Synthese (2018) 195:1337–1353
https://doi.org/10.1007/s11229-016-1274-y

Transitivity, self-explanation, and the explanatory circularity argument against Humean accounts of natural law

Marc Lange[1]

Received: 5 July 2016 / Accepted: 11 November 2016 / Published online: 25 November 2016

Abstract Humean accounts of natural lawhood (such as Lewis's) have often been criticized as unable to account for the laws' characteristic explanatory power in science. Loewer (Philos Stud 160:115–137, 2012) has replied that these criticisms fail to distinguish grounding explanations from scientific explanations. Lange (Philos Stud 164:255–261, 2013) has replied by arguing that grounding explanations and scientific explanations are linked by a transitivity principle, which can be used to argue that Humean accounts of natural law violate the prohibition on self-explanation. Lange's argument has been sharply criticized by Hicks and van Elswyk (Philos Stud 172:433–443, 2015), Marshall (Philos Stud 172:3145–3165, 2015), and Miller (Philos Stud 172:1311–1332, 2015). This paper shows how Lange's argument can withstand these criticisms once the transitivity principle and the prohibition on self-explanation are properly refined. The transitivity principle should be refined to accommodate contrasts in the explanans and explanandum. The prohibition on self-explanation should be refined so that it precludes a given fact p from helping to explain why some other fact q helps to explain why p. In this way, the transitivity principle avoids having counterintuitive consequences in cases involving macrostates having multiple possible microrealizations. The transitivity principle is perfectly compatible with the irreducibility of macroexplanations to microexplanations and with the diversity of the relations that can underwrite scientific explanations.

Keywords Scientific explanation · Grounding · Transitivity · Humean accounts of natural law · Laws of nature · Multiple realizability · Contrastive explanation

✉ Marc Lange
mlange@email.unc.edu

[1] University of North Carolina, CB#3125, Chapel Hill, NC 27514, USA

1 Introduction

According to Humean views of natural law [such as Lewis (1986) Best System Account], the lawhood of a given fact (such as the fact that all sodium salts burn yellow) is constituted by some feature of the global spacetime mosaic of Humean facts. Many philosophers, such as Armstrong (1983, p. 102), have criticized Humean views of lawhood as unable to account for the laws' important role in scientific explanations—a role generally recognized by Humeans and non-Humeans alike. This objection to Humean accounts is expressed nicely by Maudlin:

> If the laws are nothing but generic features of the Humean Mosaic, then there is a sense in which one cannot appeal to those very laws to *explain* the particular features of the Mosaic itself: the laws are what they are in virtue of the Mosaic rather than vice versa. (Maudlin 2007, p. 172)

In accordance with recent interest (e.g., Rosen 2010; Fine 2012) in "grounding" as underwriting a distinctive, metaphysical sort of explanation, Loewer (2012) has replied that the above objection to Humean views fails to attend to the distinction between *scientific* explanation and *metaphysical* explanation:

> On Lewis' account the Humean mosaic *metaphysically* determines the … laws. It metaphysically explains (or is part of the explanation together with the characterization of a Best Theory) why specific propositions are laws. This metaphysical explanation doesn't preclude … laws playing the usual role of laws in scientific explanations. (Loewer 2012, p. 131)

According to Loewer, there is no difficulty in laws being partly *scientifically* responsible for the mosaic while the mosaic is partly *metaphysically* responsible for the laws.

Lange (2013) has replied that Loewer's view violates the prohibition on self-explanation because even if we grant "for the sake of argument" the distinction between grounding and scientific explanation, these two kinds of explanation are linked by a "transitivity principle" (for facts D, E, and F):

> If E scientifically explains [or helps to scientifically explain] F and D grounds [or helps to ground] E, then D scientifically explains [or helps to scientifically explain] F. (Lange 2013, p. 256)

Suppose the fact that all F's are G is a law helps (together with the fact that Fa) to scientifically explain the fact that Ga. Suppose also (as Humean views of law maintain) that the fact that Ga (part of the Humean mosaic) helps to ground the fact that all F's are G is a law. Then (Lange argues) by the transitivity principle, the fact that Ga must help to scientifically explain the fact that Ga, violating the prohibition on self-explanation.[1]

[1] Perhaps the prohibition on self-explanation is violated in some exotic cases, such as when time-travel occurs. However, following Marshall (2015, pp. 3150–3151, n. 13), let's grant that it is not violated in routine cases of scientific explanation, even if it is violated in certain pathological cases. Following Miller (2015, p. 1325), Marshall (2015, p. 3152, n. 14) also notes that Lange's objection to the Humean view of law can be reformulated without appealing to the prohibition on self-explanation; routine scientific self-explanation

