts in the relevant kind of explanatory chains. This may be plausible in some cases. Perhaps our epistemic access to the brain states that ground pleasure or to their relationship to goodness is limited. But not all cases are like that. Consider, for instance, Michael Smith’s example of a Socratic sort of person (Smith, 2000). Her life is good because it is a life of achievement, which it is because it displays understanding. Her life in turn displays understanding because she does things like writing certain particular words (with their particular meanings) in her papers, giving lectures, and teaching and supervising students. Now, I have read enough of works like *Critique of Pure Reason* and *A Treatise of Human Nature* to see that they display great understanding. If I can, so can you. And yet those who doubt that *Pleasure* satisfies the Normative Transmission Principle may very well also doubt that the Socratic sort’s life is good because she writes the particular words (with their particular meanings) which she writes in her papers. Such intuitions cannot be explained away by appeal to epistemic gaps.

A second possible way out of the impasse is to note that more general higher-level explanations are superior to more specific lower-level explanations.¹⁹ For instance, if pleasure is multiply realizable, then (E1) is robust across a wider range of counterfactual variations than (E3). E would have been good even if the subject of E had been in some physical state other than B which realizes pleasure. But even if (E1) is in this sense superior to (E3), it doesn’t follow that (E3) isn’t also true. The underlying lower-level facts about brain states may still also combine to explain why E is good. The higher-level explanation may merely be the better of two correct explanations. The point remains even if we add

¹⁸In earlier work, I used explanatory chains similar to *Pleasure* to argue against the transitivity of normative explanation (Väyrynen, 2009, pp. 301–305). That argument may risk the kind of confusion I allude to in the text. It also relies on a (modest) epistemic constraint on explanation, unlike my argument in this paper.

¹⁹This is a common point (Jackson & Pettit, 1990; Yablo, 1992; Woodward, 2003). The less specific explanation cannot be too unspecific, though (Krämer & Roski, 2017, pp. 1207–8).

that higher-level explanations aren't just better than lower-level explanations but also irreducible to them. Explanation P1 is irreducible to explanation P2 if P1 makes some important explanatory contributions that couldn't be made by P2.²⁰ So when P1 is irreducible to P2, P2 can also be a genuine explanation of the same phenomenon. But the failure of the Normative Transmission Principle requires that (E3) be false, not merely that it is a poorer but still true explanation.

The above two solutions don't involve considering whether there is anything special or distinctive about the kind of explanations of normative facts which first-order normative theories aim to formulate and defend. An alternative way out of the impasse is to identify a feature of normative explanation which can fail to be preserved by chains of metaphysical determination of the sort we find in *Pleasure*. This is the kind of solution I propose. I'll first motivate the view that at least some normative explanations are correct only if they satisfy a kind of justification condition which doesn't apply to non-normative explanations. I'll then argue that the satisfaction of this condition isn't guaranteed to transmit in accordance with the Normative Transmission Principle.

Nothing in our general thinking about explanation rules out the idea that normative explanations of the sort which first-order normative theories aim to formulate and defend might have features lacked by other kinds of explanations. Why-questions ask for reasons. Reasons thus figure in the contents of correct answers to why-questions. Correct explanations are contents of correct answers to why-questions. Normative why-questions such as 'Why should I pay my taxes?' – let alone 'Why should I be moral?' – are often naturally heard as asking for reasons that would (also) *justify* certain actions or attitudes. If there was no reason for me to pay my taxes or if it weren't fitting for you to resent me for failing to pay my taxes, why would paying my taxes count as something that I ought to do in the first place? But we already recognize that in different contexts, why-questions ask for different sorts of reasons: sometimes causes, sometimes grounds, sometimes motives or purposes.²¹ There is no reason to think that this list must be exhaustive. There is thus no principled reason why normative why-questions couldn't ask for explanations that justify.²² Perhaps asking why some normative fact obtains doesn't in all contexts request a justification. But when it does, at least some part of a complete correct answer to the question must identify a feature that would justify, at least pro tanto, certain actions or attitudes.

As luck has it, Väyrynen (2019) defends the kind of justification condition on normative explanation which I have in mind here:

(NEJ) For any particular normative fact N, a correct complete normative explanation of why N obtains must identify features that would go at least some way towards justifying certain responses to (the constituent object of) N.

(NEJ) concerns normative justification, understood as taking actions or attitudes as its objects, not the epistemic justification of normative beliefs. If desired, (NEJ) can be modified to allow that the objects of normative justification include (also) requirements, permissions, and the like. Its upshot is that a correct complete explanation of why I ought to pay my taxes must identify features that would make it fitting or appropriate for you to resent or blame me for paying my taxes, or would provide some suitably specified agent (in this case, me) with a reason to pay my taxes, or the like. As flagged

²⁰For one version of this point, relating higher-level "program explanations" to the lower-level "process explanations" for which the former program, see Jackson and Pettit (1990).

²¹This observation is central to the theory of answers to why-questions in Skow (2016).

²²Hyman (2015, Chapter 6), Wedgwood (2017, p. 91), and Cullity (2018, pp. 12–13) think that this is what they at least normally do.

above, the ‘go at least some way’ signals that the justification may only be pro tanto or contributory, rather than all-things-considered. (NEJ) concerns complete explanation. Earlier I screened off general disputes about what it takes for something to be a complete explanation. Here I’ll assume that partial normative explanations such as those in *Pleasure* feature a non-normative fact such that if that fact doesn’t meet the condition in (NEJ), then no other part of the complete explanation will either. This looks like a fair simplifying assumption. (I address a related worry below.)

