

# Varieties of Normative Explanation

Daniel Fogal, Uppsala University

(based on joint work with Knut Skarsaune, University of Oslo)

## 1 Introduction

My overarching goal in this talk:

- To explore the nature of normative explanation, understood as a kind of metaphysical explanation, where what's being explained is a normative fact.

Some more specific goals:

- To emphasize the importance of distinguishing different types of normative facts (general vs. particular) and properties (particular- vs. property-applying), since different types may be susceptible to different kinds of explanation.
- To consider how non-naturalist normative realists can (and arguably should) explain the supervenience of the normative on the non-normative.
- To consider different ways fundamental normative facts might be 'necessary'.

'Explanation'-talk is pervasive, and tricky—not least because it is highly context-sensitive (cf. Jenkins 2008). Three important distinctions:

1. partial vs. full explanation
  - You: "Why did you order a triple espresso?" Me: "Because I'm tired."
2. pragmatic vs. non-pragmatic explanation
  - An explanation in the pragmatic sense involves "irreducible reference to facts about the interests, beliefs or other features of the psychology of those providing or receiving the explanation and/or (ii) irreducible reference to the 'context' in which the explanation occurs" (Woodward 2014). An explanation in the non-pragmatic sense does not.
3. causal vs. non-causal explanation
  - "Why is the barn red?" Two very different answers: (A) Because I painted it yesterday. (B) Because it is crimson.

"Metaphysical explanation" is a kind of full, non-pragmatic, non-causal explanation.

## 2 Varieties of metaphysical explanation

The recent literature on metaphysical explanation suggests there are at least two:

**"Grounding" explanations:** a fact is explained by saying what it holds in virtue of (its "grounds"), with a general explanatory principle (or "metaphysical law") connecting the grounds with what's grounded. (cf. causal explanation)

- (1) The barn is red. Why? Because it is crimson, and anything that is crimson is thereby red.
- (2) We're attending a conference. Why? Because we're attending a formal meeting of people with the goal of discussing particular topics of shared interest, and any meeting of such a kind is thereby a conference.

The nature of grounding explanations—as well as metaphysical laws—is a matter of controversy; for defense of the "law-based" picture, see Kment 2014, Wilsch 2015, and Glazier forthcoming.

(An internecine debate: are laws *part* of the explanation, or do they merely "underlie" or "back" it? It turns out this *might* matter; I'll return to it in §11.)

**"Essentialist" explanations:** where a fact *F* is explained by saying of some thing (object, property, etc.) *t* that *t* is by its nature (or is essentially) such that *F*.

- (3) Water contains hydrogen. Why? Because water by its very nature contains hydrogen.
- (4) Socrates' singleton set—call it *S*—contains Socrates. Why? Because *S* is essentially such that it contains Socrates.

The above list isn't meant to be conclusive, nor exhaustive. There may also be "necessitarian" explanations, for example, where a fact is explained by saying it is necessary. And it's neutral regarding so-called "reductive" explanations (e.g. facts about water "reduce" to facts about H<sub>2</sub>O—cf. Bader (ms) on "vertical" vs. "horizontal" explanations) as well as "identificational" claims such as 'to be an *F* just is to be a *G*' (e.g. to be just is to have a well-ordered soul).

- So: there are a lot of open questions. Metaphysics is messy!

## 3 The structure of grounding explanations

Suppose Billy sees a cat and kicks it for fun. What Billy did was wrong. Why? A natural answer: because he kicked a cat, and kicking cats for fun is wrong.

This is an example of a *particular normative explanation*—i.e. an explanation of a particular normative fact. In such a case, what's being explained is a particular normative fact and what's doing the explaining are (one or more) particular non-normative facts together with a (more) general normative principle.

**Explanans:** particular (non-)normative fact(s)

**Connection:** general normative principle

**Explanandum:** particular normative fact

The explanans together with the explanatory connection constitute the "full" explanation of the explanandum.