This explanatory circularity argument against the Humean view of law is like other "philosophical arguments aiming to show that some D cannot ground some E, because otherwise (since E scientifically explains F) D would have to scientifically explain F" (Lange 2013, p. 257), violating the prohibition on self-explanation. Here is an example. The difference in fitness between two types of moths (e.g., dark-colored and light-colored) is used by evolutionary biologists to help scientifically explain why, on average, members of one type (the fitter) have a higher number of offspring than members of the other type. But then it cannot be that the fitness difference between these types is grounded in the difference in their average numbers of offspring, since this would lead (by transitivity) to what Mills and Beatty (1979, p. 265) call "justifiable charges that certain explanations invoking fitness differences are circular":

> Where fitness is defined in terms of survival and reproduction success, to say that type A is fitter than type B is just to say that type A is leaving a higher average number of offspring than type B. Clearly, we cannot say that the difference in fitness of A and B *explains* the difference in actual average offspring contributions of A and B, when fitness is defined in terms of actual reproductive success. (Mills and Beatty 1979, p. 265)

By the transitivity principle, this charge of self-explanation cannot be parried by pointing out that fitness differences *scientifically* explain differences in average numbers of offspring, whereas differences in average numbers of offspring *metaphysically* explain fitness differences.

Lange's argument has recently been sharply criticized by Hicks and van Elswyk (2015), Marshall (2015), and Miller (2015). In this paper, I will reply to many of their criticisms. My replies will also have considerable independent interest—not only because they emphasize important features of grounding and scientific explanation, but also because my replies propose refinements to both the transitivity principle and the prohibition on self-explanation.

In Sect. 2, I will respond to the objection that the transitivity principle has counterintuitive consequences in cases involving multiply realizable macrostates figuring in scientific explanations and grounded, in turn, by their microrealizations. I will argue that the transitivity principle does not have counterintuitive consequences in these cases once the principle is refined to recognize the *contrasts* in the explanandum and the explanans. This refined transitivity principle can be used to give the explanatory circularity argument against the Humean view of lawhood.

In Sect. 3, I will respond to the objection that the transitivity principle fails to respect the diversity of the relations that can underwrite scientific explanations. In reply, I will argue that transitivity applies even when D *non-causally* scientifically explains E while E *causally* scientifically explains F. I will also respond to the objection that transitivity fails to do justice to the irreducibility of scientific explanations that appeal to macrolevel facts. I will argue that the transitivity principle stands in no tension with

Footnote 1 Continued
is just one of the intuitively implausible consequences of the Humean view of law when it is coupled with the transitivity principle.

macrolevel scientific explanations making explanatory contributions that could not be made, even in principle, by microlevel scientific explanations.

Finally, in Sect. 4, I will respond to a series of objections that focus on the role of laws in and around scientific explanations. These objections all argue that the circularity argument fails to undermine the Humean view of law because a given fact's lawhood is not among a Humean fact's explainers, but rather performs some other function in connection with scientific explanations. Ultimately, I will refine the prohibition against self-explanation so that it applies even if a fact's lawhood helps to explain not why a given Humean fact holds, but rather why some other fact helps to explain why the given Humean fact holds.

The objections offered to the explanatory circularity argument require that the argument be refined in various respects. But those refinements all have independent motivations. Therefore, I conclude, the objections ultimately fail to blunt the force of the explanatory circularity argument against Loewer's (2012) defense of the Humean view of lawhood.

2 Multiple realizability, transitivity, and contrastive explanation

Hicks and van Elswyk (2015, pp. 437–438) and Miller (2015, pp. 1321–1324) offer the same objection to the transitivity principle. They object that the principle faces counterexamples in which D does not help to scientifically explain F even though E helps to scientifically explain F. In these counterexamples, E cites "what are in some sense higher-level or general" facts and D helps to ground E, where D is a "particular grounding fac[t] about the mosaic" that is "incidental to" F (Miller 2015, p. 1321). That D does not help to scientifically explain F has two motivations: (i) that E (which helps to scientifically explain F) is multiply realizable and so could have obtained without D, and (ii) that D, in the absence of other lower-level facts, would have obtained without E (Miller 2015, p. 1322). Hicks and van Elswyk (2015, pp. 437–438) give this example: electron e's having a certain position helps to ground lion L's having a certain position, and L's position (in turn) helps to scientifically explain the number of prey animals in a given region. "But the position of electron e does not explain" (or even help to explain, Hicks and van Elswyk would presumably say) "the number of prey animals in region R. For if the electron were elsewhere, L would still be warding prey animals out of R".

Of course, one could grant the truth of this counterfactual (as I do) and still insist that the electron's presence at the given location helps to scientifically explain how few prey animals there are in the given region. As is well known, cases of overdetermination involve D helping to scientifically explain F even though F would still have obtained, had D not obtained.[2]

[2] Marshall (2015, p. 3152) makes exactly this reply to Hicks and van Elswyk. Also note that Hicks and van Elswyk say that in this example, D does not help to *scientifically* explain F because F would still have obtained had ∼D, and yet Hicks and van Elswyk say that D helps to *metaphysically* explain E even though E would still have obtained had ∼D. Hicks and van Elswyk need some account of why counterfactual dependence is required for *scientific* explanation but not for *metaphysical* explanation. (Thanks to Chris Dorst.)

