6 Self-Explanation

The purpose of this chapter is to explore the idea of a self-explanatory proposition and to develop a novel notion of self-explanation. The idea of self-explanation is as controversial as it is philosophically interesting: On the one hand, certain alleged fundamental facts or first principles, e.g. God's existence, have sometimes been taken to be self-explanatory.[202] As already mentioned in the introduction, the idea of a self-explanation is one way of spelling out the idea of an ultimate explanation, i.e. an explanation whose explanans does not give rise to further why-questions. On the other hand, self-explanation is frequently considered to be incoherent and unintelligible.[203]

As we will see, given the inclusive sense of 'explains', which was introduced in chapter 1 and in which both the sources and the link of an explanation can be said to explain its explanandum, two importantly different senses in which a proposition can be (at least partially) self-explanatory can be distinguished. In the following I want to focus on one of these notions, which is less often (if ever) recognized, even though it is more viable than the other. As it turns out, to define a corresponding notion of a *fully* self-explanatory proposition, the notion of an empty-base explanation is required too. This chapter argues that the resulting kind of self-explanation is possible (or at least compatible with the nature of explanation) and identifies some in principle candidates for such propositions.

This is the plan for the chapter: Section 6.1 approaches the notion of self-explanation and presents a family of arguments against its possibility. After having recapped some general assumptions about explanation from chapter 1, section 6.2 disambiguates two notions of (self-)explanation – the *restrictive* and the *inclusive* sense – the latter of which is then argued to be able to avoid the arguments from the previous section. Section 6.3 uses these findings to offer a solution to a circularity problem for Humeanism about laws of nature.

Section 6.4 then combines these previous results with the notion of an empty-base explanation to introduce the notion of an *empty-base link-self-explanation* and defend it against two further arguments against self-explanation due to Kovacs (2018). Section 6.5 develops the notion further and investigates its application to the

202 Proponents of the principle of sufficient reason (PSR) are sometimes drawn to ideas like this (cf. Guigon 2015). Spinoza for example considers God to be a *causa sui* (cf. Lærke 2011). The idea can also be found in the literature on the question why there is anything at all, e.g. Nozick (1981, 115ff.).

203 E.g. Oppy (2006, 277ff.), Kovacs (2018), and relatedly Schnieder (2015) on the asymmetry of 'because'.

https://doi.org/10.1515/9783111069500-007

ideas that first principles, God's existence, or certain grounding propositions themselves are self-explanatory. As it turns out, the notion can help make sense of Nozick's (1981, 119ff.) idea of "explanatory self-subsumption" and capture some strongly rationalist theses related to necessitarianism and the PSR. Section 6.6 concludes by showing that certain historical ideas about the explanation of God's existence give rise to a proposal for a self-explanation in the developed sense.

6.1 Approaching self-explanation

Let us approach the topic of self-explanation by observing what I take to be a conceptual platitude: For a proposition to be self-explanatory is for it to explain itself. Or, schematically:

> For a proposition *x* to be self-explanatory is for *x* to explain *x*.

Here, of course, 'explains' has to be used in the relational sense in which it expresses a relation that relates propositions or facts, viz. the kind of entities that constitute explananda and explanantia.

Before we disambiguate 'explains' further, note that the platitude already helps to distinguish self-explanation from related notions like our own empty-base explanation and Dasgupta's (2014b, 2016) explanatory autonomy, which might play a similar theoretical role or provide similar explanatory benefits as self-explanation proper. For example, one purpose of all three notions is to help satisfactorily end explanatory inquiry or avoid it all together.

Nevertheless, neither the notion of explanatory autonomy nor the notion of an empty-base explanation capture the idea expressed by the above platitude, namely that of a proposition *explaining itself*: Firstly, an explanatorily autonomous proposition is not explained, rather it is such that *qua* being autonomous it does not require an explanation.[204] Therefore, it is not self-explanatory in the proper sense. Secondly, empty-base explanations are (to foreshadow a little: at least in general) not instances of a proposition explaining itself, as is for example witnessed by the zero-grounding explanation of non-factive grounding claims (à la Litland 2018): Non-factive grounding claims are (empty-base) explained, but they do not explain themselves.[205]

---

204 Perhaps it is possible that a proposition does not *require* an explanation and nevertheless *has* an explanation, but even this case does not capture the idea of a proposition *explaining itself*.

205 Granted, a more relaxed sense of a self-explanatory proposition might exist in which for example merely empty-base explainable propositions count as self-explanatory. Perhaps such a sense functions similarly to 'self-evident': A self-evident proposition arguably need not be evidence for

Eventually, a comparison of the advantages and disadvantages of the three related notions will be desirable, but here my primary task is to investigate self-explanation proper, as captured by the platitude (although see section 6.4 for some comparison of empty-base explanation and self-explanation with respect to the idea of ultimate explanation). Indeed, the possibility of self-explanation in the platitudinous sense is heavily contested. While this is often based on raw intuition, here I focus on the following three arguments from the asymmetry of related notions:

'From "because"':

(P1) For any *P*, *Q*: If the proposition that *P* explains the proposition that *Q*, then *Q* because *P*.