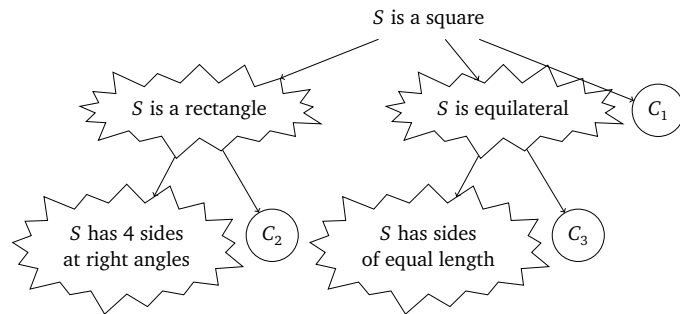
- Note: both parts of the full explanation needn't be, and usually aren't, actually cited in providing an explanation. Normally the relevant normative principle is simply taken for granted, and hence treated as part of the conversational background. This is what we'd expect given our explanatory practice.

Also: particularity and generality are both a matter of degree, and both normative principles and non-normative facts can vary in terms of their (relative) fundamentality, and so will oftentimes admit of further explanation.

- For example: kicking cats for fun is wrong because it involves harming sentient creatures and harming sentient creatures is wrong.

But both particular non-normative facts and general normative principles plausibly “bottom out” somewhere, with the fundamental normative principle(s) being maximally general.

This fits the structure of metaphysical “grounding” explanations in general, with general normative principles serving as the “explanatory connection” ( $C_1, C_2, \dots$  below) between non-normative facts and the particular normative facts... (from Glazier):



- Terminological aside: ‘ground(s)’ is used in different ways. For some, to say that ‘A grounds B’ is to just say that A metaphysically explains B (e.g. Litland 2013, Dasgupta 2014, Wilsch 2015), while for others ‘A grounds B’ states an explanatory connection to which this metaphysical explanation will appeal (e.g. Audi 2012). And still others slide back and forth unwittingly. Hence it’s best to avoid it.
- Also: the connection in a given explanation should be distinguished from the fact that the explanans explains the explanandum. A natural—but to my mind problematic—suggestion is that the former will explain the latter, at least in part (cf. Bennett 2011, deRosset 2013). A better suggestion is that the explanans explains the explanandum only in the sense that it is *part* of the full explanation, per above. (This helps avoid explanatory regress worries...)

## 4 Metaphysical and normative laws

Following Glazier (forthcoming), let *Crimson* be the general connection between an arbitrary thing’s being crimson and its being red. An “atomic” explanation of the fact that the barn is red:

Explanans: The barn is crimson.

Connection: *Crimson*.

Explanandum: The barn is red.

An important question: what is the logical form of *Crimson*? A natural thought: *Crimson* is just a universal generalization, something like

For all  $x$ , if  $x$  is crimson, then that  $x$  is crimson makes it the case that  $x$  is red.

But this is a mistake, since universal generalizations—including “non-accidental” ones like that above—are plausibly explained, at least in part, by their instances, whereas with *Crimson* and its instances the explanatory order is reversed.

If *Crimson* is not a universal generalization, what is it? Here’s Glazier:

“It clearly has a sort of generality, but it is a general fact that is not explained by its instances. Since this sort of generality is not achieved through quantification, it must instead be achieved through another variable-binding operator. I therefore propose that we recognize a new operator ‘ $\gg$ ’ [that] bind[s] any number of variables, since our intuitive understanding of a general metaphysical-explanatory connection does not support any relevant limit. And because a fact may be metaphysically explained by any number of other facts, the operator should also be variably polyadic ‘on the left.’”

A statement of a general explanatory connection (or “metaphysical law”) will thus be of the form

$$\varphi_1, \dots, \varphi_n \gg_{\alpha_1 \dots \alpha_m} \psi$$

where  $\varphi_1, \dots, \varphi_n, \psi$  are sentences and  $\alpha_1, \dots, \alpha_m$  are variables.

We can thus state the fact *Crimson* as

$$\text{Crimson} \quad x \text{ is crimson } \gg_x x \text{ is red.}$$

In terms of facts: *Crimson* states the general explanatory connection that holds between facts of the form ‘ $x$  is crimson’ and facts of the form ‘ $x$  is red’.

A plausible proposal: the relevant “normative laws” that appear in grounding explanations of normative facts should be understood in the same way as those appearing in grounding explanations more generally (cf. Rosen 2010 on “Moorean connections”). The difference is simply that normative laws involve normative properties—a difference in content, not form.