Nevertheless, the objection cannot be so easily dismissed. I recognize a strong intuition that in such examples, which could be multiplied endlessly (as Hicks and van Elswyk (2015, p. 438) rightly remark), D helps to metaphysically explain E and E helps to scientifically explain F even though D does not help to scientifically explain F. There is something correct that we might naturally convey by remarking that electron e's presence at a given location does not help to explain the number of prey creatures in a given region—even though e's presence at the given location helps to metaphysically explain a given lion's presence in the given region, which (in turn) helps to scientifically explain the number of prey creatures there. In the rest of this section, I will identify what is correct in the above remark and why it does not conflict with a transitivity principle that can underwrite the explanatory circularity argument against the Humean view of lawhood. To do so, I will appeal to an important feature of many explanations: that they are *contrastive*.[3]

Explanations often employ contrasts in both their explanans and their explanandum. To take an example (slightly amended) from Lipton (2004, pp. 33, 36): my having a liking for contemporary non-musical dramas rather than a liking for mid-twentieth century American musical comedies explains why I went out to see Stoppard's *Jumpers* rather than Bernstein's *Candide* last night, but it does not explain why I went out at all last night rather than stayed home. What does contrastive explanation require? For the fact that X obtains rather than X' to help explain why it is the case that Y obtains rather than Y', it must be that X stands in a certain explanatory relation to Y (e.g., helps to cause Y) and also it must be that had X' obtained instead of X, then X' would (or, at least, might well) have stood in a relation to Y' roughly corresponding to the relation in which X stands to Y.[4] For example, my liking for contemporary non-musical dramas helped to cause me to go out to see *Jumpers*, but if I had liked mid-twentieth-century American musical comedies, then that attitude might well in the same way have helped to cause me to go out to see *Candide*—but it would not have helped to cause me to stay home. Similarly (to use an example from Lewis (1986), pp. 229–230), that Lewis received an invitation to visit Monash rather than an invitation to visit Oxford in 1979 explains why he visited Monash rather than Oxford in 1979: his invitation to Monash was a cause of his visiting there, and had Lewis been invited to Oxford instead of Monash, then his Oxford invitation might well (in the same way) have caused him to visit Oxford.

Often the contrasts in play are safely left implicit, though they can be made explicit. Here is the natural way to refine the original transitivity principle (which made no explicit mention of contrasts) to make the contrasts explicit:

> TRANSITIVITY: If the fact that E rather than E' scientifically explains [or helps to scientifically explain] the fact that F rather than F', and if the fact that D rather than D' grounds [or helps to ground] the fact that E rather than E', then the fact

[3] Hicks and van Elswyk (2015, p. 438) acknowledge in a footnote: "Perhaps a dedicated anti-Humean could revive the circularity objection with a contrastive [transitivity] principle. Having not seen such an argument, we are agnostic of its cogency". They pursue the matter no further. I accept their invitation now. That explanation is contrastive was influentially emphasized by van Fraassen (1980) and Garfinkel (1981).

[4] Lipton (2004, pp. 42–43) makes a very similar proposal. Nothing here will turn on the details.

> that D rather than D' scientifically explains [or helps to scientifically explain] the fact that F rather than F'.[5]

TRANSITIVITY captures the fact that the contrasts in each step have to line up in order for a transitivity principle to apply. To see an example, let's assume that the location of a seesaw's center of mass (which explains whether and in what direction the seesaw tips) is grounded (at least partly) in the masses of the bodies sitting on either arm of the seesaw. The fact that the seesaw's center of mass is directly over its base of support rather than slightly to the right helps to scientifically explain the fact that the seesaw balances rather than tips to the right, and the fact that Jones (on the right arm of the seesaw) weighs 90 pounds rather than 120 pounds helps to ground the fact that the seesaw's center of mass is directly over its base of support rather than slightly to the right. According to TRANSITIVITY, it follows that the fact that Jones (on the right) weighs 90 rather than 120 pounds helps to scientifically explain the fact that the seesaw balances rather than tips to the right. This conclusion is plausible. On the other hand, if the contrasts in each step do not line up, then the corresponding conclusion does not follow. For example, the fact that Smith (on the left) weighs 90 rather than 120 pounds helps to ground the fact that the seesaw's center of mass is directly over its base of support rather than slightly *to the left*, and (as we know) the fact that the seesaw's center of mass is directly over its base of support rather than slightly *to the right* helps to scientifically explain the fact that the seesaw balances rather than tips to the right. But it does not follow (indeed, it is false) that the fact that Smith (on the left) weighs 90 rather than 120 pounds helps to scientifically explain the fact that the seesaw balances rather than tips to the right. That is because it is not the case that the *left* side's weighing more would (or even might well) have helped to cause the seesaw to tip to the *right*.

