



UNIVERSITY OF LEEDS

This is a repository copy of *Normative Explanation and Justification*.

White Rose Research Online URL for this paper:
<https://eprints.whiterose.ac.uk/143163/>

Version: Accepted Version

Article:

Väyrynen, P orcid.org/0000-0003-4066-8577 (2021) Normative Explanation and Justification. *Noûs*, 55 (1). pp. 3-22. ISSN 0029-4624

<https://doi.org/10.1111/nous.12283>

© 2019 Wiley Periodicals, Inc. This is the peer reviewed version of the following article: Väyrynen, P (2019) Normative Explanation and Justification. *Noûs*. ISSN 0029-4624, which has been published in final form at <https://doi.org/10.1111/nous.12283>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

NORMATIVE EXPLANATION AND JUSTIFICATION*

Pekka Väyrynen

University of Leeds

Abstract: Normative explanations of why things are wrong, good, or unfair are ubiquitous in ordinary practice and normative theory. This paper argues that normative explanation is subject to a justification condition: a correct complete explanation of why a normative fact holds must identify features that would go at least some way towards justifying certain actions or attitudes. I first explain and motivate the condition I propose. I then support it by arguing that it fits well with various theories of normative reasons, makes good sense of certain legitimate moves in ordinary normative explanatory discourse, and helps to make sense of our judgments about explanatory priority in certain cases of normative explanation. This last argument also helps to highlight respects in which normative explanation won't be worryingly discontinuous with explanations in other domains even though these other explanations aren't subject to the justification condition. Thus the paper aims not only to do some constructive theorizing about the relatively neglected topic of normative explanation but also to cast light on the broader question of how normative explanation may be similar to and different from explanations in other domains.

1. Introduction

Normative explanations are explanations of why things have the normative features they do. That a policy is fair because it allows access to higher education to be distributed in a manner that is sensitive to brute luck in circumstances of birth and upbringing is an example of a putative explanation of a particular normative fact. So is the claim that eating dark chocolate is good in virtue of how it is pleasurable and healthy.¹ Such claims regarding why things are

* Thanks are due to an anonymous reviewer for this journal and audiences at the Philosophical Society of Finland Colloquium 2018, Normativity and Reasoning VI workshop at NYU Abu Dhabi, University of Leeds, Metaethical Seminar at L'École normale supérieure and Institut Jean Nicod, University of St. Andrews, 2018 Bled Philosophical Conference, Central European University, and Durham University. Their questions and comments helped to improve the paper significantly. Thanks also to the Faculty of Arts, Humanities, and Cultures at the University of Leeds for awarding me study leave during which this paper was first drafted.

¹ I use 'explains' as an inverse of 'because' and 'in virtue of': A explains B just in case B holds because A holds, or in virtue of A holding. I use 'normative' as shorthand for 'normative and evaluative'.

wrong, good, unfair, and the like, are ubiquitous in ordinary normative discourse and in moral, political, and legal theory. The question of what makes for a successful normative explanation is as central to normative inquiry as the question of what makes for a successful scientific explanation is to scientific inquiry. But, whereas there is a great deal of research on scientific explanation, and an increasing amount of work on metaphysical explanation, little exists on the nature of normative explanation or its relationship to explanations in other domains. This silence might merely reflect a reasonable preference on part of normative theorists to engage with first-order issues without getting sucked into the metaphysics of the normative. But we shouldn't assume that this reflects implicit agreement among normative theorists concerning the nature of normative explanation. It might merely mean that normative explanatory talk has been fairly unreflective and unrefined.

This paper argues that normative explanation is subject to a justification condition: a correct complete answer to the question why a given normative fact obtains must identify features that would go at least some way towards justifying certain responses. Normative explanations are (perhaps among other things) justifications: at least some explanatory reasons *why* a normative fact holds must provide normative reasons *for* certain responses, or be features in the light of which those responses are apt or fitting or the like. If that is right, then the aim of normative theories to explain why the good and the right things are so and their aim to justify actions and reactions needn't be seen as distinct enterprises. Nor will there be need to interpret questions such as 'Why ought I to ___?' and 'Why is ___ good?' as systematically ambiguous between explanatory and justificatory readings. I'll also suggest that these local features of normative explanation can be at least partly accounted for in terms of certain general features of explanation. If so, the justification condition needn't make

normative explanation worryingly discontinuous with explanations in other domains, and we'll be able to see how they all might deserve the common label 'explanation'. Given the paucity of existing work on normative explanation, my aim is less to argue against any particular opponent than to do some constructive theorizing about normative explanation. Those already sympathetic to the kind of claim I'll defend may take my discussion as an attempt to spell out such a claim in some detail. (As we'll see, the details aren't straightforward.) Those more dubious about a tight connection between normative explanation and justification will, I hope, find reason to reconsider.

2. The Justification Condition on Normative Explanation

The question how normative explanation and justification are related is rarely raised. This is curious. Questions of why certain things are good and bad, or what we ought to do, can be interpreted as asking for explanations or as asking for justifications. It seems crucially important to determine when such questions request one or the other or both. When the relationship does get addressed, normative explanation and justification are sometimes identified.² This is far from obviously right, however. For instance, they have different relations: what is explained are normative and evaluative facts, whereas normative justification may take as its objects actions, attitudes, policies, norms, rules, and so on. Insofar as there is an interesting connection between things that figure in normative explanations and things that figure in normative justifications, it isn't immediately clear what that connection is or

² Wedgwood writes that "explanatory characterizations" of normative reasons "associate reasons with a *justificatory story*—that is, with a story that *explains* the truth about which action or attitude one has, all things considered, most reason to do" (Wedgwood 2017: 91). Hyman explores ways in which explanation and justification are closely related, and argues that justification is a particular kind of explanation, namely of why something is just or right (Hyman 2015: ch. 6). Elstein suggests that normative explanation "coincides" with justification (Elstein ms).

whether such a connection holds uniformly when the facts being explained are evaluative facts, such as facts about what is good, and when they are normative facts, such as facts about what we ought to do.

It is plausible that there is some interesting and close connection between normative explanation and justification. Particular normative and evaluative facts aren't brute. If an agent ought to ϕ , there must be some fact or facts that explain why she ought to ϕ . Similarly, if x is good, there must be some fact or facts that explain why it is good. But normative and evaluative facts about what we ought to do and what is good also imply that certain responses are appropriate, apt, fitting, merited, called for, or warranted—in short, justified.³ What I take to be attractive is that at least one fact among those that explain why I ought to keep my talk to its allotted time should provide normative reason to do as I ought, or make it fitting to intend to do so, or make it appropriate to disapprove of me if I go on for longer, or the like—at least in some respect or to some extent.⁴ If no such responses were justified in any respect or extent, why ought I to keep my talk to time?⁵ Similarly, at least one fact among those that explain why helping a homeless shelter is good (whether *pro tanto* good, morally good, good *simpliciter*, or what not) should go at least some way towards justifying some such responses as promoting, commending, or cherishing help to the shelter. If some such responses weren't

³ It is no accident that normative reasons are also called “justifying reasons” (e.g. Darwall 1983).