Of course, this hardly settles any of the hard metanormative questions concerning the status, content, number, and explanation (if any) of the relevant normative laws. But that's a feature of this proposal, not a bug: it focuses the metaphysical side of the metanormative debate precisely where it should be (and historically has been)—namely, on the fundamental normative laws (if such there be).

## 5 Supervenience: the “Divide and Conquer” strategy

Prominent objection against non-naturalist normative realism: unlike their naturalist (or error-theoretic) rivals, non-naturalists cannot provide a metaphysical explanation for why normative properties supervene on natural properties.

The relevant supervenience relation that is standardly (though somewhat controversially) thought to hold between the normative and the natural is as follows, where A is the family of normative properties, B is the family of “natural” (or “descriptive” or “not-normativity-involving” or “base”) properties, and  $\Box$  is metaphysical necessity (cf. McPherson 2015, Dreier MS):

**Strong Supervenience**  $\Box(\forall F \text{ in } A)(\forall x)[Fx \rightarrow (\exists G \text{ in } B)(Gx \ \& \ \Box(\forall y)(Gy \rightarrow Fy))]$

In English: necessarily, for any normative property *F*, if something is *F*, then that thing has some natural property *G*, and necessarily everything that has *G* has *F*.

Assuming Strong Supervenience, the non-naturalist (and everyone else who isn't afraid of metaphysics) faces the question:

(Q1) What, if anything, explains why Strong Supervenience holds?

A promising strategy that emerges from the above...

### The “Divide and Conquer” Strategy

First step: distinguish particular normative facts from general normative principles or “laws”.

Second step: insist that particular normative facts have grounding explanations—they are explained by non-normative facts + general normative principles.

Third step: insist that the (explanatorily fundamental) general normative principles are not contingent.

The explanation of why normative facts strongly supervene on non-normative or “natural” facts is thus straightforward (cf. Enoch 2011, Scanlon 2014, Skarsaune 2015). For whether or not a particular thing has a normative property depends on two things: (i) which non-normative properties it has and (ii) which fundamental normative principles are true. Regarding (i), it's obvious that there can't be a difference in a thing's non-normative properties without there being a difference in its non-normative properties. And regarding (ii), it's equally obvious that there can't

be a difference in which fundamental normative principles are true without there being a difference in some non-normative facts, since the relevant principles are not contingent and hence trivially supervene on everything (there can't be a difference in them, period). There will thus never be a difference in the normative facts (general or particular) without there being a difference in the non-normative facts.

- Note: even though particular normative *facts*, which involve the *instantiation* of normative properties, will always have grounding explanations, it doesn't follow that *the normative properties themselves* can be explained in any way.

## 6 Two (and a half) objections: Leary and Wilson

Leary (forthcoming) considers a similar line of response, but worries that if non-naturalists claim that the fundamental normative principles are metaphysical laws...

“...it's not clear how they can maintain their non-naturalist commitments, since the metaphysical structure of the normative facts is exactly the same as that of certain derivative natural facts. For example, consider facts about tables. If metaphysical laws are fundamental and [explain] the contingent derivative facts together with the contingent fundamental facts, then table-facts are [explained] in the same way as the normative facts: e.g. the fact that there is a table is [explained] in the fact that there are particles arranged table-wise and the fundamental metaphysical law that if there are particles arranged table-wise, then there is a table. So, it's not clear on this view why tables are the same kind of stuff as paradigmatic scientific stuff and are compatible with a scientific worldview, while normative properties are not.”

The obvious response: the difference derives from the *content* of the laws—normative laws involve normative properties, while non-normative laws do not.

- Compare a mind-body dualist who proposes the following metaphysical law: brain *b* is in state *B* >> cartesian soul *S* (i.e. the one associated with *b*) experiences pain. Leary could object: “It's not clear how they can maintain their dualist commitments, since the metaphysical structure of the mental facts is exactly the same as that of certain derivative natural facts”. But that's not very convincing...)
- Leary may reply: OK, but what makes normative properties *different* from natural ones (and hence “non-natural”)? There are at least two options: (a) primitivist response and (b) Leary's own essentialist response (following Rosen)—i.e. “the essences of some normative properties involve something irreducibly normative, which cannot be specified in non-normative terms, and...do not specify naturalistic or any non-normative sufficient conditions for their instantiation”.
- Note: the law-based theorist can accept Leary's essential explanation of non-naturalism while *rejecting* the existence of (what she calls) “hybrid properties”.