Let's now take account of the contrasts in the putative counterexample involving the lion and electron. TRANSITIVITY might seem to lead to the admittedly false conclusion that electron e's presence at rather than absence from a given location helps to scientifically explain why there are few rather than many prey creatures in the surrounding region. This false conclusion follows by TRANSITIVITY from two premises: (i) The lion's presence at rather than absence from the region helps to scientifically explain why there are few rather than many prey creatures there, and (ii) Electron e's presence at rather than absence from a given location helps to metaphysically explain the lion's presence in rather than absence from the given region. However, premise (ii) is false: Although e's presence at the given location helps to metaphysically explain the lion's position, it is not the case that e's presence there *rather than e's absence from there* helps to metaphysically explain the lion's presence there rather than absence from there. That is because had e been absent, then it is not the case that e's absence would (or even might well) have helped to metaphysically explain the lion's absence, since it is not the case that had e been absent, the lion would (or even might well) have been absent; the lion would simply have been present without e. So we cannot argue against TRANSITIVITY on the grounds that it would saddle us with the false

[5] Schaffer (2005, p. 310) argues that causal relations are contrastive and proposes a principle analogous to TRANSITIVITY, but exclusively concerning causal relations.

conclusion that e's presence at rather than absence from a given location helps to scientifically explain why there are few rather than many prey creatures in the surrounding region.

We might naturally convey this conclusion's falsehood by remarking that e's presence at a given location does not help to explain the number of prey creatures in a given region. This remark leaves all of the contrasts implicit. In this way, I vindicate the strong intuition I mentioned earlier that apparently motivated Hicks, van Elswyk, and Miller to object that the transitivity principle fails in cases of multiple realizability. I have argued that contrasts allow their intuition to be reconciled with the transitivity principle.

The false premise in the lion case that led TRANSITIVITY to a false conclusion was (ii): Electron e's presence at rather than absence from a given location helps to metaphysically explain the lion's presence in rather than absence from the given region.[6] By employing a different contrast, we can replace the false premise (ii) with a true premise (ii'): The presence of a leonine configuration of particles consisting of electron e together with all of the lion's other constituent particles, rather than the absence from the region of any leonine configuration of particles at all, helps to metaphysically explain the lion's presence in rather than absence from the given region. By applying TRANSITIVITY to this premise together with premise (i), we arrive at the conclusion that the presence of a leonine configuration consisting of e together with all of the lion's other constituent particles, rather than the absence from the region of any leonine configuration of particles, helps to scientifically explain why there are few rather than many prey creatures there. This conclusion is true, so TRANSITIVITY's role in generating it supplies no argument against TRANSITIVITY.

Inspired by premise (ii'), we could try to find a true premise involving something like the contrast that figures in premise (ii)—namely, between e's presence and e's absence. Here is a candidate: Given the rest of the actual leonine configuration of particles, e's presence in that configuration rather than absence from it helps to metaphysically

[6] In deeming this premise false, I have presumed that it is not the case that had e been absent, the lion would (or might well) have been absent; the lion would simply have been present without e. But the truth-values of counterfactuals are context-sensitive, so perhaps there are contexts where it is true that had e been outside of region R, then the lion would have been outside of region R. To see that TRANSITIVITY yields a plausible conclusion in a context where such a counterfactual holds, let's change the example to one where it is easier to imagine such a counterfactual holding. Suppose that the presence of a gold bar in room R of the museum helps to scientifically explain why there are so many visitors in R. (The bar attracts a crowd; had the bar been on display in room S of the museum instead of R, the crowd would have been in S and there would have been few people in R.) Consider a particular gold atom g that is in the bar. We can easily imagine a context where it is true that had atom g been in S rather than R, then the entire bar would have been in S rather than R. Accordingly, analogous to premises (i) and (ii) in the lion case, we have the following two premises: The bar's presence in R rather than S helps to scientifically explain why there are many rather than few people in R, and atom g's presence in R rather than S helps to metaphysically explain the bar's presence in R rather than S. TRANSITIVITY yields the conclusion that atom g's presence in R rather than S helps to scientifically explain why there are many rather than few people in R. This seems plausible to me, bearing in mind that the entire bar would have been elsewhere, had g been elsewhere. (However, by singling out atom g for special attention, this counterfactual may give a false impression; to remove this impression, it suffices to add that of course, g makes no greater contribution to attracting a crowd than any other gold atom in the bar does.)

explain the lion's presence in rather than absence from the given region. But this premise is false. It is (at least roughly) equivalent to: e's presence along with the presence of the rest of the actual leonine configuration of particles, rather than e's absence together with the presence of the rest of the actual leonine configuration of particles, helps to metaphysically explain the lion's presence in rather than absence from the given region. Once again, this premise is false, as suggested by the fact that had e been absent while the rest of the actual leonine configuration remained present, then the lion would still have been present.