(P2) For no *P*: *P* because *P*.

(P3) For any *x*: If *x* explains *x*, then there is a proposition that *P* such that the proposition that *P* explains the proposition that *P*.

(C1) For no *x*: *x* explains *x*.[206]

'From explanatory dependence':

(P4) For any *x*, *y*: If *x* explains *y*, then *y* stands in an explanatory dependence relation to *x*.

(P5) For no *x*: *x* stands in an explanatory dependence relation to *x*.

(C1) For no *x*: *x* explains *x*.[207]

'From reasonhood':

(P6) For any *x*, *y*: If *x* explains *y*, then *x* is a reason for *y*.

(P7) For no *x*: *x* is a reason for *x*.

(C1) For no *x*: *x* explains *x*.

These arguments are similar in form: The first premise establishes a link between explanation and a further notion, the second premise establishes the asymmetry of that notion, and from this the asymmetry of explanation follows. The arguments are valid, so the proponent of self-explanation has to address the premises.

Admittedly, the arguments may perhaps be of somewhat limited dialectical value: A staunch defender of self-explanation might rather take them as counting against one of their premises than be convinced by them. In particular, the

itself. Rather, no further proposition is required for it to be evident. Here, though, I want to focus on the idea captured by the platitude above.

206 For an argument like this see Oppy (2006, 277f.). Let us ignore complications that might arise from quantifying into the contexts of 'explains' and 'because': My purpose here is to present a notion of self-explanation that can avoid these arguments independently of such concerns.

207 An argument like this is suggested in Schnieder (2015).

premises (P2), (P5) and (P7) that establish the asymmetry of the respective notion related to explanation seem to come quite close to the conclusion that nothing explains itself. Nevertheless, these premises enjoy considerable intuitive appeal and are widely endorsed.[208]

Therefore, I consider denial of either (P2), (P5) or (P7) to be a significant cost that would require serious argument.[209] So instead of taking this route in defense of the possibility self-explanation, I will now employ the distinction between the restrictive and the inclusive sense of 'explains' introduced in chapter 1: While we can maintain premises (P1), (P4) and (P6) given the restrictive sense, these premises are doubtful given the inclusive sense (of course, we are still free to endorse the three simple arguments if we choose to understand 'explains' in the restrictive sense throughout).

6.2 Two notions of (self-)explanation

Recall that in chapter 1 I argued that there is an *inclusive sense* of 'explains' in which not only the reasons (i.e. elements of the base) involved in an explanation (at least partially) $\text{explain}_{\text{inclusive}}$ the explanandum, but also the link of an explanation (partially) $\text{explains}_{\text{inclusive}}$ its explanandum. As explained there, this sense of 'explains' stands in contrast to a more *restrictive sense* which corresponds more closely to because-statements and in which only the elements of the explanatory base (i.e. the reasons why), but not the link of an explanation (partially) $\text{explain}_{\text{restrictive}}$ its result.

To make this explicit, we can define the two senses as follows:

- For all x, y: x (at least partially) $\text{explains}_{\text{restrictive}}$ y $\text{iff}_{\text{def.}}$ x is in the base of an explanation whose result is y.
- For all x, y: x (at least partially) $\text{explains}_{\text{inclusive}}$ y $\text{iff}_{\text{def.}}$ x is in the base of an explanation whose result is y, or x is the link of an explanation whose result is y.[210]

208 But, of course, not universally: For example, one reason to deny causal irreflexivity may stem from the possibility of time travel and corresponding causal loops, cf. Smith (2019). For a critical discussion of the irreflexivity of metaphysical dependence see Jenkins (2011), and for the irreflexivity of grounding see Kovacs (2018) and the references therein.

209 For the case of grounding explanations, the start of such an argument might be provided by the puzzles of ground given in Fine (2010) and Krämer (2013). For some further discussion concerning the irreflexivity of grounding explanation see Bliss and Trogdon (2016, sec. 6.2).

210 This should be understood as a definition of *immediate* explanation to avoid ruling out here that x may in an inclusive sense explain y by in the restrictive sense explaining a link z of an explanation of y (and assuming a principle of transitivity, cf. the next section).

Correspondingly, we can distinguish self-explanation in the inclusive sense from self-explanation in the restrictive sense: Proposals for self-causing or self-grounding facts concern self-explanation in the restrictive sense, while we will encounter candidates for self-explanations in the inclusive (but not restrictive) sense below.