However, there's another worry in the vicinity (inspired by Wilson 2015):

“Law-based grounding explanations just say something (together with a law) explains something else without saying anything more specific about *how* the former explains the latter. If a law of the form ‘x is in brain state B  $\gg$  x is conscious’ can be agreed upon by somebody who is a Cartesian dualist about consciousness, by somebody who thinks consciousness *emerges* from B, by somebody who thinks consciousness is *constituted* by B, by somebody who thinks consciousness is *functionally realized* by B, etc., then a law-based explanation is silent on many important metaphysical issues.”

Perhaps so, but that's not an objection, for at least two reasons:

- First, there will be metaphysical laws associated with each kind of metaphysical “dependence”/“building”/“construction” relation that we admit (composition, emergence, constitution, set-membership, realization, etc.), since each will be a particular way in which some things explain other things, and the full statement of such “grounding” explanations will plausibly involve metaphysical laws per above. The difference is just that in such cases the laws will hold (partly) in virtue of facts about the relevant relation. But just because a metaphysical law is susceptible to explanation doesn't mean it isn't a law, or that it doesn't do any explanatory work. (cf. Schaffer forthcoming, Berker MS)
- Secondly, when we turn from (e.g.) the philosophy of mind to the normative realm the range of (remotely plausible) candidate “construction” relations shrinks dramatically, so the worry isn't terribly worrisome.

## 7 The dependence intuition

Clearly distinguishing kinds of normative facts and kinds of explanation also clarifies the sense in which the normative “depends” on the non-normative.

“It is very common to think that actions and other things have their normative and evaluative properties *in virtue of* their non-normative, non-evaluative properties. It is similarly very common for those who are allergic to talk of normative properties nonetheless to agree that things are good or bad, or right or wrong, *because* of some non-normative properties. There is, in other words, a strong intuition that normative facts are dependent on and explained by other facts. Call this ‘the dependence intuition’.” (Värynen 2013)

The dependence intuition derives most (and maybe all) of its plausibility from considering *particular* normative facts, not general ones.

- For example, the wrongness of Billy's action depends on the non-normative facts, including the fact that he kicked a cat. But it doesn't *fully* depend on them. It also depends on the normative fact that kicking cats for fun is wrong.

When it comes to the fundamental general normative principles, however, the dependence intuition is far less compelling. Why is suffering bad, or intentional deception wrong-making? A natural answer: *it just is*.

Moral: here as elsewhere we should be wary of the gap between true generic claims (“normative facts depend on non-normative facts”) and their corresponding universal generalizations (“all normative facts depend on non-normative facts”).

## 8 The “grounding argument” against non-naturalism

Bader's (forthcoming) “grounding” argument against non-naturalism:

- 1'. normative properties are (fully) grounded in non-normative properties.
- 2'. the set of non-normative properties is closed under infinitary disjunction and conjunction.
- 3'. disjunctive properties are (individually) grounded in their disjuncts; conjunctive properties are (collectively) grounded in their conjuncts  
∴ for every normative property there is a non-normative property that is hyperintensionally equivalent.
- 4'. hyperintensionally equivalent properties are identical  
∴ every normative property is identical to a non-normative property .

The non-naturalist should deny premise 1'. (And perhaps 4' too.) Normative properties are not fully grounded in non-normative properties. Of course, particular normative *facts* (i.e. ones involving the instantiation of particular-applying normative properties) are “grounded” in particular non-normative facts, but even so they aren't *fully explained* by such—general normative principles (or “laws”) are also involved, which aren't explained by anything.

**Upshot**: there's no need to admit Bader's distinction between “normative grounding” relations and “metaphysical grounding” relations, with the former being “governed” (but not grounded) by normative laws and the latter being “governed” by metaphysical laws. For according to Bader:

“...normative laws do not ground normative properties but govern the grounding relations connecting non-normative grounds to normative properties...Rather than normative properties being metaphysically grounded in non-normative properties together with the normative laws, they are normatively grounded in the non-normative properties via these laws...they modify the grounding relation rather than featuring as relata.”