The Normative Transmission Principle is false if there are contexts where normative explanation is subject to the condition stated in (NEJ) but the satisfaction of that condition isn’t guaranteed to transmit in accordance with the principle. So *Pleasure* is a counterexample to the Normative Transmission Principle if (E1) satisfies the condition but (E3) doesn’t. In what follows I’ll offer such an argument. The argument doesn’t require that (NEJ) hold in full generality. It only requires that normative explanation be subject to the condition stated in (NEJ) in some contexts. From now on, let (NEJ) refer to this weaker thesis. I cannot defend it here in full, but some quick motivations will hopefully justify using it as a premise.²³

We can think of (NEJ) as telling us that at least some normative why-questions are such that the content of a complete correct answer to the question must identify features which would justify certain actions or attitudes, at least pro tanto. One motivation for this claim is its fit with the enterprise of normative theory. When moral theorists formulate and defend views about why certain acts are morally right and others morally wrong, their concerns normally include justification for the corresponding requirements and permissions to act. For instance, one worry that is sometimes raised against non-consequentialist moral theories is that it is unclear why agents sometimes may prioritize their own interests so that it is morally permissible for them to act in ways that don’t have the best outcome overall. This is to ask for a justification for (actions that exercise) such agent-centered prerogatives. Whether such a justification holds isn’t sensitive to agents’ interests or background beliefs or to whether it induces understanding. Rather, the justification should just be there in principle to be given. And if that is case, then exercising such a prerogative can be normatively justified even if the agent isn’t able to understand or give the justification. So the relevant notion of normative justification seems objective in the same kind of way as the relevant notion of explanation. The view that normative explanation, in the relevant objective sense, is subject to (NEJ) fits well with the justificatory dimension of first-order normative theories. It therefore seems safe to assume that, among the explanations of normative facts which first-order normative theories aim to formulate and defend, at least some are subject to (NEJ).

Another motivation for (NEJ) concerns the nature of normative facts.²⁴ One way in which normative facts are often thought to differ from non-normative facts is that normative facts ipso facto mean that some suitably specified agent would have normative reason to respond in a certain way, at least in certain suitably specified circumstances, or that certain responses would be appropriate, apt, fitting, or warranted – that is, justified.²⁵ That E is an experience of pleasure leaves it a further question what responses to E would be fitting or apt. But if E is good, then it is ipso facto appropriate to promote, commend, or cherish E.²⁶ This link between normative facts and the appropriateness of certain responses looks tight enough for the latter to be part of what is being explained in explanations of

²³For something more like arguments for (NEJ), see Väyrynen (2019).

²⁴I lift this motivation from Väyrynen (2019). I take it to summarize a widely held view.

²⁵Views might vary on whether normative facts are this way in virtue of the nature of normative concepts or in virtue of the nature of normative properties. For my purposes, we needn’t choose.

²⁶This doesn’t imply that goodness itself provides reasons for such responses, at least in any non-derivative sense. The reasons might instead be provided by the properties that make things good.

normative facts as normative. Such explanations are plausibly among those which first-order normative theories aim to formulate and defend. If an explanation of why x has property P doesn't identify features that would justify promoting, commending or cherishing x , then it isn't an explanation of why x is *good*. If some such responses to E weren't apt or fitting, why would E count as good in the first place? If this is right, there is no reason to expect that (NEJ) will apply to explanations of non-normative facts. It can therefore claim to specify a relevantly special feature of normative explanation, one not shared by other kinds of explanations.

The next step is to argue that we shouldn't expect the satisfaction of the justification condition in (NEJ) to transmit in accordance with the Normative Transmission Principle. Suppose that *Pleasure* is a case where the explanation of why E is good is subject to (NEJ). (NEJ) and (E1) jointly suggest that (J1) is true and (NEJ) and (E3) jointly suggest that (J2) is true:

- | | |
|------|---|
| (J1) | [E is an experience of pleasure] would go at least some way towards justifying some such responses as promoting, commending, or cherishing E . |
| (J2) | [The subject of E is in brain state B] would go at least some way towards justifying some such responses as promoting, commending, or cherishing E . |

I'll argue that (J1) and (J2) needn't stand or fall together. More precisely: at least given certain plausible first-order normative assumptions, it can be that (J1) is true but (J2) is false, even under the objective notion of normative justification discussed above. If the explanation of [E is good] is subject to (NEJ), then it can be that (E1) is a true normative explanation but (E3) is false, in that the complete explanation corresponding to (E3) doesn't satisfy (NEJ). But then *Pleasure* is a case where normative explanation can fail to transmit from (E1) via (E2) to (E3) in accordance with the Normative Transmission Principle.

The reason why (J1) and (J2) needn't stand or fall together is that [E is an experience of pleasure] and [The subject of E is in brain state B] may differ justificatorily with respect to [E is good]. Explanation is widely regarded as hyperintensional: substitution of necessary equivalents in explanatory contexts may fail to preserve truth-value.²⁷ For instance, 'It is true that grass is green because grass is green' is a true explanation but 'Grass is green because grass is green' is false, even as, necessarily, it is true that grass is green if and only if grass is green (Schnieder, 2011, p. 445). There is no reason why normative explanation wouldn't be hyperintensional as well. But then, when the truth of a normative explanation is constrained by (NEJ), it is reasonable to expect the relevant kind of normative justification also to be hyperintensional. But then it should be perfectly possible that (J1) is true while (J2) is false. If so, there should be at least one instance of a *Pleasure*-style example in which (NEJ) isn't met.