⁴ One might worry that the idea articulated here is too close to obvious or trivial. For whatever one thinks about the connection between explanations of evaluative facts and normative justification, one might find it hard to see how there could, even in principle, fail to be a connection between an explanation of why I ought to ϕ and a justification for ϕ -ing. The discussion that follows will indicate respects in which substantial assumptions are required for connecting even explanations of why one ought to ϕ to the justification for ϕ -ing.

⁵ If there are many different oughts (moral, prudential, legal, etc.), then something that I ought to do in one respect (e.g. prudentially) might, of course, not be justified in another (e.g. morally). My points here don't require that whenever some agent ought to ϕ , there is some fact or facts that count in favor of ϕ -ing. (Some “oughts of rational requirement” might not require this, for instance.)

apt or fitting in the light of the reasons why helping the shelter is good, why count helping it as good in the first place? So there is at least some initial motivation for a justification condition on normative explanation. I'll first articulate some assumptions implicit in these motivations, and then formulate the condition I'll defend.

The above examples suggest that there is a close connection between the obtaining of a normative fact and certain actions' or reactions' being justified (apt, fitting, or the like) in response to the state of affairs (or its constituent object) in question. Normative facts embed or imply ideals or standards for actions or attitudes. Adapting a fairly common way of understanding normative standards, a normative fact specifies (implicitly or explicitly) a condition that is to be met by someone or something (perhaps under certain other conditions), and against which it is at least in principle possible to appraise the thing in question positively or negatively on the basis of whether it meets the condition in question.⁶ In this way, normative facts play a role in directing or regulating our actions and attitudes. It seems hard to avoid characterizing normative facts without reference to what responses are apt under what conditions.⁷ Take, for instance, the relation between pleasure and goodness. That an experience is pleasant doesn't *ipso facto* mean that certain responses are apt or fitting, or that anyone has reason to react in any particular way. But the fact that the experience is good does *ipso facto* mean this. The contrast remains at least so as long as pleasure and goodness are conceptually (even if not metaphysically) distinct aspects of situations. That an experience is pleasant isn't itself a normative fact, but a normatively significant non-normative fact, one

⁶ See e.g. Copp (1995: 19). While standards are typically taken to be general, my formulation in the text is meant to leave room also for particularist views that don't countenance general normative principles.

⁷ If someone thinks that facts about what one ought to do or about what is good are no more tightly linked to the appropriateness of certain responses than facts about what is pleasant, I would have to hear their characterization of normative facts to assess to what extent the rest of my discussion might still apply.

that contributes to the obtaining of a normative fact. Normative properties can be distinguished from normatively significant non-normative properties in part by how they relate to certain responses.

The justification condition on normative explanation which I propose takes this connection between normative facts and their consequences for what responses are justified to be tight enough for the latter to be included as part of what gets explained in normative explanation.⁸ This should seem plausible to those (many) who think that a distinctive feature of normative concepts is that their satisfaction *ipso facto* means that someone or other has reason to respond in certain ways or that certain responses would be fitting or appropriate.⁹ (The idea should suit also expressivists and others who have no truck with normative facts in any metaphysically robust sense.) It should also seem plausible to those who think the same about normative properties. On either view, it is plausible that an explanation of why x is F isn't an explanation of why x is *good* unless the explanation goes some way towards justifying responses like promoting, cherishing, or protecting. It is similarly plausible that an explanation of why x is F isn't an explanation of why x is right unless it goes some way towards justifying responses like performing x , resenting those who don't do x , or the like.

This idea can be turned into a general justification condition as follows:

⁸ Thanks to Geoff Sayre-McCord for a helpful discussion here.

⁹ This allows but doesn't require that the fact that something is good or the fact that something is right is itself a reason. (See section 4.) Nor does it require a reasons-first or fittingness-first approach to normativity.

(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.¹⁰

Here ‘N’ holds the place for a that-clause that expresses a fact.¹¹ Instantiated for one of my examples above, (NEJ) says that a correct complete normative explanation of why helping a homeless shelter is good (why that normative fact obtains) must identify features that would go at least some way towards justifying certain responses either to the fact that helping the shelter is good or to the relevant constituent, helping the shelter.¹² The substantial claim of (NEJ) is that the truth of normative explanations of the form ‘N because ____’ requires that the features that replace the blank include features in the light of which it would be apt or fitting to act or react in a certain way or which would provide reasons to do so. The ‘would’ is meant to avoid problems such as “solitary goods”: good states of affairs that entail that there are no past, present, or future subjects to adopt the fitting responses (Bykvist 2009). A putative normative explanation that fails (NEJ) is a failed normative explanation: false, incorrect, or no explanation at all, depending on our ideology of explanation. If we think of an explanation of why N obtains as the content of an answer to the question ‘Why N?’, then (NEJ) is equally a constraint on correct complete answers to the questions of why particular normative facts obtain.¹³ I restrict (NEJ) to particular normative facts because singular explanation may differ

10 (NEJ) is an attempt to spell out a very quick suggestion in Väyrynen (2013: 173), inspired by conversations in the distant past with Christian Coons and Daniel Elstein. Thanks also to Walter Pedriali for helpful suggestions about formulating (NEJ).

11 Sometimes I’ll take the liberty of putting in for ‘N’ a term that denotes a fact. Nothing deep hangs on this.

12 Whether the proper object of a response is a fact or its constituent object may vary across different responses.

13 For reasons to focus on answers to why-questions in theorizing about explanation, see e.g. Skow (2016: ch. 2).

in important ways from explanations of laws. (The two are commonly kept separate in discussions of scientific explanation.)

(NEJ) makes a general claim about normative explanation. It doesn't threaten to overgeneralize to explanations of non-normative facts. We don't expect an explanation of why H₂O boils at 100°C at sea level to identify features that would (in the relevant sense) justify any particular response to that fact, and the same goes for the non-normative fact that an experience is pleasant. But within the normative domain I have so far tried to make (NEJ) plausible only in the special cases of facts about what we ought to do and facts about what is good. Generalizing from these cases to (NEJ) in full generality raises various complications.

First, to just what range of facts (NEJ) applies will depend on what facts count as normative. Views vary, for instance, on whether the disjunctive fact that stealing is wrong or grass is green, or the negative fact that there is no reason not to do something, are normative facts. I cannot settle these matters here, so my defense of (NEJ) will remain hostage to how the contours of the normative and the evaluative end up being drawn. Here I aim to defend (NEJ) with respect to a paradigmatic range of particular normative facts.

Second, it is unclear what responses are apt to such normative facts as that one has normative reason to do something, or that something is (merely) permissible, or that something is a good instance of a kind, or that something is supererogatory? A systematic pairing of normative properties and apt responses to them is again beyond the scope of one paper. But I remain optimistic. It doesn't seem implausible that if the Wegner Wishbone is a good dining chair, it is fitting to commend it for use as a dining chair, or at least that if there is reason to commend things for use as dining chairs, there is reason to commend the Wegner Wishbone. Nor is it implausible that if spending a Sunday reading fiction is merely

permissible, it is appropriate to tolerate that activity. And if I have a reason to go and see a friend's band, it seems fitting to recommend doing so, at least provided that there is no decisive reason to do something else. Beyond this, we recognize plenty of room for substantial debate about which sorts of actions or reactions, of which subjects, under what conditions, are justified responses to instances of different normative properties. Whatever else these complications may end up showing, they should at least ensure that (NEJ) isn't trivial.