However, the analogous premise in the case of natural laws, on the Humean view of lawhood, is true: Given the rest of the actual Humean mosaic (including Fa), Ga's presence in it rather than ∼Ga helps to metaphysically explain the fact that it is true rather than false that all F's are G, and thereby helps to metaphysically explain the fact that it is a law rather than not a law (because false). This premise is (at least roughly) equivalent to: Ga's presence along with the rest of the actual Humean mosaic (including Fa), rather than ∼Ga's holding along with the rest of the actual Humean mosaic (including Fa), helps to metaphysically explain the fact that it is a true rather than false that all F's are G, and thereby helps to metaphysically explain the fact that it is a law rather than a non-law. This premise is correct: had ∼Ga obtained while the rest of the Humean mosaic (including Fa) remained in place, then it would have been false (and hence not a law) that all F's are G.

With this premise, TRANSITIVITY can be used to mount our argument against the Humean view of lawhood. That view entails that given the rest of the actual Humean mosaic (including Fa), Ga's presence in the mosaic rather than ∼Ga helps to metaphysically explain the fact that it is a law that all F's are G rather than not a law. But the fact that it is a law that all F's are G, rather than not a law, helps to scientifically explain why Ga rather than ∼Ga. It follows by TRANSITIVITY that given the rest of the actual Humean mosaic, Ga's presence in the mosaic rather than ∼Ga helps to scientifically explain why Ga rather than ∼Ga, contrary to the view that facts do not routinely help to scientifically explain themselves.

Thus, I have reconciled the putative counterexamples to the original transitivity principle with a contrastive transitivity principle (TRANSITIVITY) that underwrites the explanatory circularity argument against the Humean view of lawhood. The intuitions motivating the putative counterexamples have been respected without undermining the explanatory circularity argument.

3 The diversity of scientific explanations and the irreducibility of macroexplanations

Some objections to the transitivity principle accuse it of failing to respect the diversity of scientific explanations and the distinctness of various levels in nature.

Hicks and van Elswyk (2015, pp. 438–440) recognize that there are many different kinds of explanations (e.g., causal, metaphysical) backed by different kinds of relations (e.g., causal, grounding) between explanans and explanandum. Hicks and van Elswyk use this diversity to argue against transitivity: "When one backing relation connects

D and E, and a distinct backing relation connects E and F there is no reason to think that either backing relation somehow connects D to F" (p. 439).[7]

But transitivity does not require that D be connected to F either by the kind of backing relation connecting D–E or by the kind of backing relation connecting E–F. Rather, transitivity is satisfied if the D–E and E–F backing relations together suffice to back an explanatory connection of D–F. In other words, transitivity is satisfied if there are "hybrid explanations" where D is connected to F neither by the kind of backing relation connecting D–E nor by the kind of backing relation connecting E–F, but by the combination of the two.

There are many such hybrid explanations in science. Consider, for instance, a helium balloon rising in the atmosphere. The initial conditions (e.g., that the balloon is filled with helium), the relevant force laws (such as Archimedes' Principle giving the buoyant force), and Newton's laws of motion scientifically explain (in classical physics) why the balloon begins to rise. This explanation works by describing the relevant causes (the forces on the balloon) and how they act. One way to explain Archimedes' Principle, in turn, is by appealing to energy conservation. But this explanation does not work by describing *causes* (such as air pressure differences). Rather, this explanation works by showing that Archimedes' Principle is a consequence of the *constraint* that energy conservation imposes on the kinds of forces there could be, so that whatever forces may be acting on the balloon, they must conserve energy.[8] By combining these two explanations, we can use energy conservation to help explain the balloon's rising. This explanation is neither a wholly causal explanation nor a wholly non-causal explanation that works by citing constraints. Instead, this explanation involves both of these "basing relations".

Here is another example of a hybrid explanation—involving a causal component and a non-causal, "distinctively mathematical" component. Suppose I bet my friend, a teacher named Jones, that if he picks eight students at random from his class, then he will find that at least two of them were born on the same day of the week. Jones foolishly thinks this result unlikely, takes the $1 bet, loses, and pays me. Why did Jones pay me $1? There is a causal explanation involving Jones's believing that he lost his bet (because he did lose it) and Jones's desiring to honor his commitments. Why, in turn, did Jones lose his bet? There is a non-causal, distinctively mathematical explanation appealing to the pigeonhole principle: with seven days of the week and eight students, it is certain that two students' birthdays fall on the same day of the week. The combination of these two explanations can explain why Jones paid me $1. This explanation is supported not just by causal relations and not just by distinctively mathematical relations, but rather by the two in combination.

Hicks and van Elswyk (2015) offer another objection to transitivity. They see transitivity as treating the microlevel facts D as qualified to explain F merely by virtue of grounding macrolevel facts E that explain F. Hicks and van Elswyk see this view as running roughshod over the distinction "between cases when explanations from one

[7] I have adjusted the letters in this passage to match the letters in my earlier statement of the transitivity principle.

[8] For more on this particular explanation (and on non-causal explanations by conservation laws more generally), see Lange (2011).

science can be reduced to another, and cases in which they cannot" (p. 441). They see this view as treating all macrolevel explanations as "*reducible* to those at a lower level" (p. 441) even though macroeconomic explanations "do not appear to be reducible to the explanations of physics, nor do the explanations given by Darwinian evolutionary theory" (p. 440).