It can be that (J1) is true but (J2) is false even under tighter metaphysical assumptions than *Pleasure* requires about how [E is an experience of pleasure] and [The subject of E is in brain state B] are related. For example, this can hold even if (J1) and (J2) concerned the same fact under different modes of presentation or concerned worldly non-representational states of affairs that are only hyperintensionally distinct.²⁸ That is, it should be perfectly possible that [E is an experience of pleasure] makes it apt or fitting to promote or commend E but [The subject of E is in brain state B] doesn't, even if the

²⁷On the hyperintensionality of explanation, see, e.g., Ruben (2012, Chapter 5). Grounding is also widely regarded as hyperintensional.

²⁸The former would be the case if experiencing pleasure were identical with being in brain state B . The latter would be the case if properties or facts were individuated hyperintensionally (Nolan, 2014, p. 158) but it were metaphysically necessary that an experience is one of pleasure if and only if its subject is in B . Note that even if explanation is hyperintensional because its truth-value is sensitive to modes of presentation, its correctness needn't depend on anyone grasping the modes of presentation under which facts are introduced into the explanation.

two were necessarily equivalent. After all, even if maximizing general happiness and being commanded by God were necessarily equivalent, it wouldn't follow that they are equivalent with respect to normative justification. If even the tightest assumptions about how the two facts are related metaphysically allow justificatory divergence, such divergence should be perfectly possible under weaker such assumptions.

This argument for justificatory divergence has a hole as it stands.²⁹ (NEJ) concerns complete explanation, whereas (E1)–(E3) state partial explanations and (J1)–(J2) state partial justifications. Above I sought to mitigate this complication by assuming that if [The subject of E is in brain state B], as it occurs in (E3), doesn't meet the condition in (NEJ), then no other part of the complete explanation corresponding to (E3) will do so either. The problem is that even if no other part of the complete explanation on its own satisfies (NEJ), it might be that some other part of the explanation together with [The subject of E is in brain state B] does satisfy (NEJ). So even if (J2) is false, the following might be true:

(J2*) [The subject of experience E is in brain state B] would, perhaps with some other facts, go at least some way towards justifying some such responses as promoting, commending, or cherishing E.

The truth of (J2*) would suffice to show that (E3) satisfies (NEJ). To assess whether (J2*) is sufficiently more plausible than (J2), we need some idea of what the other facts might be. The objection is silent on this, and I cannot survey all possibilities. What I'll argue is that on a natural kind of way to fill in the details, (J2*) isn't sufficiently more plausible than (J2).

What we are looking for are facts whose conjunction with [The subject of experience E is in brain state B] constitutes a set of facts which would justify, at least pro tanto, some such responses as promoting, commending, or cherishing E. A natural thought here is to invoke facts which together with [The subject of E is in brain state B] wholly make it the case that E is an experience of pleasure. There are some controversies about what this takes, such as whether the relevant set must include only all facts that play a distinctive grounding role or all facts that play either a distinctive grounding role or an enabling role.³⁰ But what will hopefully do as a representative example is [Brain state B plays functional role F], where F is the functional role of pleasure. So a representative version of the objection claims that if (J1) is true under my assumptions, then so is the following:

(J2**) [The subject of experience E is in brain state B] and [B plays functional role F] would go at least some way towards justifying some such responses as promoting, commending, or cherishing E.

To assess (J2**), recall what my claim about (J2) was: at least given certain plausible first-order normative assumptions, it can be that (J1) is true but (J2) is false. So I am committed to claiming that, at least given certain plausible first-order normative assumptions, it can be that (J1) is true but (J2**) is false. This claim strikes me as defensible. (I'll describe the role that first-order normative assumptions play in my argument more fully shortly.)

One line of defense is that my argument from hyperintensionality applies equally well to (J2**) and other variants of (J2*). If normative justification is hyperintensional in the way that normative explanation is, then it should be equally possible that [E is an experience of pleasure] would pro tanto justify promoting, commending, or cherishing E but {[The subject of E is in brain state B] & [B plays functional role F]} wouldn't do so – even if the two were necessarily equivalent. So (J2**) isn't more plausible than (J2) in this respect.

²⁹I am grateful to the reviewer for *PPR* for raising this excellent objection.

³⁰Wygod Cohen (2020, p. 80) distinguishes between “full exclusive grounds” and “full inclusive grounds” in this way.

Another line of defense flags my focus on the sort of normative explanations which first-order normative theories aim to formulate and defend. Since (NEJ) is a necessary condition on normative explanations of that sort, it is reasonable to impose a corresponding restriction on the relevant notion of normative justification. The objection thus requires that (J2**) be sufficiently plausible specifically under a such a notion of justification. But why should first-order normative theories, as such, care about all of the metaphysical grounds of the features which they designate as justifying certain responses?³¹ If it turned out that pleasures didn't have metaphysical grounds (in the form of functional realizers or otherwise), this wouldn't require revisions to normative theories on which the fact that an experience is one of pleasure would pro tanto justify certain positive responses to it. Similarly, why should a discovery that pleasure plays a certain functional role which can be fulfilled by brain state B prompt revisions to normative theories which designate pleasure as justifying certain responses but had bracketed questions about the metaphysical basis of pleasure? Whether the functional role of pleasure plays the kind of role in normative justification which (J2**) says it does depends on what that role is – which dispositions mark experiences as ones of pleasure, and so on. But in (J2**) the functional role of pleasure matters only insofar as it relates brain state B to the nature of E as an experience of pleasure.