Nothing in our thinking about explanation in general rules out (NEJ). Why-questions ask for reasons, and reasons are what figure in the contents of correct answers to why-questions. When R is the content of a correct answer to 'Why Q?', R is at least part of a complete explanation of why Q. (Read 'Q because R' and 'R is a reason why Q' accordingly.¹⁴) We already recognize that in different contexts, why-questions ask for different sorts of reasons: sometimes for a cause, sometimes a ground, sometimes a motive or purpose. There is then no principled reason why normative why-questions couldn't ask for reasons that justify. A question like 'Why should I pay my taxes?'—let alone the 'Why be moral?' question—is naturally heard as asking for a justification for a requirement on action. (NEJ) makes this compatible with the truism that why-questions ask for reasons without implying that normative why-questions are systematically ambiguous between requests for explanation and requests for justification.

(NEJ) concerns complete explanations of normative facts. But my examples of normative explanations have featured partial rather than complete explanations. For instance,

¹⁴ 'R [partly] explains why Q' is factive with respect to 'Q'. But it is controversial whether 'R is a reason why Q' is also factive with respect to Q. There can be reasons why I ought to do something even though it isn't the case that I ought to do it (Nebel forthcoming). I'll focus on cases where the fact that's being explained obtains. So I won't need to worry about factivity failures when I use the term 'reason why'.

that E is an experience of pleasure is presumably not the full story about why E is good. It is a matter of controversy in the theory of explanation what a complete explanation of a phenomenon consists in. (It is also a matter of controversy what a complete answer to a why-question consists in.) One issue is what factors are required to be in place for R to be a complete explanation of Q. Another question is which of these factors are parts of such an explanation, and which (if any) play some other role with respect to the complete explanation. For instance, it is controversial whether complete explanation in some way requires general laws or principles. Moreover, even if it does, it is controversial whether laws are parts of complete explanations or only “back” them.¹⁵ It is also controversial how facts that “enable” others to make something the case relate to a complete explanation of it.¹⁶ Are they parts of such explanations, or do they explain why some other facts are parts of such explanations, or do they play some other role altogether? Fortunately I needn’t take a stand on these questions here, and can continue to work with partial explanations. For it would be odd if, for a given normative fact N, no partial explanation of N had to identify features that would go at least some way towards justifying certain responses to it. Again, if none of the reasons why something is good could go any way towards justifying certain positive responses to it, why think it is good in the first place? But if at least one partial explanation of why it is good must identify such features, then a complete explanation must do so—as (NEJ) says.

15 For an overview of this kind of debate, see Hillel-Ruben (2012: ch. 6).

16 On enablers, see Dancy (2004: ch. 3). For a brief discussion of what role enablers might play in normative explanation, see Väyrynen (forthcoming). We’ll also need to settle what role factors that “intensify” reasons to ϕ without themselves being reasons to ϕ play in normative explanations of why there is more reason to ϕ than to ψ .

The notion of explanation which I take to be relevant to (NEJ) is objective in three senses. First, I'll take the fundamental relation of explanation to be facts rather than representations.¹⁷ Those allergic to talk of normative facts are welcome to give a deflationary interpretation of such talk or replace it with their preferred substitute. Second, normative explanation is objective in that its correctness isn't relative to the interests or background knowledge of an audience, at least not once context has done what it may do to influence which why-question an interrogative sentence of the form 'Why Q?' is being used to ask in the context in the first place.¹⁸ First-order normative theories aim to supply explanations that are objective in this sense. For instance, if some form of act-consequentialism is the correct moral theory, then wrong actions are wrong in virtue of failing to maximize the value of outcomes, irrespective of how that relates to our interests or knowledge.¹⁹ Third, normative explanation is objective in that it isn't a constraint on correct normative explanation that it induce understanding in a given audience. Explanations often are in a good position to do so. But if act-consequentialism is the correct moral theory, then wrong actions are wrong in

¹⁷ I take facts to be structured non-representational entities composed of worldly items such as objects, properties or relations, and the like, and individuated by their constituents and their mode of composition. However, if desired, I could accommodate the view that facts are true propositions. Nothing too deep hangs on the first assumption.

¹⁸ It is widely recognized that one and the same interrogative sentence can be used to ask different why-questions in different contexts of utterance (see e.g. van Fraassen 1980: ch. 5). There is less consensus on what makes why-interrogatives context-sensitive in this way (cf. Skow 2016: 62–4). A fuller theory of normative explanation than I can offer here would say more about how such context-sensitivity relates to normative explanation. (NEJ) may imply that normative contexts restrict the domain of possible complete answers to 'Why Q?' so that it must include facts that are capable of justifying the relevant responses. I would see this as a feature, not a bug. Also note that an explanation that is objective in the sense I specify can be subject to certain further kinds of epistemic constraints, such as, perhaps, contributing to the contextually salient explanatory project. I won't rely on this, but can allow it.

¹⁹ The question whether normative explanation is objective in this sense shouldn't be confused with the question whether the normative facts that are being explained themselves reflect epistemic or pragmatic factors. For instance, if moral obligation is evidence-relative or facts about what is right concern what maximizes expected value, their explanation can still be objective in the sense I specify.

virtue of failing to maximize value, irrespective of whether this explanation induces understanding of why the wrong actions are wrong in the given audience.

These assumptions concern correct normative explanations. Among such explanations, some may be better or worse relative to an audience, depending on how well they fit with the audience's interests and background knowledge or how liable they are to induce understanding. If an audience isn't asking a moral question when they ask 'Why is killing bad?', the answer that killing is bad because it fails to maximize value might be a poor answer in that conversational context. But that would make it no less true as an answer to the moral question which this why-interrogative can be used to ask. Distinguishing the notion of a correct explanation from that of a better/worse explanation allows an objective notion of normative explanation to accommodate the idea that normative theories aim also to provide useful practical guidance. Perhaps meeting the epistemic and pragmatic conditions required for useful practical guidance makes a normative explanation a better explanation. It doesn't follow that meeting them is a correctness condition on the explanation. Nor does it follow that justifying actions or reactions requires meeting the sorts of epistemic or pragmatic conditions which would make explanation non-objective. A justification, too, may be correct or sound without always being available for useful guidance.

I'll leave (NEJ) intentionally generic in further respects of interpretation. It sets out a schematic condition that can be filled out in different ways. For instance, I'm happy to allow various views regarding which objective relation of the general type I have specified is at work in normative explanation. Irrespective of whether 'in virtue of' and 'because' stand for grounding, constitution, counterfactual invariance, or something else, (NEJ) merely requires instances of these relations (or the explanations they back) to have a certain justificatory

feature when the facts to be explained are normative facts. So far as (NEJ) goes, normative explanation needn't be distinctive regarding what *kind* of explanation it is, but only regarding what it is an explanation of—namely, facts that *ipso facto* mean that certain responses are apt or fitting, or that someone or other has reason to act or react in some way.