The difference between normative and metaphysical grounding explanations *does* boil down to a difference in the laws involved in each, but it's merely a difference in content. No need to multiply “grounding” relations or talk of laws “governing” such.

## 9 Brute connections?

Suppose the Divide and Conquer strategy is successful. Two obvious questions:

- (Q2) Isn't the non-naturalist still saddled with "brute connections" between quite different properties?
- (Q3) What, if anything, explains the *non-contingency* of the brute "connections" between normative and non-normative properties?

Short answer to (Q2): yes—this is simply the bullet non-naturalists bite.

Longer answer: it depends on what is meant by 'connections'...

- **Distinguish:** co-occurrence (iff) vs. general-'making' ( $\gg_x$ ) vs. instantiation (is)

Co-occurrence is not brute. Are the normative laws/explanatory connections brute? It depends on the non-naturalist. It's worth distinguishing between:

- (A) **Nomological Fundamentalists:** yes, the normative laws/explanatory connections are brute. (cf. Enoch 2011, Scanlon 2014)
- (B) **Normative Platonists:** no, they're not, since "the basic normative properties are *second-order*; they take *universals* [i.e. *properties*], rather than particulars, as their bearers" (Skarsaune 2015).
  - You: list everything that is bad. Me: OK...pain, suffering, the war in Syria, unrequited love, genocide, the fact that Bill cheated on Hilary, etc.
  - Thus: "the property *causes needless pain* has the property *wrongmaking* [and] the property *pleasant* [instantiates] the property *goodmaking*...The necessary connection between them is not *co-occurrence*, but *instantiation*...a first-order [property] has a second-order [property]."

Possibly interesting aside: What should Normative Platonists say about first-order, particular-applying normative properties? There are at least two options here, corresponding to two ways of interpreting claims of the form 'To be *F* is to be *G*':

Option 1 (Identity): for any particular-applying normative property *F*, *F* = the property *having an f-making property* (where '*f*' is an adjective expressing *F*).

- On this view, particular-applying *goodness* is (i.e. =) the property *having a good-making property*, particular-applying *wrongness* is the property *having a wrong-making property*, and so on. (cf. Skarsaune 2015)
- Property-applying properties (e.g. *goodmaking*) are basic.

Option 2 (Instantiation): *F* is not to be *identified* with the property *having an f-making property*—instead, the latter simply provides the *instantiation conditions* of the former (i.e. what it takes for something to have *F*).

- Particular-applying *goodness* might thus be taken as basic, even though

particular facts involving *goodness* (i.e. particular instantiations of *goodness*) won't be—they'll be explained in way indicated above.

- What about property-applying normative properties (e.g. *good-making*)? Two more (sub-)options (these aren't exhaustive):
  - Option 2(a): They are defined in terms of the relevant normative laws (e.g. for *P* to have the property *good-making* is just for *P* to make things that have *P* good—i.e. for *P* to be such that  $x$  is  $P \gg_x$   $x$  is good).
  - Option 2(b): They are taken as basic, and as that which explains the relevant normative laws (e.g. the property *good-making* is essentially such that  $x$  is  $P \gg_x, P: P$  is *good-making*  $X$  is good).

What's common to all these options is the need to distinguish particular-applying normative properties (e.g. *goodness*) and property-applying ones (e.g. *good-making*), though only Option 1 and Option 2(b) are versions of Normative Platonism. End of possibly interesting aside.

- (C) **Essentialists:** no, because the normative laws/explanatory connections can be given essentialist explanations. (cf. Leary forthcoming)

An Essentialist might claim, for example, that the property *painfulness* is essentially such that  $x$  is painful  $\gg_x$   $x$  is bad. (This is compatible with Option 2(a) above. Option 2(b)—assuming it's coherent—is another possible story. One can therefore be both a Normative Platonist and an Essentialist.)