More therefore seems to be needed to support the thought that (J2**) is sufficiently more plausible than (J2) under the relevant notion of normative justification, than that (J2**) features a full metaphysical ground of [The subject of E is in brain state B]. This response to (J2**) should apply also to many other ways of filling in the “other facts” in (J2*) which generate something that wholly makes it the case that E is an experience of pleasure. There is then reason to think that wholly making it the case that an experience is of a certain kind is different from doing the justifying work that may be done by experiences of that kind.³² But it is a key assumption of the version of the objection I am addressing that if [E is an experience of pleasure] does a certain kind of justifying work, then the facts that wholly make it the case that E is an experience of pleasure also do that justifying work. So, although there may be other ways of filling out (J2*) and thus I don't claim to have addressed the objection definitively, I hope to have said enough to indicate why this key assumption is resistible.

I'll now return to a point flagged earlier: if (J1) and the variants under (J2) needn't stand or fall together with respect to (NEJ), then which way they may break depends on the details of our normative theory. If pleasure were normatively neutral, then *Pleasure* should be replaced with a different example. Either E wouldn't be good in the first place or it would be good for some other reason. But suppose pleasure isn't

³¹This question is different from whether certain metaphysical discoveries should make first-order normative theories reconsider what features they designate as good and bad making. I'll consider this scenario below. I say ‘all of the metaphysical grounds’ because the objection seems to be committed to claiming further that if microphysical state M grounds brain state B, then [The subject of E is in M] would, with some other facts, go at least some way towards justifying some such responses as promoting, commending, or cherishing E.

³²I am committed to this claim being robust across the distinction between mediate and immediate grounding (see, e.g., Fine, 2012, pp. 51–52). I can acknowledge that in *Pleasure*, the facts that wholly make it the case that E is an experience of pleasure are a mediate ground of the fact that E is good, mediated through such more immediate relationships of ground as that [E is an experience of pleasure] grounds [E is good]. The argument I have given against the Normative Transmission Principle implies that mediate metaphysical grounds of normative facts may fail to satisfy (NEJ) even if their more immediate grounds satisfy it. So a metaphysical ground of a normative fact which is derivative in this way may not thereby constitute a derivative normative explanation of that fact (even if it constitutes some kind of metaphysical explanation of it – see §4 below). That said: whether we should countenance derivative normative explanations, and of what kind(s), are under-explored issues in normative theory which may be relevant to how (NEJ) should more precisely be understood (cf. Väyrynen, 2019, §4). Here I can only flag that the relevant issues are distinct from how reasons, values, norms, or principles can be derived from other reasons, values, norms, or principles via instrumental relations, subsumption, or specification. (For one recent discussion of normative derivations, see Cullity, 2018, Chapter 4). I cannot here close off the possibility that some possible outcomes of inquiry into these issues might help resurrect the objection I have been addressing.

morally neutral. Then we have a choice regarding which is a normatively more fundamental reason why E is good: [E is an experience of pleasure] or [The subject of E is in brain state B]. The choice isn't settled by the fact that being in the brain state is metaphysically more fundamental than having pleasure. For what is metaphysically more fundamental needn't be more fundamental in the normative order. Rather, the choice depends on first-order normative assumptions. I have assumed that [E is an experience of pleasure] is more fundamental in the normative order than [The subject of E is in brain state B]. This isn't an unreasonable assumption. As I noted, we would still think that pleasure is good-making even it turned out that pleasures didn't have metaphysical grounds, and discovering that pleasure has a certain kind of metaphysical ground wouldn't by itself require extensions or revisions in normative theory.

My argument against the Normative Transmission Principle doesn't, however, require that the above is how the normative chips fall. For instance, imagine we conclude that pleasure is really a matter of dopamine activation and doesn't itself have a role to play in a mature science of the mind. Especially if we satisfied ourselves that pleasure talk has at most some heuristic value, we might then well conclude that dopamine activation is normatively more fundamental with respect to goodness than pleasure. Then it might be that (J2) is true but (J1) is false, and *Pleasure* couldn't be used to show that the Normative Transmission Principle is false. But the core of my argument is robust across such complications. The failure of the Normative Transmission Principle requires just that explanations of normative facts which are subject to (NEJ) discriminate among the different nodes in a chain of metaphysical determination, excluding some nodes from counting as correct normative explanations.³³ (I'll return to this issue in §4.) My claims about (J1) and (J2) imply that this is what we can expect of normative explanation if (NEJ) is true. So long as there are some examples with the structure of *Pleasure* where the normative chips fall on a metaphysically less fundamental level, my argument should go through.