That (NEJ) remains schematic in some respects doesn't mean that we lack independent grip on what it amounts to. Consider some intuitive characterizations of normative justification. Normative justification for ϕ -ing is sometimes understood in terms of a contribution to *closing deliberative questions*, such as whether to ϕ . In that case, (NEJ) requires a correct complete normative explanation to identify features that would contribute to closing such questions.²⁰ This seems plausible. Suppose that an experience E is good because it is pleasant. That E is pleasant contributes to closing such questions as whether to pursue or commend E. Of course, that E is pleasant might not do so if there were no bridge assumption to the effect that pleasure (of at least certain kinds) is good. But in that case the fact that E is pleasant might not partly explain why E is good in the first place. The fact that a glass has a molecular composition M goes no way towards explaining why the glass is fragile in the absence of a bridge law connecting fragility to its categorical base.²¹ Given such a law, having M partly explains why the glass is fragile. (NEJ) makes good sense under this kind of way of thinking about normative justification.

A different way to understand normative justification is that justified responses, requirements, or norms are *authoritative* regarding how to act and react, not arbitrary. (NEJ) would then require a correct complete normative explanation to identify features that would

²⁰ Different views are, of course, available regarding what it takes to close a normative question. For one particularly demanding proposal, see Korsgaard (1996).

²¹ This is of course a controversial view on the metaphysics of dispositions. But I hope it'll do for illustration.

go at least some way towards making directives for certain responses relevantly authoritative. This may amount to different things in different cases. The phenomenology of normative authority normally presents such authority as external to one's will, sentiments, and desires. For instance, suppose it is good to minimize our consumption of plastic and a certain policy would reduce our consumption significantly. That a policy would do so would be authoritative with respect to what policy to promote or commend in a way that stems from outside of our will or desires.

Not all cases are like this, even if the phenomenology suggests otherwise.²² That a standard of etiquette has currency in a community is external to one's will and desires. But the normative authority of such a standard is widely seen as hypothetical on one's ends. Instrumental and prudential normativity are also relevantly authoritative. If something is in my best interest, a directive to pursue it is authoritative in a way that directives to pursue things that aren't in my interest aren't, irrespective of whether the normative authority of self-interest resides in one's will or desires. For the instrumental case, suppose that I have some suitably significant desire that ϕ and that ψ -ing is a necessary means to ϕ , so that I have a reason to ψ . Here (NEJ) would require a correct complete explanation of why I have reason to ψ to identify features that would go at least some way towards making a directive to ψ relevantly authoritative for me. What features those would be depends on what the best theory of instrumental reasons is. For instance, they might be facts that explain why ψ -ing is a necessary means to ϕ . In that case (NEJ) makes good sense. If ϕ is a suitably significant end of mine and R explains why ψ -ing is a necessary means to ϕ , then R meets the condition

²² Daniel Star writes that "the problem of authority is the problem of the possible justification of one being subject to directives originating outside of oneself" (Star 2015: 42). I suggest this isn't quite right. Etiquette directives originate outside of me, and other people may issue me with directives based on facts about what is instrumental to my ends. It doesn't follow that the normative authority of such directives is external to my will or desires.

imposed by (NEJ). For instance, R would explain why not ψ -ing would open me to criticism in the light of my own ends, and to that extent would make a directive to ψ authoritative for me. The cases of etiquette and instrumental reason also show that (NEJ) is neutral on whether normative justification is sometimes categorical or always hypothetical. (NEJ) takes no stand on whether a correct normative explanation of why ϕ -ing would be morally wrong must identify features that provide categorical reasons for action.

3. (NEJ) and Theories of Reasons

In articulating (NEJ) I have already suggested that it is attractive in certain central cases and fits with a common conception of what is distinctive about normative concepts or properties. In the next three sections I'll offer further support for (NEJ). The first piece of support is that (NEJ) falls out of many prominent theories of normative reasons.

One example are theories according to which normative reasons for actions or attitudes are facts that make certain normative facts hold.²³ One such view is that reasons why an agent ought to ϕ are normative reasons for her to ϕ .²⁴ Broome, for instance, holds that a reason for S to ϕ is a fact that plays the 'pro ϕ -ing' role in a weighing explanation of how S ought to act (Broome 2013). We don't need to accept this as an *analysis* of normative reasons to agree that facts that explain why you ought to do a thing are reasons for you to do it. Suppose we do agree with that. A normative reason for something just is a consideration that justifies it, at least to some extent or in some respect. Thus, accounts of normative reasons

²³ This claim is non-trivial. Favoring an act (which is what normative reasons do) and making an act right (which is what a reason why an act is right does) are distinct relations (Dancy 2004: 79–80).

²⁴ See e.g. Alvarez (2010), Broome (2013), and Nebel (forthcoming). Hyman says that a reason to act is an explanation of why it is right to act, but uses a non-objective notion of explanation (2015: 136).

like this imply that facts that explain why you ought to ϕ go at least some way towards justifying ϕ -ing. This also extends beyond 'ought'. A reason why a thing is good is something in the light of which it would be fitting to act or react in certain ways, and thus justifies those responses at least in some respect.²⁵

(NEJ) also falls out of views according to which normative reasons are evidence of a certain type (Kearns and Star 2009; Whiting 2018). If a fact is a reason why you ought to ϕ , it is evidence that there is at least some respect in which you ought to ϕ . (The converse fails. A fact can be evidence that P without being a reason why P.) If normative reasons are evidence, then such a fact is a reason for you to ϕ . But again, normative reasons to ϕ go at least some way towards justifying ϕ -ing. So a fact partly explains why an agent ought to ϕ only if it goes at least some way towards justifying her ϕ -ing.

Even views on which normative reasons are facts that would explain a possible response on the part of some relevant agent which would count as an ideal response may at least fit with (NEJ). For instance, if x 's being F explains why x is good, and something like promotion, admiration, or approval is an ideal response to Fs, then ' x is good because it is F' could be said to identify features that would go at least some way towards justifying certain responses to x .²⁶

In sum, there are several theories of normative reasons which fit well with (NEJ). This gives us reason to accept (NEJ) insofar as these theories of reasons are credible.

²⁵ Raz (1999) analyzes reasons for action as explanations of why it is good that a person acts.

²⁶ It is less clear whether (NEJ) fits in the same way with certain other views, such as that normative reasons are premises of good reasoning.

4. (NEJ) and Normative Discourse

My second piece of support for (NEJ) comes from certain features of our ordinary normative discourse. Bits of ordinary discourse are of course subject to pragmatic influences. But this doesn't mean that they cannot provide evidence that normative explanation (in the kind of objective sense I specified) works as (NEJ) suggests. Consider a Socratic sort of person, one whose life is characterized by a certain sort of intellectual activity.²⁷ Suppose hers is a good life. Suppose it is a life of accomplishment. Suppose it is a life that displays understanding. And suppose that in her life she does things like writing the particular words with their particular meanings (with an understanding of their meanings) which she writes in her papers, giving talks, teaching and mentoring, and so on. (I'll abbreviate these as 'writes the stuff she writes'.)

Why is the Socratic sort's life a good life? Here are some possible answers to this question, construed as requesting a normative explanation:

- (1) The Socratic sort's life is good because it is a life of accomplishment.
- (2) The Socratic sort's life is good because it displays understanding.
- (3) The Socratic sort's life is good because she writes the stuff she writes (etc.).
- (4) Each of (1)-(3).
- (5) Not each of (1)-(3) but more than one of them.