## 10 Fundamentality and autonomy

No matter which route she takes, the non-naturalist realist will insist that at least some normative facts and/or properties are *metaphysically fundamental* insofar as they lack any metaphysical explanation (and play a role in explaining all the rest).

In addition to being fundamental, it's a further question whether they are metaphysically *autonomous*. To be autonomous is to not be the sort of thing that can, in principle, have a metaphysical explanation—it's to be such that the question of what explains it "does not legitimately arise" (Dasgupta 2014; cf. Schaffer 2009).

- Analogies: not everything is apt for causal explanation (e.g. mathematical facts), and not everything in a formal system is apt for being proved (e.g. definitions).

Not all metaphysically fundamental facts or properties are autonomous, though some arguably are. For example: essentialist truths concerning properties.

- E.g. Knowledge is essentially factive, and crimson is essentially red. Why? Natural response: that's just (part of) *what knowledge and crimson are*.

Also plausibly autonomous: facts about what (fully) metaphysically explains what.

(Though perhaps an essentialist explanation is possible...)

It's less clear whether, say, facts about which properties are good-making are autonomous (even if they're fundamental). It depends on whether, for example, they're susceptible to essential explanation.

## 11 Transcendence

Suppose the Divide and Conquer strategy is successful, *and* we're willing to bite the non-naturalist bullet. Next question:

(Q3) What, if anything, explains the *non-contingency* of the "connections" between normative and non-normative properties?

(To repeat: this is *not* a question about supervenience.)

A promising response for the non-naturalist: whether or not the fundamental normative laws (and/or facts involving property-applying normative properties) are autonomous, they are nonetheless *transcendental*.

Huh? A quick detour through language (following Skarsaune 2015, inspired by Fine 2005, Ch. 10): the basic normative statements (e.g. 'suffering is bad', 'pleasure is good') are "timeless" and "unworldly": they're not to be evaluated with respect to any particular time or world. (Compare: '7 is prime'...)

The metaphysical analogue of this is not necessity, but transcendence. A "transcendental" fact is not one that holds *in* all possible worlds, but rather one that holds *independently* of the worlds. As Skarsaune notes, Fine thus makes a distinction within the class of what is usually thought of as the necessary truths:

"The necessary truths proper are those that engage with each world, but... come out true every time. 'P or not-P' is one example; in some worlds it holds because 'P' is true, in other worlds because 'not-P' is. Another example is 'Nothing is both red and green all over'. Whichever world it is evaluated at, it looks at every object in that world, but never finds a counter-instance.

The transcendental truths, on the other hand, do not even engage the different worlds; there is nothing *in* the worlds that they answer to for their truth. Mathematical claims are good examples: 'Seven is a prime number' can only in a degenerate sense be said to be true 'in' a given world, for there is nothing about any of the worlds that makes it true."

We thus need to distinguish between two ways of thinking about possible worlds:

"On what we can call the *tractarian* conception, a world is a totality of facts. The actual world is everything that is the case; a merely possible world is a way everything could have been."

"On the second conception, a world is a totality of concrete substances and their

attributes. If you specify all the concrete substances there ever are, and all the properties and relations they have throughout their careers, then you have specified a world. Let us call this the *substantive* conception of a world."

The modal status of many theses of philosophical interest are best understood in terms of the latter conception:

"...we can imagine God pondering which world to create. His alternatives are the substantive possible worlds, not the tractarian ones. Seven being a prime number is not somehow a feature of every alternative, it is not a feature of any of them. The mathematical facts are already there, before he creates anything.

Likewise, when non-natural realists say the basic normative facts are necessary... they are not well understood as saying that these facts hold, again and again, in every world. They are part of the invariable framework. God may decide which world to create, but he does not get to say how good it will be if created. The basic normative facts are already there."

## 12 Transcendental, Fundamentalist Non-Naturalism

I've lazily borrowed the following diagram from Leary (forthcoming). There are some **important alterations** that need to be made, though. In particular:

- change 'fundamental' to 'fundamental *and* transcendental'
- change 'partially grounds' to 'partially *explains*'
- note: the explanans + laws *fully explain* the explanandum, with *that* fact being autonomous (like all facts about what metaphysically explains what, assuming they can't be given an essential explanation).

Diagram 1: Fundamentalist Non-naturalism