Various epistemic and pragmatic considerations don't threaten my argument against the Normative Transmission Principle either. Suppose pleasure is identical to, or even just constituted by, dopamine activation and this connection becomes epistemically transparent to us. In such an event, [E is an experience of pleasure] and [The subject of E is in a dopamine activation state] might guide and regulate our actions and reactions in the same way, thanks to their epistemic (near-)equivalence. But normative justification doesn't reduce to its action-guiding role, especially in the kind of objective sense at issue in (NEJ). There can also be various epistemic and pragmatic notions that are closely connected to but distinct from normative justification in that sense. Most obviously, someone who requests a justification for promoting or commending E might be satisfied just by evidence that E is good. Given (NEJ), this implies that *there is* a normative justification. That the subject of E is in a dopamine activation state, for instance, can be evidence that there is a normative justification without itself constituting one. Finally, just as there are epistemic and pragmatic notions of explanation, there can be notions of normative justification which are constrained by epistemic or pragmatic factors. Justifications in such senses just wouldn't meet the condition in (NEJ). In short, epistemic and pragmatic considerations don't threaten the idea that pairs of facts like [E is an experience of pleasure] and [The subject of E is in brain state B] can differ with respect to normative justification in the way that my argument implies.

4 | NORMATIVE AND METAPHYSICAL EXPLANATION

My argument against the Normative Transmission Principle is restricted to explanations of normative facts of a sort which first-order normative theories aim to formulate and defend. I have argued

³³DePaul (1987, pp. 436–438) argues that what he calls “moral dependence” is “exclusive” in such a way.

that if such explanations are constrained by (NEJ), then we'll have principled reasons to expect that normative explanation needn't transmit downward via chains of metaphysical determination. For the corresponding kind of normative justification may not so transmit. But the argument comes with a qualification: explanations of normative facts may not be subject to the condition in (NEJ) in all contexts. I'll now argue that this qualification doesn't render my argument ad hoc or deprive it of its interest or significance.

If correct answers to the question of why some normative fact obtains aren't constrained by (NEJ) in all contexts, then my argument against the Normative Transmission Principle is compatible with a distinct mixed principle:

Metaphysical Transmission Principle: For any normative fact N and any non-normative facts A and B, if N obtains (in part) because_N B obtains and B obtains (in part) because_M A obtains, then N obtains (in part) because_M A obtains.³⁴

The same goes for a variant where the first conjunct of the antecedent uses 'because_M'. First-order normative theories aim to identify the most basic right- and wrong-making features of actions in the normative order. Such features can fail to coincide with various metaphysically more ultimate grounds of particular rightness and wrongness facts. But then a condition like (NEJ) may not apply in contexts where the question 'Why is experience E good?' or 'In virtue of what is E good?' concerns some metaphysically more ultimate grounds of [E is good].³⁵ When *that* is the question, [E is good] could be grounded, in some non-rigged-up sense, in [The subject of E is in brain state B], and in *that* sense explained by it. For a partially analogous case, note that if the question at issue is 'Why is the subject of this brain scan having a good experience?', a correct answer might be 'Because she is in brain state B'. Such explanations aren't subject to (NEJ). We shouldn't then expect that normative explanation proper must transmit downward via such a grounding relation in accordance with the Normative Transmission Principle.³⁶ This is so irrespective of whether normative explanation arises from metaphysical or generic grounding plus (NEJ), or in some other way.

In case claiming that the Metaphysical Transmission Principle can be true although the Normative Transmission Principle is false seems like too much nuance, note what might hang on it. Insofar as the sort of normative explanations which first-order theories aim to formulate and defend are subject to a constraint which doesn't apply to bare metaphysical explanation, normative theory enjoys a kind of explanatory autonomy with respect to general metaphysical inquiry. For instance, if a normative theory has designated F as a right-making feature, but metaphysics reveals that F-facts are grounded in G-facts, it doesn't follow that the normative theory must now be extended to say that G is a right-making feature. This is compatible with thinking that a properly explanatory first-order normative theory

³⁴What I say about this principle should go also for the principle Berker (2018, p. 751) calls '(Tran_{met/nor})'. The antecedent of that principle is a ground-theoretic parallel of the Metaphysical Transmission Principle and its consequent concerns grounding in the "non-rigged-up" sense discussed in §2 above.

³⁵It is a good question whether the idea that [The subject of experience E is in brain state B] can metaphysically ground [E is good] without normatively explaining it is a special case of some more general distinction. One option is to distinguish facts that explain some fact F from facts that supply further relevant information about facts that explain F but don't thereby themselves explain F (cf. Litland, 2013). Perhaps [The subject of E is in B] can supply further relevant information about [E is an experience of pleasure] without thereby normatively explaining, in the relevant sense, why E is good. Unfortunately I cannot pursue this further here.

³⁶In a ground-theoretic framework this means denying that for any A and B, if A metaphysically grounds B, then A normatively explains B. The Grounding-Explanation Link stated in (G3) above must then be either rejected or modified in some way. This is a natural upshot of (NEJ).

is also a theory in normative metaphysics (Berker, 2018, p. 770). What follows is that the relevant sort of normative explanations aren't generic metaphysical explanations with a normative explanandum, not that they aren't also a kind of metaphysical claims. For all I say, normative explanations may (or may not) be metaphysical explanations under a constraint.

Rejecting the Normative Transmission Principle while leaving the door open to the Metaphysical Transmission Principle is therefore not ad hoc if 'because_N' and 'because_M' express explanations that are subject to different constraints. This result isn't unique to (NEJ). Among ground-theoretic approaches to normative explanation, one would expect grounding pluralists who think that there is a *sui generis* normative grounding relation to deny the Normative Transmission Principle. Grounding monists can also deny the principle if they think that, given the distinctive features of normative facts, the single fundamental kind of grounding relation may yield normative explanation only subject to a condition like (NEJ). Those grounding monists who would still insist that we have a normative explanation of a sort in which first-order normative theories traffic whenever a normative fact is grounded in other facts need in any case to tell us how they would address the considerations that support (NEJ). Even if a grounding relation yields *an* explanation of why normative facts obtain, it may not be of a sort which first-order normative theories aim to formulate and defend.