For the present purposes, I needn't deny either (4) or (5). Perhaps the Socratic sort's life is one of accomplishment at least in part because it displays understanding and displays

²⁷ I adapt the example from a discussion of moral supervenience in Smith (2000). It may be a simplistic example, but its simplicity is useful for my purpose of illustrating a certain kind of structure.

understanding at least in part because she writes the stuff she writes. If the facts in the right-hand sides of (1)-(3) form this kind of metaphysical hierarchy, neither (4) nor (5) should entail any objectionable kind of overdetermination of the fact that the Socratic sort's life is a good life. Also assume, for simplicity, that the explanatory facts identified in (1)-(3) exhaust the options, however each would have to be supplemented to constitute a complete explanation. If you already accept (NEJ), I can also grant that each of, or more than one of, (1)-(3) identify features that would go at least some way towards justifying certain responses. In what follows I address those who don't already accept (NEJ). I'll argue that certain legitimate responses to (1)-(3) as putative explanations of why the Socratic sort's life is good point to reasons to accept (NEJ).

Consider (3) as an answer to the question why the Socratic sort's life is good. One legitimate response to this answer would be to ask: what difference does the fact that she writes the stuff she writes make to her life's being good? One answer that would be to the point of this query is that the stuff she writes displays understanding. After all, had it not displayed understanding, it wouldn't have been part of what makes her life good. Either her life wouldn't have been good, or it would have to have been good in some different way.²⁸ If this query is a legitimate response to (3), what does that tell us?

The query isn't a merely epistemic query arising from lack of knowledge regarding the metaphysically lower-level facts reported in (3) or their relation to the higher-level facts reported in (2). Epistemic gaps might suffice to account for other cases, but that needn't be so here. I've read enough of Hume's *Treatise* and Kant's *Critiques* to grasp that they display great understanding. If I can, so can you. More plausibly, the query about (3) suggests that

²⁸ For purposes of illustration, I'll bracket cases where the stuff the Socratic sort does exemplifies some other good-making feature than displaying understanding.

certain facts qualify as explaining normative facts only insofar as they are suitably related to certain other facts that explain normative facts. For it suggests that the fact that the Socratic sort writes the stuff she writes plays a distinctive role in making her life good only insofar as the stuff displays understanding. Since (2) is itself a putative explanation of why her life is good, the idea isn't merely that the fact that her life displays understanding enables the fact that she writes the stuff she writes to explain why her life is good. Nor are (2) and (3) distinct parts of the same complete explanation; they operate at different levels. What I suggest our responses to indicate instead is that (2) is a more basic explanation than (3) in the order of normative explanation, even if (3) may be more basic than (2) in the order of metaphysical fundamentality.

This claim about how (2) and (3) are related involves particular first-order normative assumptions. It assumes that a principle linking understanding to the relevant kind of goodness is normatively more basic than a principle that links writing the stuff the Socratic sort writes to the relevant kind of goodness. This strikes me as a plausible first-order assumption. But the essential point for my purposes is a structural point which is independent of particular first-order normative assumptions. We—or, at any rate, I—don't yet have well-developed conceptual resources for fully explicating this structural point. But let me try to bring out the general idea.

(3) says that the fact that the Socratic sort writes the stuff she writes is a reason why her life is good. It is legitimate to ask why *that* is so. The answer will be a higher-level reason why: a fact (or a set of facts) that is a reason why <the fact that the Socratic sort writes the stuff she writes is a reason why her life is good>.²⁹ Those facts might not be still more

²⁹ I use corner brackets to keep iterations of reasons why readable. Skow (2016: ch. 5) offers a general discussion of higher-level reasons why. Väyrynen (forthcoming) discusses its application to the normative case.

metaphysically fundamental than the fact that the Socratic sort writes the stuff she writes. The answer will consist either in facts that are more basic in normative significance (irrespective of their relative metaphysical fundamentality) or, if we have already hit the bedrock of normative significance, in something like a general principle encoding the normative significance of the fact in question, or perhaps in some transcendental condition.³⁰ Our query about (3) suggests that we don't take the normative significance of the fact that the Socratic sort writes the stuff she writes to be bedrock. Rather we think its normative significance is based on its standing in some suitable (such as causal or constitutive) asymmetric relation to something else of normative significance.

For my point, it doesn't matter whether this "something else" is the fact that the Socratic sort's life displays understanding or that her life is one of accomplishment, or some still other fact. The general form of my suggestion above is that <the fact that the Socratic sort writes the stuff she writes is a reason why her life is good> obtains because [the stuff that the Socratic sort writes is F and <the fact that her life is F is a reason why her life is good>]. What goes in for 'F' will depend on the correct first-order normative theory. It could be that the normative significance of accomplishment is more basic than the normative significance of understanding.³¹ Or it could be the other way around. And given what I said about (4) and (5) above, there may not always be a unique normatively the most basic node in a chain of

³⁰ By a normatively significant fact/property I mean a fact/property that contributes to the instantiation of a normative property. Being right (good, etc.) making and being wrong (bad, etc.) making are two paradigm cases of such normative significance.

³¹ A legitimate response to (2) as an answer to the question why the Socratic sort's life is good would be to ask: what difference does the fact that her life displays understanding make to her life's being good? If our response is that displaying understanding is an important kind of accomplishment, this suggests that we regard the normative significance of understanding as less basic than the normative significance of accomplishment. If our response is merely something like 'Understanding is a really good thing', this suggests we don't take such a view.

metaphysical determination. The way these normative chips in fact fall would be reflected in how the correct first-order normative theory relates the principles which link understanding and accomplishment, respectively, to the relevant kind of goodness.

Suppose the above is broadly on the right track. How is it supposed to support (NEJ)? Recall the query about (3): what difference does it make to whether the Socratic sort's life is good that she writes the stuff she writes? If explaining why the Socratic sort's life is good only required identifying some or other facts that make her life is good, this query shouldn't arise. So the query seems to be asking for something more than just any metaphysical ground of the fact that her life is good. (NEJ) identifies a plausible candidate for this "something more". The fact that the Socratic sort writes the stuff she writes doesn't seem by itself to justify such responses as commending or cherishing the Socratic sort's life. This is why it seemed plausible that an answer that would be to the point of the query about (3) would be something like: the fact that the Socratic sort writes the stuff she writes matters to her life's being good insofar as her life displays understanding. (NEJ) can explain why this would be to the point by saying that (2) identifies features that go at least some way towards justifying the relevant responses.³² Or, given different first-order assumptions, that role might be played by (1) instead. (What normative justifications hold depends on what normative principles hold just as much as what normative explanations hold so depends.) But, either way, the fact that (NEJ) can make sense of legitimate reactions to putative normative explanations of why the Socratic sort's life is good is evidence that at least one of (1)-(3) must identify features that justify certain responses. We aren't satisfied that (3) does so, and move on to consider whether (2) does. But if at least one of (1)-(3) must identify features that would go at least

³² In view of what I said above about first-order normative assumptions, a more precise claim would be that (2) identifies features that at least do better in justifying the relevant responses.

some way towards justifying certain responses, then a complete correct explanation of why the Socratic sort's life is good must do so.