Given this distinction, we can try to defend one type of self-explanation by arguing that the arguments against self-explanation only apply to the other type of self-explanation. Indeed, it can be argued that the first premise of each of the three arguments above is false given the *inclusive* sense of 'explains': For example, so understood, (P1) is false because if a proposition that *P* explains$_{\text{inclusive}}$ a proposition that *Q*, then it is not in general the case that *Q* because *P*. The two sentential arguments of a 'because'-statement correspond to the base and result of an explanation and it is normally not the case that the explanatory link of an explanation is also in the base of the relevant explanation and thereby occurs in the corresponding 'because'-statement in this capacity. Rather, explanatory links correspond in a different way to 'because'-statements, for example by being tracked by the latter.[211]

Analogous considerations arise for (P4) and (P6) of the other arguments: If *x* explains$_{\text{inclusive}}$ *y*, then it is not in general the case that *y* suitably depends on *x*: For example, the explanandum of a causal explanation does not causally depend on the causal connection or law of nature connecting it and its cause. Likewise, the grounding connection between a ground and a groundee does not normally also ground the groundee.[212] Explanatory links involve the explanatory priority relation between an explanation's sources and its result, but in general do not themselves stand in such a relation to the result. Similarly, (P6) is false because if *x* explains$_{\text{inclusive}}$ *y* (viz. by being the link of an explanation of *y*), then it is not in general the case that *x* is a reason for *y*. The base of an explanation consists of reasons for the explanation's result, but links normally do not play this role; instead links connect the reasons that constitute the explanation's base with its result.[213]

There is a more general lesson here: 'explains$_{\text{inclusive}}$' does not necessarily share the structural features of 'explains$_{\text{restrictive}}$'. On the tripartite view of explanation and 'because' introduced in chapter 1, structural features often ascribed to explanation (e.g. asymmetry and transitivity) are captured by ascribing corresponding structural features to the link component. Additional analogous constraints on, e.g., the relation between explanatory links and results are unmotivated on this view: According to it,

211 Cf. Schnieder (2010).

212 Cf. Bolzano (1837, secs. 199, 344f.) and Litland (2018).

213 See chapter 1 and the discussion in Skow (2016).

what the relevant structural features of explanation come down to are the structural features of explanatory links. But normally, no additional explanatory links hold between the link and result of an explanation, so there appears to be no reason to assume corresponding structural features to govern the relation between link and result. In fact, stipulating corresponding constraints in addition to the structural features of the links would result in a disjoint account.

These considerations allow (but do not compel) us to maintain that self-explanation$_{\text{restrictive}}$ falls prey to versions of the three arguments in which each occurrence of 'explains' is understood in the restrictive sense while maintaining the intelligibility of self-explanation$_{\text{inclusive}}$:

'From "because" – revised':

(P1*) For any P, Q: If the proposition that P explains$_{\text{restrictive}}$ the proposition that Q, then Q because P.

(P2) For no P: P because P.

(P3*) For any x: If x explains$_{\text{restrictive}}$ x, then there is a proposition that P such that the proposition that P explains$_{\text{restrictive}}$ the proposition that P.

(C1*) For no x: x explains$_{\text{restrictive}}$ x.

'From explanatory dependence – revised':

(P4*) For any x, y: If x explains$_{\text{restrictive}}$ y, then y stands in an explanatory dependence relation to x.

(P5) For no x: x stands in an explanatory dependence relation to x.

(C1*) For no x: x explains$_{\text{restrictive}}$ x.

'From reasonhood – revised':

(P6*) For any x, y: If x explains$_{\text{restrictive}}$ y, then x is a reason for y.

(P7) For no x: x is a reason for x.

(C1*) For no x: x explains$_{\text{restrictive}}$ x.

Thus, we are free to deny the intelligibility of self-explanation$_{\text{restrictive}}$ while maintaining the intelligibility of self-explanation$_{\text{inclusive}}$, candidates for which we will look at in what follows.

6.3 On a circularity problem for Humeanism about laws of nature

Before we combine the notions of self-explanation$_{\text{inclusive}}$ and empty-base explanation to investigate the possibility of *fully* self-explanatory$_{\text{inclusive}}$ propositions, I want to show how the previous result applies to matters that the reader might

consider a bit more grounded. According to Humeanism about laws of nature (as I will understand them here), laws of nature are universal generalizations (or at least partially grounded in such). This idea is confronted with the following circularity problem that the distinction from the previous section can help solve:

Consider an explanation of [Ga] whose explanatory link is identical to or grounded in the universal generalization [$\forall x(Fx \rightarrow Gx)$], and whose explanatory base contains [Fa].[214] Together, the link and the base explain the result, so in particular:

(1) [$\forall x(Fx \rightarrow Gx)$] partially explains [$Ga$].