However, I agree with Hicks and van Elswyk that when a macrolevel fact E helps to explain a macrolevel outcome F, then that "macroexplanation" is often irreducible to a "microexplanation" that replaces E with its microlevel ground D. It is important, though, to distinguish the claim that the macroexplanation of F is irreducible to an explanation appealing to D from the claim that D cannot help to explain F. The macroexplanation's irreducibility is its making important explanatory contributions that cannot be made, even in principle, by microexplanations. The macroexplanation's irreducibility therefore does not require that there be no corresponding microexplanation of the same explanandum. The macroexplanation's irreducibility requires merely that any such microexplanation fail to render the macroexplanation superfluous.

Consider one of the examples mentioned by Hicks and van Elswyk: that changes in the frequencies of various traits in biological populations are explained by differences in those traits' fitnesses and, ultimately, by the fact that certain traits are being selected for. Such a macroexplanation supplies explanatory information that cannot be supplied, even in principle, by microdescriptions of the causes of the outcomes of various particular attempts at mating, predation, birth, and so forth. (For instance, the macroexplanation may work partly by specifying what would have happened to the trait frequencies, had the environment been different so that a certain trait supplied less effective camouflage.) Whereas macroexplanations make this distinctive explanatory contribution, microexplanations of the outcomes of various particular attempts at mating, predation, and so forth may nevertheless also combine to explain the overall outcome. Many authors, such as Jackson and Pettit (1992), have argued for the irreducibility of macroexplanations without denying that the corresponding microexplanations are explanatory (even if those microexplanations are, in practice, generally unavailable). Such authors embrace explanatory pluralism.

One of the key explanatory contributions that macroexplanations make (and that cannot be made, even in principle, by the corresponding microexplanations) lies in their unifying microstructurally diverse cases by giving them macroexplanations that all appeal to the same macrolevel properties and laws. For example, explanations that appeal to differential fitness, selection for various traits, and other natural kinds and natural properties in evolutionary theory unify explananda that cannot be unified, even in principle, at the microlevel; the natural kinds and natural properties at the macrolevel correspond to wildly heterogeneous disjunctions of natural kinds and natural properties at the microlevel. This is one of the main lessons famously taught by Fodor (1974). But once again, none of this denies that a particular macrolevel outcome can be explained by a sufficiently full account at the microlevel.[9] It denies only

[9] It is important that the explanandum here is a *particular* macrolevel outcome rather than a macrolevel *regularity*. To insist that a particular macrolevel outcome has a microlevel explanation is compatible with accepting the antireductionist idea (e.g., Kitcher 1984, p. 350) that certain macrolevel "patterns" can be explained only by macrolevel, "structural" features—*not* at the microlevel.

that its macrolevel explanation is reducible to its microlevel explanation; it denies that the microlevel explanation (were we to know it and to grasp it) would supply all of the understanding that the macroexplanation supplies and so would render the macroexplanation superfluous.

Of course, Fodor (1974) maintains that there is no microlevel explanation of the macrolevel outcome. I have suggested that the existence of a microlevel explanation is nevertheless consistent with one of the main lessons of Fodor's paper: that a macrolevel explanation is irreducible because it makes explanatory contributions that could not be made, even in principle, by a microlevel description. Thus, whether transitivity "summarily dismisses the anti-reductionism of Fodor (1974)" (Hicks and van Elswyk 2015, p. 441; cf. Miller 2015, p. 441) depends on what one takes the most important part of Fodor's anti-reductionism to be.[10]

4 The role of laws in and around scientific explanations

Finally, let's examine some objections to the transitivity principle (when coupled with the prohibition on self-explanation) that concern the role of laws and regularities in and around scientific explanations. These objections will help us to refine the prohibition on self-explanation.

Miller (2015, p. 1325) presents a worry arising from "cases in which we want to at least help explain the fact that a particular individual has some feature by citing the fact that all individuals of a relevant sort have this feature: this particular atom has mass m because all hydrogen atoms have mass m, for instance. … Lange's prohibition on self-explanation rules this out, at least given the assumption that the fact that this atom has mass m can *help* ground the fact that all atoms of the relevant sort have mass m, plus his transitivity principle".

Of course, non-Humeans about lawhood may reply that in these cases, the explainer is not the *regularity* that all individuals of the relevant sort have the given feature—a fact that is partly grounded in the explanandum. Rather, the explainer is the fact that it is a *law* that all individuals of the relevant sort have the given feature (or some other metaphysically weighty fact, such as a fact about the essence of being a hydrogen atom). That explainer is not partly grounded in the fact being explained, according to non-Humeans. The objection to the transitivity principle thus begs the question by presupposing a Humean view of lawhood.