This argument for (NEJ) claims that (3) doesn't *by itself* identify features that would justify the responses that are apt to the fact that the Socratic sort's life is good. This move can be spelled out in two different ways, depending on our other commitments.³³ One is that (3) identifies *no* normative reason to commend or cherish her kind of life, or *no* feature that would make these responses apt or fitting. The other is that (3) does identify a justification, but one that is in some sense *derivative*. On this view, the fact that the Socratic sort writes the stuff she writes is a reason to commend or cherish her life, or would make these responses apt or fitting, but it has such a normative status only insofar as, and because, it is suitably related to some distinct fact that has such a normative status (perhaps, that her life displays understanding).³⁴

I won't try to decide between these options, not least because I doubt intuitions about reasons will help much. So far as (NEJ) goes, the choice between the options depends, among other things, on whether to countenance derivative explanations. To illustrate the structure of a derivative explanation, we might say that the Socratic sort's life is good because she writes the stuff she writes, where this explanation holds insofar as, and because, the stuff she writes displays understanding and her life is good because it displays understanding. Derivative explanations are ruled neither in nor out by (NEJ) as such. The issue matters because

³³ Many further issues about normative reasons arise in this neighborhood which I cannot address in this paper. For one interesting recent discussion of some of the issues I have in mind, see Johnson King (2019).

³⁴ There are many possible ways for a reason to be derivative, besides the kind of constitutive structure that we find in the Socratic sort example, depending on what count as relevantly "suitable" relations. If I have a reason to hydrate myself, then my reason to drink water is, plausibly, a derivative reason thanks to a certain causal relation. If the fact that *x* is right is any reason to do *x* in the first place, it looks like a derivative reason to do *x* thanks to its relation to reasons to do *x* which are overall stronger than reasons not to do *x*. And if teachers have

countenancing derivative explanations would require one of two things: either restricting (NEJ) only to non-derivative explanation and justification, or distinguishing that reading of (NEJ) from one that connects derivative justification to derivative explanation.³⁵ Either way, if (3) identifies derivative reasons to commend or cherish the Socratic sort's life, (3) wouldn't satisfy the non-derivative reading of (3) but would plausibly satisfy the derivative reading of (NEJ) if we countenanced such a thing. So (NEJ) doesn't require us to distinguish between derivative and non-derivative explanations or justifications, but it can accommodate such distinctions.

One might think it isn't plausible that there is any non-arbitrary way to treat (2) but not (3) as identifying facts that would justify certain responses when the facts they feature are metaphysically related in the way they are in the Socratic sort case. This worry is resistible once we clarify what (NEJ) does and doesn't entail. (NEJ) doesn't require (2) to constitute a unique correct explanation of why the Socratic sort's life is good. I granted that (3) can be true, and that (1) can be true. My argument only requires that a complete correct explanation must identify features that would go at least some way towards justifying certain responses, and that not all features that may figure in a correct complete normative explanation are bound to satisfy this requirement. It doesn't entail that explanations that don't by themselves satisfy this requirement are false, so long as some other explanation of the same normative fact does satisfy it. I also noted that first-order assumptions about which normative principles are more basic than others matters to which explanations satisfy (NEJ). It doesn't seem

reason to set work that is appropriate to their students' level of capacity, then first-grade teachers' reason to set work that develops basic literacy and numeracy skills is, plausibly, a derivative reason thanks to a certain kind of specification relation. (Some of these more basic reasons may themselves derive from some still further reasons.) On the face of it, these examples don't require there to be a unique fact from which a derivative reason derives its reasonhood. Nor do I take them to suggest that to classify a reason as derivative is to dismiss it, for instance as "not real" or as incapable of fulfilling certain job descriptions of normative reasons for action.

³⁵ I assume that (NEJ) should in this case feature parallel qualifications on its left-hand and right-hand sides.

arbitrary which principle identifies features whose normative significance is more fundamental: one connecting facts about the stuff that the Socratic sort writes to goodness or one connecting understanding to goodness.

5. (NEJ) and the Euthyphro Dilemma

My third argument for (NEJ) concerns our judgments about the direction of normative explanation in Euthyphro-style cases. In Plato's dialogue *Euthyphro*, Socrates raises the question whether pious things are pious because gods love them or whether gods love them because they are pious.³⁶ The question is widely thought to generate a dilemma against those who endorse the former answer. If gods love the pious, either they do so for reasons independent of their love or not. If they have no independent reason to love the pious, then their love seems arbitrary and it's unclear why their love should be normative or authoritative for us. But if gods love the pious for an independent reason, then what really explains why the pious things are pious are the reasons that explain why gods love them. This dilemma is widely thought to generalize beyond piety and gods' love to normative and evaluative properties in general and the attitudes or stances of any subjects towards instances of such properties. I'll now argue that (NEJ) helps to make sense of our judgments about Euthyphro-style cases, and that insofar as these judgments are plausible, (NEJ) supports them. Interpreting reasonable reconstructions of typical kinds of reasoning in Euthyphro-style cases as relying on (NEJ) may not be the only way of making sense of and supporting our judgments, but I hope to show that it is one good way of doing so.

³⁶ In what follows, I'm not making interpretive claims about the dialogue *Euthyphro*, but only summarize what I take to be a widely shared conception of a metaethical problem which *Euthyphro* can be read as raising. For one representative textbook presentation of the problem, see Timmons (2002: 29–30).

For concreteness of illustration, suppose that the following are necessary truths: any act is right if and only if it maximizes general pleasure (GP, for short); any act is right if and only if it is commanded by God; and God commands all and only what maximizes GP. These necessary biconditionals take no stand on the direction of normative explanation. The possibilities are that actions are right in virtue of maximizing GP and that actions are right in virtue of being commanded by God. These cannot both be correct. (The structure here is different from the example of the Socratic sort.) Overdetermination aside, one putative normative explanation excludes the other (DePaul 1987). If ‘*x* is right because *x* maximizes GP’ is true, then ‘*x* is right because God commands *x*’ is false, and *vice versa*.

Many of us (including some who believe in God) believe that it is the divine command explanation that is false. The typical reasoning is that we reject ‘*x* is right because God commands *x*’ because God’s commands would be arbitrary unless they were based on independent reasons, such as perhaps that the action maximizes GP. Does it make sense to reject ‘*x* is right because God commands *x*’ on the basis that God’s commands would be arbitrary unless they were based on independent reasons? That would make sense if God’s commands wouldn’t in that case normatively justify the responses that are appropriate to right action: doing it, resenting those who don’t do it, or the like.³⁷ And God’s commands wouldn’t do so if they were arbitrary, since then they would lack the requisite normative authority. But if so, then ‘*x* is right because God commands *x*’ fails (NEJ): this putative explanation doesn’t identify features that would go at least some way towards justifying the relevant responses. So (NEJ) helps to make sense of the typical Euthyphro-style reasoning for the falsity of ‘*x* is right because God commands *x*’. Moreover, insofar as that conclusion is

³⁷ That an action is commanded by God would of course be excellent evidence that it is right. But again, the relevant notion of justification isn’t that of the epistemic justification of a normative belief.

plausible, (NEJ) provides a way to support it, by providing a connection between a putative normative explanation being false and its failing to specify features that would justify certain responses.