But it is a widely accepted grounding principle about (true) universal generalizations that they are (partially) grounded in their instances, so [$Fa \rightarrow Ga$] partially explains [$\forall x(Fx \rightarrow Gx)$]. Equally, it is widely accepted that if a material conditional has a true consequent, the former is grounded in the latter. So [Ga] explains [$Fa \rightarrow Ga$], and an application of transitivity for grounding yields:

(2) [Ga] partially explains [$\forall x(Fx \rightarrow Gx)$].[215]

But (1) and (2) constitute an instance of symmetric (partial) explanation and an application of transitivity would even yield an instance of (partial) self-explanation.[216]

While several solutions to this problem have been discussed in the literature, the observations from the previous section afford a particularly straightforward solution: The derivation of a symmetric instance of 'explains' can only succeed given the inclusive sense of 'explains': (1) is true only in this sense. But as we have seen, there is reason to believe that structural features of explanation such as asymmetry only apply to the restrictive ('because'-corresponding) sense of 'explains', so the problem is avoided.[217]

---

214 As always, I use '[. . .]' to refer to the proposition or fact expressed by the sentence within the brackets.

215 For proponents of the relevant grounding principles see for example Fine (2012, 59ff.), Schnieder (2011, 406f.), Correia (2013a, 44f.), and for discussion in the present context Roski (2018). Note that for the problem to arise, all the Humean has to postulate is that laws are sometimes partially grounded in what they explain. This arguably already follows from the idea of *Humean supervenience*, championed by David Lewis, according to which nomic facts arise from a 'mosaic' of particular, non-nomic facts (cf. Weatherson 2016, sec. 5).

216 For discussion of this problem see, e.g., Loewer (2012), Lange (2013b), and Roski (2018), as well as the latter's bibliography.

217 Note that, alternatively, the application of transitivity in deriving a (partial) self-explanation from (1) and (2) could also be blocked like this.

6.4 Empty-base *self*-explanation

Self-explanations promise to be *ultimate* explanations, i.e. explanations that end explanatory regresses and do not give rise to further why-questions. Explanations by status (and thus empty-base explanations) may play a similar role: They explain without involving reasons why that could give rise to further why-questions. Nevertheless, empty-base explanations are (generally) not self-explanations in the platitudinous sense. Still, the notion of an empty-base explanation can be used to characterize a particular kind of full self-explanation$_{\text{inclusive}}$ that is not a self-explanation$_{\text{restrictive}}$, namely that of an empty-base explanation whose explanatory link is identical to its explanatory result.[218] Schematically, such an 'empty-base self-explanation' has this form:

Base: $\varnothing$
Link: P
Result: P

In such an explanation, the result explains$_{\text{inclusive}}$ itself by being the link of its own empty-base explanation. Note that since there are no explanations without a link, self-explanations in the *restrictive* sense will likely involve a proposition that is distinct from its result, i.e. the explanatory link.[219] In contrast, an empty-base self-explanation would only involve one proposition, namely its explanatory result and link. Thus, in a sense, only an empty-base self-explanatory proposition would be *fully* self-explanatory in the sense of having an explanation with just it as a constituent, and only such explanations could be truly ultimate in that they do not involve any propositions that are unexplained or only explained by further explanations.

Before we consider candidates for empty-base self-explanations, let me address two arguments by Kovacs (2018, sec. 4) against the possibility of self-explanation that do not follow the pattern from section 6.1. In his first argument, Kovacs argues that just like circular ordinary arguments, circular explanatory arguments are objectionable, because just like ordinary arguments, explanatory arguments are supposed to provide reasons for their conclusions, but circular (ordinary as well as explanatory) arguments do not provide such reasons. Since Kovacs further assumes that every case of self-explanation corresponds to a circular explanatory argument, he concludes that self-explanation is objectionable.

---

218 We could in principle also consider explanations whose link and result are identical, but whose base contains different propositions, but these would not be *full* self-explanations.

219 'Likely' since we could in principle consider explanations whose reason, link, and result are identical. I will set aside this issue for what follows.

In response note first that an explanation whose result and link are identical is structurally related to the notion of rule-circular justification: In such an explanation, an explanatory link (partially) explains itself. Therefore, the corresponding explanatory argument has a conclusion that corresponds to the explanatory rule that governs the argument.[220] Similarly, a rule-circular justification of an inference principle is provided by an argument to the conclusion that the principle in question holds (or perhaps to a conditional that corresponds to the inference principle), but which uses the inference principle in question to establish this.[221]

While some (e.g. Boghossian 2001) have endorsed the idea that rule-circular arguments may provide justification for their conclusions, their epistemic value is doubtful (for a recent criticism see Carter and Pritchard 2017). But note that even if the possibility of rule-circular *justification* is denied, the impossibility of empty-base self-explanation does not obviously follow: From the impossibility of rule-circular justification it would *prima facie* merely follow that if empty-base self-explanation is possible, then there are possible explanatory arguments that do not *justify* their conclusion, but they might still *explain* it.