Why accept that correct answers to the question of why some normative fact obtains may not be subject to (NEJ) in all contexts? Nothing in our general thinking about explanation rules it out: since why-questions ask for different kinds of reasons in different contexts, why not think that normative why-questions sometimes ask for justifications but sometimes don't? This supposition is supported also by an independently plausible general view of why-questions. The view is that why-interrogatives are context-sensitive, in the sense that "one and the same interrogative sentence can be used to ask different why-questions in different contexts of utterance" (Skow, 2016, p. 62). To borrow an example from van Fraassen (1980, p. 131), consider (A1) and (A2) as answers to 'Why is the porch light on?':

- | | |
|------|--|
| (A1) | The porch light is on because the porch switch is closed and the electricity is reaching the bulb through that switch. |
| (A2) | The porch light is on because we are expecting company. |

The truth-values of (A1) and (A2) as answers to 'Why is the porch light on?' may vary with context. When we are interested in the human intentions or expectations that led to the depressing of the switch, (A2) can qualify as a correct answer to the question but (A1) won't. But in certain contexts the reverse is true. Imagine, for instance, that we are rewiring the house and upon seeing the porch light on we fear that we have caused a short circuit that bypasses the porch light (van Fraassen, 1980, p. 131).

We can explain this if the question which 'Why is the porch light on?' is used to ask may vary with context. In this example, what varies is what kind of causes of the porch light's being on count as possible answers to the question. In the first kind of context, the contents of possible answers are restricted to causes involving human intentions and expectations. In the second kind of context, the contents of possible answers are restricted to causes other than human intentions or expectations. This implies contextual variation in what 'Why is the porch light on?' means, assuming that the meaning of a why-question in some way depends on what propositions correspond to a possible answer.³⁷

³⁷Another issue arising from the semantics of why-questions is that the class of possible answers may be relative to a contrast space. Walden (2016) argues that such contrast-relativity has important implications for ethical explanation.

Giving clear criteria for when a response counts as a “possible answer” is a complicated topic in the semantics of questions. As far as I can tell, however, this minimal assumption is widely accepted.³⁸

The source of the context-sensitivity of why-interrogatives is controversial. Van Fraassen proposes that it has to do with what “relation of explanatory relevance” is “salient” in the context (van Fraassen, 1980, Chapter 5). Skow proposes that it derives from contextual variation in the domain of quantification (over reasons why some fact occurs) (Skow, 2016, p. 63). No doubt other views are possible, too. But whatever its source may be, the context-sensitivity of why-interrogatives generates a straightforward explanation of why why-questions ask for different kinds of reasons in different contexts. Some contexts restrict the contents of possible answers to causes, or certain kinds of causes. In other contexts, only grounds and not causes qualify. In some contexts, only certain kinds of grounds and not others qualify. In yet other contexts, only functions, or motives, or purposes, qualify.

If questions as to why a normative fact obtains sometimes request reasons that are (also) justifications, then there will be some contexts where only justifications and not other sorts of reasons qualify as possible answers to such a question. The context-sensitivity of why-interrogatives thus supports a claim I made in §3: explanations of normative facts which are subject to (NEJ) discriminate among nodes in a chain of metaphysical determination. Nodes which don’t meet that condition are excluded from counting as correct normative explanations. If [E is an experience of pleasure] and [The subject of E is in brain state B] can differ justificatorily, then it can happen that (E1) is true but (E3) is false in contexts where a correct answer to ‘Why is E good?’ must be a justification, even if both can be true in contexts where possible answers to ‘Why is E good?’ include any metaphysical grounds of [E is good]. I simply suggest that these are distinct sorts of contexts and that normative theorizing at least often, if not always, generates the former sort of contexts.

We can now see that explanatory discrimination between pairs of facts based on whether those facts differ with respect to normative justification is a special case of a more general phenomenon. Suppose the salad I had for lunch was good for me. Eating it might be good for me because it contains carrot, or because it contains beta-carotene, or because it contains essential nutrients for humans, or because it is healthy. If the different non-normative facts about my salad differ justificatorily, then not all of them may qualify as possible answers to ‘Why was eating the salad good for me?’ in contexts where the possible answers are restricted to justifications of one or another sort. But this isn’t the only kind of discrimination among the different facts about my salad which the context-sensitivity of why-interrogatives might generate. For instance, in contexts that restrict the possible answers to ‘Why was eating the salad good for me?’ to nutrients, ‘The salad is good for me because it contains beta-carotene’ can qualify as a correct explanation but ‘The salad is good for me because it contains carrot’ and ‘The salad is good for me because it is healthy’ don’t.

The salad example is useful in clarifying another feature of my proposal regarding explanatory discrimination. The example makes vivid that in contexts that focus on normative justification, there might be no non-arbitrary way to identify a unique fact that provides the “real” explanation of why eating the salad is good for me. So it is important to see that my argument doesn’t require otherwise. If normative explanation screens out some nodes, that doesn’t mean that it must or will screen out all but one node. Nothing in principle stands in the way of acknowledging the possibility of correct normative explanations at more than one level. Which such possibilities obtain will again be determined by the details of one’s normative theory. For instance, which facts about my salad would go some way towards justifying certain positive responses to eating it, and not just

³⁸See, e.g., the survey in Cross and Roelofsen (2018).

constitute evidence that it has features which would do so, will be determined in part by the details of one's theory of well-being.