However, the objection cannot be dismissed so easily. We can pose the objection in a case that does not involve laws at all and so does not risk begging the question about lawhood. Consider this apparent example of a regularity that is not a matter of natural law explaining one of its instances:

[10] Putnam is often coupled with Fodor as defending anti-reductionism. Putnam says that he does not care whether we call the microdescription a non-explanation or "a terrible explanation" (1975, p. 296). For Putnam, I think, the most important aspect of anti-reductionism is not that a singular occurrence explained by a higher-level science cannot be explained by a lower-level science, but that the higher-level explanation is "autonomous" because it brings great explanatory benefits (such as unification) that cannot be supplied by any lower-level description, no matter how complete (and so the higher-level explanation is far better).

> Suppose I walk into my classroom. My watch reads 11:10 a.m. My class starts at 11:20 a.m. I am trying to get to class early to put some examples on the board. Much to my surprise, the entire class is already present, ready to begin. I ask, 'Why are you here now?' A bit puzzled at what to say, they answer, 'We are always here at 11:20 a.m.' (my watch was slow). (Carroll 1999, p. 79)

Suppose that in Carroll's example, a regularity genuinely explains one of its instances. The regularity is partly grounded in its instances. So the example runs afoul of the transitivity principle coupled with the prohibition on self-explanation. We avoided this problem above by suggesting that it is actually the regularity's lawhood that is doing the explaining. But this solution is inapplicable to Carroll's example, since the regularity there is not a matter of natural law.

I suggest, however, that in Carroll's example, the regularity is not doing the explaining. Rather, the regularity is supplying relevant information about the facts that are doing the explaining. By citing the regularity, the students are telling their teacher that they are "here now" for the same reason as they are always in the classroom at 11:20 (presumably, because of various beliefs and desires that they have, such as the desire to arrive on time for class). Carroll agrees:

> On the one hand, my students' response to me … appear[s] to be [a] good respons[e] to the questio[n]that [was] asked. On the other hand, it is not clear that the success of [this] respons[e] is a consequence of [its] having supplied the requested explanation. In fact, it's pretty clear it's not. The fact that my students are always there at 11:20 may show *that* they were there at 11:20 that day, but it doesn't show *why* they were there then. (Carroll 1999, p. 79)

The students' response was a good response to their teacher's why question, since it supplied contextually relevant information about the reasons why the students were present at that time. Nevertheless, the students' response fails to give any of those reasons.

I have just appealed to the distinction between facts that are doing the explaining and facts that are supplying relevant information about the facts that are doing the explaining. This distinction is familiar from many other examples in the philosophical literature on scientific explanation. For instance, Lewis (1986, p. 220) considers the example, "Why was the CIA man there when His Excellency dropped dead? Just coincidence…" The response, "Just coincidence," is a good response to the question because it supplies conversationally relevant information about the causal histories of the CIA agent's presence and His Excellency's death (namely, that their causal histories have nothing important in common). But "Just coincidence" supplies conversationally relevant information about the explainers without identifying any of them (e.g., any cause of the CIA agent's presence).

The prohibition on self-explanation should be understood as a prohibition on a fact's helping to explain itself. It should not be understood as a prohibition on a fact's helping to provide relevant information about the facts doing the explaining. The fact being explained can be a ground of a fact that helps to provide relevant information about the facts doing the explaining without violating the transitivity principle coupled with the prohibition on self-explanation.

By appealing to the distinction between the facts doing the explaining and facts supplying relevant information about the facts doing the explaining, we have just disarmed an objection to the explanatory circularity argument against Humean views of law. But we have thereby made room for a different objection to that argument. The explanatory circularity argument presupposed that the fact that it is a law that all F's are G is a fact that helps to explain why Ga obtains. But what if this fact appears in good responses to why questions not because it is itself an explainer, but rather because it supplies relevant information about the facts doing the explaining? If the fact that it is a law that all F's are G is not an explainer, then Ga's helping to ground this fact would appear to pose no threat of self-explanation, even if the transitivity principle holds.

How could the fact that it is a law that all F's are G be a fact about the explainers rather than an explainer itself? According to Skow (2016, p. 140), sometimes when we respond properly to "Why Ga?" with "Because Fa and it is a law that all F's are G", we are giving none of the reasons why Ga. Rather, we are conveying that a is G for the same sort of reason that every other F is G, though we are not giving any of those reasons. This may indeed be true in some cases; for instance, to use one of Skow's examples, if we say that a is black because a is a raven and because it is a law that all ravens are black, we may be conveying the fact that for each raven, there is a set of reasons why it is black and the same kind of reason figures in each of these sets. (The kind of reason that interests the interlocutors might be genetic traits common to all ravens or instead events in their common evolutionary history.) Likewise, consider the question "Why is it that, when we threw the powder into a Bunsen burner flame, the powder burned yellow (rather than with a flame of some other color)?" According to Skow, a good reply to this question is that the powder is a sodium salt and that it is a law of nature that all sodium salts burn yellow. Skow says that this reply does not identify any of the causes of the powder's burning yellow and therefore does not identify any of the reasons why the powder burned yellow. But it conveys that the reasons why it burned yellow have "the property that, whenever a sodium salt is burning yellow, a reason of that kind is a reason why it is burning yellow" (Skow 2016, p. 140).