Moreover, *pace* Kovacs, the premises of a good ordinary (or *epistemic*) argument *justify* its conclusion, viz. they are *epistemic* reasons for its conclusion, but the premises of a good explanatory argument *explain* its conclusion, they are reasons *why* the conclusion obtains. Kovacs appears to conflate these two notions of reasons and assumes that good explanatory arguments must justify (i.e. provide *epistemic* reasons for) their conclusions, but in many cases (e.g. many instances of inference to the best explanation), it is rather the case that a conclusion of an explanatory argument (i.e. an explanandum) justifies a premise of said argument (i.e. part of a corresponding explanans).

Kovacs' (2018, 1170) second argument turns on considerations about the relation between explanation and understanding:

> [For] a statement such as '*p* explains *q*' to express a genuine explanation, there should be a possible cognitive state of non-understanding, best expressed by the question 'Why *q*?', and an answer, '*p*', learning of which replaces this state of non-understanding with a state of understanding. To achieve this goal, explanations have to be informative in the sense that the explanans clause conveys information not provided by the explanandum clause, or at least conveys information in a way not provided by the explanandum clause. Note that this requirement doesn't mean

220 Cf. Litland's (2017) calculus for explanatory arguments.

221 The analogy is not perfect: The result of an empty-base self-explanation is a proposition that is identical to its link. In contrast, the conclusion of a rule-circular argument is a proposition stating that a certain inference principle (that, moreover, arguably is not a proposition) holds. Thanks to an anonymous commenter on the paper on which this chapter is based.

> that the explanans clause conveys information to every audience that the explanandum clause doesn't convey in the same way, only that it's capable of doing so in the right circumstances.

In the above paper, Kovacs wants to argue that self-grounding is impossible and he does so by first arguing that self-grounding would give rise to self-explanation and then providing arguments against self-explanation. Thus, the intended targets of this argument are, in our terminology, self-explanations in the restrictive sense. But the argument is not convincing:

First, recall chapter 1, the tripartite account of explanation and its connection to why-questions and because-sentences: The explanans of an explanation (properly understood) has two components: The base component which is comprised of reasons why the explanandum obtains and which can be used to answer corresponding why-questions, and the link component which connects base and explanandum. Given these assumptions, the proponent of self-explanation in the restrictive sense can grant that the explanans needs to convey information not provided by the explanandum clause '*P*', while maintaining that '*P*' (or rather '*P* because *P*') is a possibly correct answer to 'Why *P*?': The additional information conveyed by the explanans is then located in its link component, e.g. a proposition to the effect that [*P*] grounds [*P*].

Second, and supporting this point, observation of cases reveals that the step from a lack of understanding why towards understanding why often does not consist in coming to know the base-elements of the corresponding explanation why, but rather in coming to know (or to grasp) its link. For instance, many situations involving inference to the best explanation are like this: Sherlock may already know that the window is broken and that both Watson and Moriarty threw balls at the window, but coming to understand why the window broke involves grasping the causal (or law-like) link between Moriarty's throwing his ball and the window's breaking (see chapter 5 for more discussion).

While I am skeptical of self-explanation in the restrictive sense, Skiles (manuscript) has pointed out that Kovacs' argument might apply to explanations involving zero-grounding (and, we can add, empty-base explanations generally): Since the base of such explanations is empty, it does not contain any information that might lead to understanding. It seems then that Kovacs' argument would have us conclude that empty-base explanations are not possible. But again, empty-base explanations do involve another component, namely the link, grasping which can amount to understanding why the corresponding explanandum obtains. As argued in chapter 2, if we answer 'Why *P*?' with '*P* just because' or '*P* because' in the senses proposed there, we can communicate the relevant link.

Now, Kovacs' information constraint on explanation is more problematic for the notion of empty-base self-explanation, with which we are concerned here, because

such explanations have an empty base (so there is no information to be found there), while the explanandum and the link are identical. Thus, neither base nor link provide information beyond the explanandum. In response, recall that in chapter 1 I have argued that mere knowledge of an explanatory link (plus base) need not be sufficient for understanding why the corresponding explanandum obtains. Rather, a mental state of grasping the link plus some associated cognitive control over the relationship is required. If this is correct, then there can be a possible cognitive state of non-understanding why *P* that is compatible with knowledge of [*P*] and [*P*] being the link and result of an empty-base self-explanation. For now, let us proceed to develop this notion further and look at candidates for empty-base self-explanations, I will say a bit more about this argument once we discuss generalized explanatory links below.

6.5 Candidates for empty-base self-explanations

Now, what would empty-base self-explanations look like? Recall the suggestion that explanatory links of empty-base explanations have the form '■*P*', where '■*P*' stands for the result of the corresponding empty-base explanation. Since explanatory links of empty-base self-explanations are identical to the result of their explanation, it follows from this that their links have the form '■*P*' and that the proposition [*P*] is identical to the proposition [*P*]. Call this the *formal criterion*.