The above isn't sufficient to establish that 'x is right because God commands x' in fact is false. It isn't fully clear what completion of it which doesn't make God's commands redundant, and thus would destroy the explanation, would identify the relevant kinds of justifying features.³⁸ But one sort of completion might specify certain relevant features of God, such as that God is immutably omnibenevolent (Adams 1987). If one took such additional information to show that 'x is right because God commands x' was a true partial explanation all along or to turn it into a fuller true explanation, (NEJ) again would help to make sense of such reasoning. For the crucial point would seem to be that commands rooted in God's immutably omnibenevolent nature aren't arbitrary, and so they are eligible to justify the relevant responses to what God commands. This would be to say that the modified divine command explanation meets (NEJ) after all. Moreover, insofar as the claim that the modified divine command explanation isn't false were plausible, (NEJ) would again provide a way to support it, based on the connection it provides between the truth conditions of normative explanations and features that would justify certain relevant responses. This is an advantage: as a general condition, (NEJ) should work under a wide range of assumptions about which particular normative explanations are true and which false.

My argument in this section has also a broader lesson. It illustrates how normative justification is hyperintensional: the truth-value of a sentence concerning what justifies what

³⁸ Following Salmon (1970), it is widely accepted that the inclusion of explanatorily redundant information destroys explanation. To illustrate, consider an explanation of why a piece of sugar dissolved in a sample of liquid which cites the relevant law connecting water and the dissolution of sugar plus the fact that the liquid is holy water (Ruben 2012: 172). The richer information that the liquid is holy water doesn't explain why the sugar dissolved even though the weaker information which it includes does explain why the sugar dissolved.

may not be preserved by the substitution of necessary equivalents. For instance, that God commands x and that x maximizes GP differ with respect to justifying doing x . In cases where necessary equivalents differ in what they justify, (NEJ) implies that the truth-value of a normative explanation is also sensitive to distinctions that are finer than necessarily equivalence: if ‘ x is right because God commands x ’ fails to meet the necessary condition on correct explanation specified by (NEJ), it doesn’t follow that ‘ x is right because x maximizes GP’ is also false. Thus (NEJ) can explain why the substitution of necessary equivalents might change the truth-value also of a normative explanation. But hyperintensionality is widely regarded as a mark of explanation: when ‘ Q because R ’ states an explanation, substituting ‘ R ’ with a necessarily equivalent expression may fail to preserve truth-value.³⁹ So (NEJ) can account for a widely accepted feature of explanation in the normative case.⁴⁰ In that case (NEJ) is capable of securing a significant dimension of continuity between normative explanation and explanations in other domains. There will thus be at least two respects in which scientific, metaphysical, and normative explanation all merit a shared label

³⁹ One example is that (holding the laws of nature fixed) ‘has temperature t ’ is necessarily equivalent with ‘is composed of molecules with mean kinetic energy m ’. Although ‘ b has temperature t because b ’s molecules have mean kinetic energy m ’ looks like a true explanation, ‘ b has temperature t because b has temperature t ’ is no explanation at all. Nothing explains itself. Another example is that ‘This figure is triangular because it is a closed plane figure with three angles’ can be a true explanation whereas ‘This figure is triangular because it is a closed plane figure with three sides’ isn’t, although the predicates flanking the ‘because’ on the right are again necessarily equivalent.

⁴⁰ Explanation may be hyperintensional in two ways. One option is that the truth of an explanation is sensitive to the mode of presentation under which a fact is introduced to the explanation (cf. Ruben 2012: ch. 5). Two necessarily equivalent sentences or propositions may represent the same non-representational state of affairs, thus differing only in how the state is conceptualized. Another option is that at least some non-representational states of affairs which figure in explanations are themselves individuated hyperintensionally, so that pairs of necessarily equivalent propositions may represent distinct states of affairs (Schnieder 2011; Nolan 2014). This interpretation of the hyperintensionality of explanation says that the truth of an explanation is sensitive to which (necessarily equivalent but distinct) facts figure in it. I won’t need to decide between these options here. Even when explanation is sensitive to modes of presentation, it can still be objective in the sense that it relates worldly facts without sensitivity to the background knowledge or interests of an audience or the explanation’s capacity to induce understanding in the audience. Nor does the correctness of such an explanation depend on whether anyone is interested in conceptualizing the world in those ways or even grasps them. It then seems that first-order normative theories won’t need to care too much about fact individuation when constructing their normative explanations.

‘explanation’: each answers why-questions by providing reasons (even as why-questions may request different kinds of reasons in different contexts) and each sort of explanation is hyperintensional.

6. Getting Too Objective?

I’ll now address an objection to (NEJ). The objection says that even if normative explanation is an objective relation in the sense I have specified, normative justification is sensitive also to some epistemic or pragmatic factors, and so isn’t objective in a parallel sense. Then whether a normative explanation identifies features that would go at least some towards justifying those responses wouldn’t—contrary to (NEJ)—be a condition on correct normative explanation. It would at most bear on how good or bad (relative to a context or audience) an explanation is.

It isn’t immediately obvious how to make this objection more precise with respect to the relevant non-objective factors. For instance, the thought cannot be that normative justification is sensitive to epistemic factors because it isn’t appropriate to criticize you for doing something you couldn’t be reasonably expected to have known to be wrong. The action is still wrong. Your ignorance at best excuses your wrongdoing. For all that this says, a complete correct explanation of why the action was wrong must still identify features that would go at least some towards justifying not doing it, resenting those who do it (unless, perhaps, excused), and so on. A better thought might be that features which justify the relevant responses must be ones that it would be appropriate to cite in a normative justification, but that what features these are depends on the conversational context. Even here, however, it is unclear why we should collapse the question whether a normative

justification is good or bad, relative to an audience, to the question whether it is correct. (Read 'correct' as a placeholder for whatever we select as a success term for normative justifications.) Since (NEJ) allows us to distinguish derivative and non-derivative explanations and justifications, it can happily allow that what features it is appropriate to cite as justifications in different conversational contexts may depend on epistemic or pragmatic factors.

I'll leave it for those who want to press this kind of objection to make it more precise, and will instead respond just to the core suspicion that the truth/correctness conditions of normative explanation and justification are differentially sensitive to pragmatic and epistemic factors. We have already seen important respects in which the two are parallel. Normative explanation and normative justification both admit of the distinction between a correct X and a good/bad X (for an audience). In section 2 we saw that the correctness of the kinds of explanations of normative facts which first-order normative theories aim to give isn't relative to anyone's interests or background knowledge or to whether it induces understanding in the audience. It is similarly unclear why the normative justifications generated by such theories should depend for their correctness on such conditions. If act-consequentialism is the true moral theory, the appropriateness of the responses it prescribes to acts that maximize the value of outcomes won't be relative to anyone's interests or background knowledge. A further way to defend (NEJ) against the objection is to draw out some implications of the widely accepted view (discussed above) that explanation (and thus normative explanation) is hyperintensional.