The context-sensitivity of why-interrogatives is also compatible with an objective notion of explanation. Once context has done its bit to determine which why-question is being asked in the first place, whether something is the content of a correct answer to the question at issue needn't be constrained by any further facts about the interests or background beliefs of the given audience. The corresponding explanations can also well be relations between worldly facts. Nor does van Fraassen's porch light example by itself undermine this kind of objectivity. The example suggests that our interests are one contextual factor that can play a role in determining what question a why-interrogative is used to ask in a given context of utterance. However, in the case of first-order normative theories, an interest of the relevant sort needn't amount to anything more than a characteristic explanatory goal of such theories. I have suggested that such theories normally aim to formulate and defend explanations which are objective in a certain sense and meet the justificatory condition in (NEJ).

I'll close with an issue left lingering in §2. The issue is whether the Normative Transmission Principle is established if the following principle holds by necessity: for any normative fact *N* and any non-normative facts *A* and *B*, if *B* normatively grounds/explains *N* and *A* metaphysically grounds/explains *B*, then *A* thereby explains *N*, in a "non-rigged-up" sense. This depends on whether the explanation of *N* by *A* is of a sort which first-order normative theories aim to formulate and defend. Grounding monists accept the principle if they accept Berker's (*Tran_{nor/met}*). Litland argues that grounding pluralists can also say that chaining metaphysical and normative grounding gives rise to explanation (Litland, 2018, p. 103). For instance, if [E is an experience of pleasure] normatively grounds [E is good] and [The subject of E is in brain state B] metaphysically grounds [E is an experience of pleasure], then [The subject of E is in B] "metaphysically/normatively grounds" [E is good], and thereby is an explanation of why E is good.³⁹

My position on this issue is easy to anticipate by now. Even if [The subject of E is in B] is *an* explanation of why E is good and this is an explanation in a non-rigged-up sense, it doesn't follow that it meets the condition in (NEJ) or is otherwise of a sort which first-order normative theories seek. If normative explanations in this sense work on the back of a non-rigged-up generic relation, they do so under a constraint. So, I see little reason to expect that the kind of explanation which can be generated merely by chaining together normative and metaphysical explanation must be of the requisite sort. This way of securing the Normative Transmission Principle would require a fuller alternative account of what kind of explanations first-order normative theories traffic in – one which addresses the considerations that support the idea that such normative explanations must often be justifications.

5 | CONCLUSION

I have argued that normative explanation of a sort that first-order normative theories are concerned with fails to satisfy the Normative Transmission Principle. A true normative explanation may fail to transfer to lower-level non-normative facts via chains of metaphysical determination of the sort found in examples like *Pleasure*. I have offered this argument to resolve an impasse in existing discussions of normative explanation. I have sought to base it on assumptions which strike me as either independently plausible or else relatively modest and defensible. The account I have offered aims to reconcile

³⁹By 'A metaphysically/normatively grounds B' Litland means that B can be derived from A via subsumption of A under metaphysical and normative laws.

the popular idea that particular normative facts have an ultimate ontological basis in some fairly low-level non-normative facts with the idea that those low-level non-normative facts may not be facts which normatively explain why the right and the good things are right and good. No doubt more needs to be said about many of the issues raised in the paper, including the very notion of normative justification. Even so, I hope that its results will enjoin further reflection on how normative explanation may be similar to and different from explanations in other domains.⁴⁰

ORCID

Pekka Väyrynen  <https://orcid.org/0000-0003-4066-8577>

REFERENCES

- Baker, D. (2021). If you're quasi-explaining, you're quasi-losing. *Oxford Studies in Metaethics*, 16.
- Berker, S. (2018). The unity of grounding. *Mind*, 127, 729–777. <https://doi.org/10.1093/mind/fzw069>
- Berker, S. (2019). The explanatory ambitions of moral principles. *Noûs*, 53, 904–936. <https://doi.org/10.1111/nous.12246>
- Cullity, G. (2018). *Concern, respect, and cooperation*. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198807841.001.0001>
- Cross, C., & Roelofsen, F. (2018). Questions. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2018 Edition). Retrieved from <https://plato.stanford.edu/archives/spr2018/entries/questions/>
- Dancy, J. (2004). *Ethics without principles*. Oxford: Clarendon Press. <https://doi.org/10.1093/0199270023.001.0001>
- DePaul, M. (1987). Supervenience and moral dependence. *Philosophical Studies*, 51, 425–439. <https://doi.org/10.1007/BF00354046>
- Enoch, D. (2019). How principles ground. *Oxford Studies in Metaethics*, 14, 1–22. <https://doi.org/10.1093/oso/9780198841449.003.0001>
- Fine, K. (2012). Guide to ground. In F. Correia & B. Schnieder (Eds.), *Metaphysical grounding* (pp. 37–80). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139149136.002>
- Fogal, D., & Risberg, O. (2020). The metaphysics of moral explanations. *Oxford Studies in Metaethics*, 15, 168–192.
- Hyman, J. (2015). *Action, knowledge, and will*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198735779.001.0001>
- Jackson, F., & Pettit, P. (1990). Program explanation: A general perspective. *Analysis*, 50, 107–117. <https://doi.org/10.1093/analysis/50.2.107>
- Jenkins, C. (2008). Romeo, René, and the reasons why: What explanation is. *Proceedings of the Aristotelian Society*, 108, 61–84. <https://doi.org/10.1111/j.1467-9264.2008.00236.x>
- Kovacs, D. M. (2017). Grounding and the argument from explanatoriness. *Philosophical Studies*, 174, 2927–2952. <https://doi.org/10.1007/s11098-016-0818-9>
- Krämer, S., & Roski, S. (2017). Difference-making grounds. *Philosophical Studies*, 174, 1191–1215. <https://doi.org/10.1007/s11098-016-0749-5>
- Lange, M. (2018). Transitivity, self-explanation, and the explanatory circularity argument against Humean accounts of natural law. *Synthese*, 195, 1337–1353. <https://doi.org/10.1007/s11229-016-1274-y>
- Lipton, P. (2004). *Inference to the best explanation* (2nd ed.). London: Routledge. <https://doi.org/10.4324/9780203470855>
- Litland, J. E. (2013). On some counterexamples to the transitivity of grounding. *Essays in Philosophy*, 14, 19–32. <https://doi.org/10.7710/1526-0569.1453>