Now the question is whether there can be propositions of this form. Using 'is *R*-related to' as a placeholder for relational predicates used to express explanatory links and 'is zero-*R*' as a placeholder for predicates used to express corresponding empty-base links. We can state the form of self-explanatory links as 'The proposition that *P* is zero-*R*', where the proposition expressed is identical with the proposition that *P*. Consider grounding as an example. Predicational zero-grounding statements have the form 'The proposition that *P* is zero-grounded'. Thus, if there are empty-base self-explanations of the grounding variety, the corresponding self-explanatory propositions have the form 'the proposition that *P* is zero-grounded', where the proposition that *P* is identical with the proposition that the proposition that *P* is zero-grounded. Indeed, here is a candidate that has this form:

(3) This proposition is zero-grounded.

Here, the expression 'This proposition' in (3) is intended to refer to the proposition expressed by (3). Note that while some propose that certain self-referential (e.g. paradoxical, liar-type) sentences do not express propositions, the self-referential nature of (3) alone is presumably not sufficient to assume that (3) expresses no proposition;

after all, many (apparently) unproblematic self-referential sentences exist.[222] But now note how (3) resembles the truth-teller 'This sentence is true': If we had to speculate about the truth-value of (3), it would not seem unreasonable to assign it the same truth-value as the truth-teller, which, many are inclined to believe, is defective and neither true nor false.[223] And even if (3) were true, it presumably could not fulfill the high hopes some philosophers have put into self-explanatory propositions: Intuitively, (3) is somewhat thin in content, which is, perhaps, exactly what is to be expected of a zero-grounded proposition. Consequently, it is hard to see how it could serve the idea that a substantial class of truths are eventually explained by self-explanatory propositions.

One might perhaps think that instances of the following schema could do better in this regard (let '*P*' stand for an arbitrary proposition and '*4*' express the proposition labeled by '(4)'):

(4) The proposition that (*P* and *4*) is zero-grounded.

But this is problematic because (4) seems to fail the formal criterion: If we eliminate the zero-grounding operator from (4), we obtain '*P* and *4*', which does not seem to be identical with (4), in part because (4) expresses a proposition with a zero-grounding operator having largest scope, whereas in '*P* and *4*', the conjunction operator has largest scope. We could perhaps allow that *some* conjunctions are identical (or at least suitably equivalent) to one of their conjuncts, this is for example possible according to certain worldly modes of identifying propositions or facts (e.g. Correia 2016). Then to vindicate the possibility of self-explanations of the above form, one would have to find a mode of individuation suited to deliver instances of (4) satisfying the formal criterion, but such an investigation goes beyond the scope of this chapter.

222 E.g. 'This proposition is a proposition', 'Every proposition is a proposition' and 'This proposition is such that 1+1=2'. Cf. Rosenkranz and Sarkohi (2006). As an anonymous commenter on the paper on which this chapter is based has stressed, it could be thought that the candidates considered here and in the next subsection would amount to *objectionably* ill-founded propositions. Development of a theory of propositions that would vindicate the existence of the candidates would go beyond the scope of this book, but let me note that the candidates are not obviously defective in this way and that at least with respect to (3), I am not alone in this assessment, cf. Lovett (2020). One reservation here might stem from an understanding of propositions as mereological wholes, but first this understanding is not mandatory, and second see Kearns (2011) for an argument that on such a view we should simply accept that at least certain (otherwise unproblematic) self-referential propositions are parts of themselves. For an investigation into the non-well-founded mereology required for this, see Cotnoir and Bacon (2012).

223 Cf. Field (2008), but note also Field (2008, 277).

Instead, here are three further options to find (perhaps more substantial) candidates for empty-base self-explanations: First, one could attempt to find an explanatory relation *R* such that 'This fact is zero-*R*' is more substantial and less like the truth-teller than (3). The second option invokes Dasgupta's (2014a) proposal that grounding is irreducibly plural, and the third considers laws as explanatory links.[224] Setting aside the first option, we will now look at the second and third options in turn.

6.5.1 Irreducibly plural grounding

According to Dasgupta (2014a), grounding is irreducibly plural in the following sense: (predicational) grounding statements have the form 'The *Y*s are grounded in the *X*s', where '*Y*' and '*X*' are schema-letters for expressions denoting pluralities of facts, and it is possible that the *Y*s are grounded in the *X*s, without any of the *Y*s on its own being grounded in the *X*s. For example, Dasgupta argues that the individualistic facts (i.e. facts concerning particular individuals, like [Socrates is a philosopher] or [Obama is 75 kgs]) are together irreducibly plurally grounded in purely qualitative facts. That is, for example, individualistic facts about the mass of particular individuals are plurally grounded in purely qualitative facts capturing the mass relations between things, but no single fact about the mass of a particular individual is grounded in such facts on its own.