Consider the kind of metaphysical hierarchy we find in the Socratic sort's case, where some facts that might be thought to make her life good (and thereby explain why it is good)

hold in virtue of other facts that might be thought to make her life good. Call such a hierarchy M. Take a special case M' where the facts in M are (type or token) identical. For example, suppose that the fact that an experience E is good is identical with the fact that E is an experience pleasure and this in turn is identical with the fact that the subject of E is in brain state B. Now either the fact that E is good is a reason to promote, commend, or cherish E if and only if the features that make E good are, or not. Some philosophers advocate the latter view: features that make E good can be reasons to promote E (etc.) without its being the case that the fact that E is good is a reason to do so. But suppose we don't. That is, suppose:

- (G) In any metaphysical hierarchy of the form M', the fact that something is good is a reason to promote it (etc.) if and only if the features that make it good are reasons to promote it (etc.).⁴¹

In cases under (G), we have a choice in what to cite as a normative reason for a given response. We might cite the fact that a thing is good, or we might cite a good-making feature. If (G) were true, then plausibly the only choice here would be the pragmatic choice of which fact or facts it would be conversationally appropriate to cite in the context. But then normative explanation and justification would be differentially sensitive to pragmatic factors: the fact that an act is good cannot explain why it is good, but (G) would allow it to go at least some way towards justifying the relevant responses. However, I see no compelling reason to accept (G).

⁴¹ I adapt (G) from a discussion by Zoë Johnson King. Parallels to (G) can be constructed for rightness and other normative properties. If facts or properties are individuated hyperintensionally, hierarchies of the form M' look rare.

One problem for (G) has already come up. In discussing Euthyphro cases, we saw that hyperintensional distinctions seem to be able to make a justificatory difference independently of pragmatic factors. There is no clear reason why this should be confined to Euthyphro style cases. So even if the fact that E is good is identical to the fact that it has a good-making feature F , substituting one for the other in a normative justification may fail to preserve truth-value. But even if we bracket this issue, (G) is unstable.⁴² To say that a feature F is good making is to say that it at least partly explains why some object o is F is good. If the fact that o is good can be identical with the fact that o has a good-making feature F , we must allow that normative explanation is hyperintensional. Otherwise we would fail to secure the result that ‘ x is good because x is F ’ can be a correct normative explanation although ‘ x is good because x is good’ isn’t. But if ‘That x is F (partly) explains why x is good’ is hyperintensional with respect to ‘ F ’, it is hard to see why ‘That x is F is a normative reason for S to ϕ ’ should fail to be hyperintensional with respect to ‘ F ’. For instance, such a failure would make it difficult to hold that if the fact that x is F (non-derivatively) explains why x is good, then the fact that x is F is a (non-derivative) reason to promote, cherish, or commend x . So there is significant pressure on (G) to deny that normative explanation is hyperintensional—a costly move. But if normative explanation and justification aren’t differentially sensitive to epistemic and pragmatic factors in this special case, why they should be so in general? I conclude that we have yet to see compelling reason to think that normative explanation and justification are differentially sensitive to epistemic and pragmatic factors.

⁴² Those who think that the fact that x is good is a derivative normative reason (cf. note 34) should think also that (G) is either false or ambiguous. (Parallel points can again be made about rightness, etc.) (G) will be ambiguous unless it distinguishes derivative and non-derivative reasons. But consistent disambiguation will render (G) false on this view. For it will be false that goodness is a reason if and only if the good-making features are, irrespective of whether the reasons on each side are derivative or non-derivative.

7. Conclusion

I have argued that normative explanation is subject to a kind of justification condition to which explanations of non-normative facts aren't subject, but this doesn't make normative explanation significantly discontinuous with explanations in other domains. The case isn't decisive, based as it is primarily on certain intuitive motivations for (NEJ) and its capacity to make sense of certain aspects of our normative practices. But given the relatively undeveloped state of work on normative explanation, a positive case for (NEJ) should constitute a valuable contribution.

References

- Adams, Robert M. (1987), “A Modified Divine Command Theory of Ethical Wrongness”, in *The Virtue of Faith and Other Essays in Philosophical Theology* (Oxford: Oxford University Press), 97–122.
- Alvarez, Maria (2010), *Kinds of Reasons* (Oxford: Oxford University Press).
- Broome, John (2013), *Rationality through Reasoning* (Oxford: Wiley-Blackwell).
- Bykvist, Krister (2009), “No Good Fit: Why the Fitting Attitude Analysis of Value Fails”, *Mind* 118: 1–30.
- Copp, David (1995), *Morality, Normativity, and Society* (Oxford: Oxford University Press).
- Dancy, Jonathan (2004), *Ethics without Principles* (Oxford: Clarendon Press).
- Darwall, Stephen (1983), *Impartial Reason* (Ithaca, NY: Cornell University Press).
- DePaul, Michael (1987), “Supervenience and Moral Dependence”, *Philosophical Studies* 51: 425–439.
- Elstein, Daniel (ms), “Normative Regress and Normative Relevance”, unpublished.
- Hyman, John (2015), *Action, Knowledge, and Will* (Oxford: Oxford University Press).
- Johnson King, Zoë (2019), “We Can Have Our Buck and Pass It, Too”, *Oxford Studies in Metaethics* 14: 167–188.
- Kearns, Stephen, and Star, Daniel (2009), “Reasons as Evidence”, in *Oxford Studies in Metaethics* 4: 215–242.
- Korsgaard, Christine (1996), *The Sources of Normativity* (Cambridge: Cambridge University Press).
- Nebel, Jacob M. (forthcoming), ‘Normative Reasons as Reasons Why We Ought’, *Mind*.
- Nolan, Daniel (2014), “Hyperintensional Metaphysics”, *Philosophical Studies* 171:149-160.

- Plato, *Euthyphro*. Several editions.
- Raz, Joseph (1999), *Engaging Reason* (Oxford: Oxford University Press).
- Ruben, David-Hillel (2012), *Explaining Explanation*. Updated and Expanded Second Edition (Boulder, CO: Paradigm Publishers).
- Salmon, Wesley C. (1970), “Statistical Explanation”, in *The Nature and Function of Scientific Theories*, ed. R. Colodny (Pittsburgh: University of Pittsburgh Press), 173–231.
- Schnieder, Benjamin (2011), “A Logic for ‘Because’”, *The Review of Symbolic Logic* 4: 445–465.
- Skow, Bradford (2016), *Reasons Why* (Oxford: Oxford University Press).
- Smith, Michael (2000), “Does the Evaluative Supervene on the Natural?”, in *Well-Being and Morality*, ed. R. Crisp & B. Hooker (Oxford: Clarendon Press), 91–114.
- Star, Daniel (2015), *Knowing Better* (Oxford: Oxford University Press).
- Timmons, Mark (2002). *Moral Theory: An Introduction* (Lanham: Rowman and Littlefield).
- van Fraassen, Bas C. (1980), *The Scientific Image* (Oxford: Clarendon Press).
- Väyrynen, Pekka (2013), “Grounding and Normative Explanation”, *Proceedings of the Aristotelian Society Supplementary Volume* 87: 155–178.
- Väyrynen, Pekka (forthcoming), “Reasons Why in Normative Explanation”, *Inquiry*.
- Wedgwood, Ralph (2017), *The Value of Rationality* (Oxford: Oxford University Press).
- Whiting, Daniel (2018), “Right in Some Respects: Reasons as Evidence”, *Philosophical Studies* 175: 2191–2208.