⁴⁰The material in this paper began life as a paper on whether normative explanation is transitive. I received valuable feedback on versions of that paper from audiences at the Universities of Bristol, Groningen, Helsinki, Oxford, Stockholm, Virginia, and Wuhan, as well as from Derek Baker, David Enoch, Daniel Fogal, Stephanie Leary, Jonas Olson, Olle Risberg, Henrik Rydén, Justin Snedegar, Tuomas Tahko, and anonymous reviewers for several journals. My thanks to them all. Subsequently I realized that the paper didn't need to be framed around the issue of transitivity; here is the result. I am especially grateful to an anonymous reviewer for *PPR* for sharp and rigorous comments which led to many improvements.

- Litland, J. E. (2018). In defense of the (moderate) disunity of grounding. *Thought*, 7, 97–108. <https://doi.org/10.1002/tht3.273>
- Maurin, A. (2019). Grounding and metaphysical explanation: It's complicated. *Philosophical Studies*, 176, 1573–1594. <https://doi.org/10.1007/s11098-018-1080-0>
- Nolan, D. (2014). Hyperintensional metaphysics. *Philosophical Studies*, 171, 149–160. <https://doi.org/10.1007/s11098-013-0251-2>
- Pallies, D. (2020). An honest look at hybrid theories of pleasure. *Philosophical Studies*. Advance online publication. <https://doi.org/10.1007/s11098-020-01464-5>
- Rosen, G. (2017). Ground by law. *Philosophical Issues*, 27, 279–301. <https://doi.org/10.1111/phis.12105>
- Ruben, D. (2012). *Explaining explanation*. Updated and expanded second edition. Boulder: Paradigm Publishers.
- Schaffer, J. (2012). Grounding, transitivity, and contrastivity. In F. Correia & B. Schnieder (Eds.), *Metaphysical grounding* (pp. 122–138). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139149136.005>
- Schnieder, B. (2011). A logic for 'because'. *The Review of Symbolic Logic*, 4, 445–465. <https://doi.org/10.1017/S1755020311000104>
- Shafer-Landau, R. (2003). *Moral realism*. Oxford: Oxford University Press. <https://doi.org/10.1093/0199259755.001.0001>
- Skow, B. (2016). *Reasons why*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198785842.001.0001>
- Skow, B. (2017). Against understanding (as a condition on explanation). In S. Grimm (Ed.), *Making sense of the world* (pp. 209–231). Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780190469863.003.0011>
- Smith, M. (2000). Does the evaluative supervene on the natural? In R. Crisp & B. Hooker (Eds.), *Well-being and morality* (pp. 91–114). Oxford: Clarendon Press.
- Strandberg, C. (2008). Particularism and supervenience. *Oxford Studies in Metaethics*, 3, 129–158.
- Strevens, M. (2013). No understanding without explanation. *Studies in History and Philosophy of Science Part A*, 44, 510–515. <https://doi.org/10.1016/j.shpsa.2012.12.005>
- van Fraassen, B. C. (1980). *The scientific image*. Oxford: Clarendon Press. <https://doi.org/10.1093/0198244274.001.0001>
- Väyrynen, P. (2009). Normative appeals to the natural. *Philosophy and Phenomenological Research*, 79, 279–314. <https://doi.org/10.1111/j.1933-1592.2009.00279.x>
- Väyrynen, P. (2019). Normative explanation and justification. *Noûs*. Advance online publication. <https://doi.org/10.1111/nous.12283>
- Walden, K. (2016). The relativity of ethical explanation. *Oxford Studies in Normative Ethics*, 6, 189–214. <https://doi.org/10.1093/acprof:oso/9780198790587.003.0010>
- Wedgwood, R. (2017). *The value of rationality*. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198802693.001.0001>
- Woodward, J. (2003). *Making things happen*. Oxford: Oxford University Press. <https://doi.org/10.1093/0195155270.001.0001>
- Wygoda Cohen, S. (2020). Not all partial grounds partly ground: some useful distinctions in the theory of grounding. *Philosophy and Phenomenological Research*, 100, 75–92. <https://doi.org/10.1111/phpr.12524>
- Yablo, S. (1992). Mental causation. *Philosophical Review*, 101, 245–280. <https://doi.org/10.2307/2185535>
- Zangwill, N. (2018). Epistemic/non-epistemic dependence. *Noûs*, 52, 836–857. <https://doi.org/10.1111/nous.12197>