Correspondingly, plural zero-grounding statements can be expressed by having '*X*' denote an empty plurality; alternatively, 'The *Y*s are zero-grounded' can be used. Dasgupta's proposal then allows for more contentful candidates for empty-base self-explanation by allowing for a plurality of propositions to occur as (joint) groundees in a grounding statement like this:

(5) This fact, [*P*] are zero-grounded.

Here, 'This fact' refers to the fact expressed by (5). Assuming with Dasgupta that there are irreducibly plural instances of grounding, an instance of (5) might in principle obtain without it being singularly zero-grounded, while at the same time being plurally zero-grounded together with [*P*].

---

224 A fourth option could perhaps be this: Returning to the assumption that links of empty-base explanations have the form '■*P*', one might consider the possibility of prefixing a right-side infinite sequence of '■*P*'s to a sentence '*P*' like this: '■■■ ... *P*'. Here, when the outermost '■' is eliminated, arguably, a sentence of the same form '■■■ ... *P*' remains; but to my knowledge, a theory of non-well-founded propositions like this would yet have to be motivated and developed.

Now, is there any reason to assume there being self-explanatory facts of the form of (5)? What kind of facts would be suitable to be collectively zero-grounded, where one of the collectively zero-grounded facts is the corresponding collective zero-grounding fact itself? Dasgupta's examples for collectively grounded facts all involve facts that are similar in some respect (like the individualistic facts).

Therefore, a natural candidate for our collectively zero-grounded facts are other (non-factive) grounding facts. According to this idea, all non-factive grounding facts would be irreducibly collectively zero-grounded, including this collective non-factive grounding fact itself. One tentative advantage this proposal has over Litland's (2017) original proposal (according to which non-factive grounding facts are zero-grounded) is that it avoids the following somewhat awkward regress: According to Litland's proposal, [$P \rightarrow Q$] is zero-grounded, [[$P \rightarrow Q$] is zero-grounded] is zero-grounded, [[[$P \rightarrow Q$] is zero-grounded] is zero-grounded], etc.; according to the present proposal there is just one self-referential collective zero-grounding fact here.[225]

6.5.2 Generalized explanatory links

Let us finally consider how generalized links, such as laws of the following form might help (let '$\Box_L$' stand for a law operator like the metaphysical law operator):

(LAW) $\Box_L \forall x(Fx \rightarrow Gx)$

The idea is this: An ordinary generalized explanatory link can serve as an explanatory link of many explanations by linking different bases with different results. A generalized link of an empty-base explanation could in turn figure in explanations with several different results. Thus, in principle, there might be such a link which is the result of an empty-base explanation and which thus explains itself, but which in addition is the link of a further (possibly empty-base) explanation with a different result. Incidentally, the idea is reminiscent of Nozick's idea of "explanatory self-subsumption":

225 If one considers this regress to be more problematic than merely somewhat awkward, one might additionally reason as follows. What the regress shows is that some explanatory work remains to be done at each step and is hence deferred ad infinitum. Hence, a non-factive grounding fact cannot be fully zero-grounded *on its own*. But given the present idea, Litland's proposal can be amended: Non-factive grounding facts might not be individually zero-grounded, but they are all collectively zero-grounded.

> The objectionable examples of explanatory self-deduction (total or partial) involve deductions that proceed via the propositional calculus. Would the explanation of a law be illegitimate automatically if instead the law was deduced from itself via quantification theory, as an instance of itself? If explanation is subsumption under a law, why may not a law be subsumed under itself? (Nozick 1981, 119ff.)

Here, Nozick appears to suggest that the permissibility of self-explanation somehow depends on whether the involved explanatory steps correspond to rules of the predicational calculus as opposed to the propositional calculus, but this does not seem very convincing: Just consider the question of whether universal generalizations are grounded in their instances or whether they ground their instances: While both options *may* have some initial plausibility, we should not accept both on pain of violating the asymmetry of grounding.

But we can ignore this part of Nozick's suggestion, and then the above considerations about empty-base self-explanation can help capture his idea of a self-subsuming explanatory law. Nozick (1981, 119) does not properly distinguish between the roles of explanatory link and base; for example, he takes a self-subsuming principle to be an (explanatory) reason of itself. But if we make the distinction and understand explanatory self-subsumption as a kind of empty-base self-explanation, we can explain why explanatory self-subsumption may seem possible, namely because the simple arguments against self-explanation then do not apply to it. Moreover, the idea of explanatory self-subsumption gives us a further resource to address Kovacs' second argument: Someone who knows a self-subsuming explanatory principle might not have grasped it fully and thus might not have realized that it is self-subsuming. Thus, such a person may wonder why the self-subsuming principle obtains. To understand why the principle obtains, this person then need not obtain further information, rather they need to grasp that the principle is self-subsuming and can thus take them from the empty base of reasons to the principle itself.

Let us think a little about the form self-explaining links à la Nozick would have to take. Let us consider unconditional links involving both quantification over entities and into sentence position. We can furthermore consider ordinary quantification or quantification into sentence position. Empty-base law-like links could then for example have one of the following forms (let '*O*' schematically stand for a sentential operator):

(L1) $\Box_L \forall x(Gx)$
(L2) $\Box_L \forall p(Op)$

It is unclear to me whether there could be an instance of (L1) that satisfies the formal criterion, i.e. an instance such that one of the instances of the involved

quantification is identical to the proposition that is the whole link.[226] But consider (L2): Could there be an instance for '*O*' and a proposition [*P*] such that the proposition $\Box_L[\forall p(Op)]$ is identical to the proposition [*OP*]? Well, such instances are provided by the $\Box_L$-operator and the proposition $[\forall p(\Box_L p)]$:

(L3) $\Box_L \forall p(\Box_L p)$

If the quantifier is understood as ranging over all propositions, the result is absurd because for no false proposition [*P*] is it the case that $\Box_L P$. This problem can be avoided if we instead understand the quantifier as ranging over all *facts*. The result is a candidate explanatory link according to which every fact is a law. While this will strike many as only marginally more plausible, the result is still interesting: Some philosophers have been moved to admit self-explanatory facts by their acceptance of the PSR. The PSR has also moved some to endorse necessitarianism, the idea that every fact is necessarily the case.[227] (L3), properly understood, embodies these two rationalist ideas: It is self-explanatory and it states a variant of necessitarianism according to which every fact is a law.[228]

Let us take stock: While it is unclear whether there are more plausible candidates for empty-base self-explanation, we have made progress towards answering whether empty-base self-explanation is possible by clarifying what it would take for them to exist. If we are pessimistic about the prospects of empty-base self-explanation, we have at least gained a better understanding of why this kind of self-explanation does not exist: Not because 'explains$_{\text{inclusive}}$' is irreflexive, as the arguments of section 6.2 would have it, but because it is hard to find substantial and plausible propositions of the required form.

226 If we assume, e.g., that [*P*] and [[*P*] is the case] to be identical, then '$_L\forall x(x$ is the case$)$' is an instance of (L1) that satisfies the criterion, but this example faces similar issues to those discussed below. The issue here is to find an instance that satisfies the formal criterion without being too implausible.

227 Spinoza is an example for both moves, cf. Della Rocca (2010) and Lærke (2011), but see Schnieder and Steinberg (2015) on how proponents of the PSR can avoid either consequence.

228 One idea worth considering might be to restrict the quantifier in (L3) such that it still ranges over (L3) itself, but does not range over all facts, thereby avoiding the consequence that every fact is a law.

6.6 Empty-base self-explanation meets philosophical theology

Let me end the chapter by showing how the notions of empty-base explanation and empty-base self-explanations might inform our understanding of certain ideas about the explanation of the existence of God. According to many scholastics like Aquinas, but also according to some later philosophers like Spinoza, God's essence involves God's existence.[229] This alone suggests a way in which God's existence might be explained, namely by its status as being part of the essence of God. Using the conceptual apparatus developed above, the idea can be put like this: God's existence is empty-base explained, and the explanatory link of this explanation is the fact that it is part of God's essence that God exists.

Now, both Aquinas and Spinoza go further in that they also believe that God's existence is *identical* to God's essence.[230] But this provides the material for a proposal for an empty-base self-explanation of God's existence: God's essence, i.e. the fact that it is part of God's essence that God exists, would be the empty-base link of this explanation and God's existence would be the explanatory result of this explanation. But according to both Aquinas and Spinoza, God's essence *just is* God's existence. If we understand this identity as the identity between the fact that God exists and the fact that it is part of God's essence that God exists, then the result is a proposal for an empty-base self-explanation.

Some remarks: First, by understanding their proposal as concerning empty-base self-explanations, both Aquinas and Spinoza might avoid the arguments against the intelligibility of self-explanation, as I have argued above. Second, the proposal is confronted with an issue we have encountered already: It is unclear that the required claim concerning the identity between the explanandum and the explanatory link can be made sense of. Third, while Aquinas' and Spinoza's shared assumptions allow for a proposal for a self-explanation of God's existence without the need to claim that God's existence is its own reason why (e.g. its own ground or cause), Spinoza appears to explicitly want to claim that God is her own cause, i.e. a *causa sui* and thus reason why.[231]

229 Lærke (2011, 447f.).
230 Cf. McInerny and O'Callaghan (2018, sec. 11.3) for Aquinas and Lærke (2011, 456) for Spinoza.
231 Cf. Lærke (2011).

6.7 Conclusion

Let us recapitulate: Using the tripartite account of the structure of explanations, I have distinguished two notions of self-explanation, defended one against several arguments against the possibility of self-explanation, and applied it in a solution of the circularity problem for Humeanism about laws of nature. In the remainder of the chapter, I have developed and defended the notion of an empty-base self-explanation and suggested some applications for it.