In these two examples, membership in a given natural kind F (ravens, sodium salts) is not *causing* the characteristic trait being explained (having black plumage, burning yellow). Rather, the possession of some *other* property (the raven's genetic trait, the sodium salt's having electrons in certain energy levels) is doing the causing and thus explains why the powder burns yellow. These other properties are left unspecified by a good response to "Why Ga?" that cites the law that all F's are G. In citing the law, the response informs its audience that these unspecified properties are the same for all members of the natural kind F. Other laws link those properties to the characteristic trait being explained. For instance, there are laws taking reasons why a given sodium salt burns yellow (such as its having electrons in certain energy levels) and linking those reasons to their effects (and, ultimately, to the salt's burning yellow). Suppose we respond to "Why does the powder burn yellow?" by citing those laws. That response (unlike the appeal to the powder's being a sodium salt and the law that all sodium salts burn yellow) identifies some of the reasons why the powder burns yellow and cites laws connecting those reasons to their effects. By forging those

connections, these laws themselves seem to be among the reasons why the salt burns yellow.

Nevertheless, even for such a law, there remains a way to resist regarding the fact that it is a law that all F's are G as among Ga's explainers. Marshall (2015, p. 3160) says that Humeans can propose that "while the fact that All *F*s are *G*s is a law cannot partly explain [why] *a* is *G*, it can partly explain why All *F*s are *G*s partly explains [why] *a* is *G*." Similarly, Skow (2016, p. 81) maintains that laws of nature are not first-order reasons why Ga, but rather are higher-level reasons why: reasons why some C is a cause of Ga. On this proposal, once again, the fact that p is a law figures in good responses to why questions because it supplies relevant information about the facts doing the explaining (which typically specify causes of the event being explained). But although the fact that p is a law is a reason why some other fact is a reason why Ga, it is not itself a reason why Ga. On this proposal, then, suppose we try to make the explanatory circularity argument by applying the transitivity principle (that if D helps to ground E and E helps to scientifically explain F, then D helps to scientifically explain F). The fact that all F's are G is a law cannot serve as E (with Ga serving as both D and F). Therefore, the explanatorily circularity argument against Humean views of law fails to get off the ground.

On this proposal, the fact that it is a law that p (all F's are G) does not help to explain why Ga. But there are reasons why Ga, such as Fa. However, suppose we shift the explanandum from Ga to the fact that all F's are G. It seems that in scientific practice, an explanation of the fact that all F's are G can be that it is a fundamental law that all F's are G. But if the fact that p is a fundamental law does not help to explain why Ga, it seems that by the same token, the fact that p is a fundamental law does not help to explain why p. In that event (by contrast to the case where Ga is the explanandum and Fa is the explanans) there is nothing to serve as the explanans! This seems contrary to scientific practice. For example, every attempt to build a machine that produces more energy than it consumes has failed. Why has every attempt failed? Scientists do not reply that this fact has no explanation or that every attempt has failed for lack of scientific ingenuity. Rather, scientists reply to this why question by citing the fact that a law of nature (the conservation of energy) makes such a machine impossible.

Advocates of the proposal that p is not explained by p's lawhood could attempt to reconcile it with scientific practice. They could suggest that when we reply to "Why p?" with "Because p is a fundamental law", we are not purporting to give a reason why p. Instead, we are informing our interlocutor that p has no explanation, where p's status as a fundamental law explains why p has no explanation. On the other hand, if advocates of the above proposal instead grant that in such a case, p's lawhood helps to explain why p, then this admission is enough to drive the explanatory circularity argument against the Humean view of law: that p is a law scientifically explains why p, and (on the Humean view) p helps to ground p's lawhood, so (by transitivity) p helps to scientifically explain p, violating the prohibition on self-explanation.[11]

[11] Indeed, casting the explanandum as the fact that all F's are G, rather than the fact that Ga, may be the cleanest way to put the explanatory circularity argument in the first place. By putting the argument in this

Fortunately, to press the explanatory circularity argument against the Humean view of law, we do not have to resolve the question with which I have just been wrestling; that is, we do not have to figure out whether the lawhood of All F's are G helps to explain why it is that Ga or merely helps to explain why it is that Fa helps to explain why it is that Ga. The prohibition on self-explanation should be interpreted not only as prohibiting a fact q from helping to explain itself, but also as prohibiting q from helping to explain why (if q obtains) some other fact helps to explain q. Both of these are too circular to qualify as explanations.[12]

There is some precedent for this broad view of what would constitute unhealthy circularity—although in connection with inferences rather than explanations. Salmon (1967) evaluated the success of Black's response to the Humean problem of justifying induction. Black's response, roughly speaking, purports to offer a means of avoiding any unhealthy circularity in justifying belief in the Principle of the Uniformity of Nature (PUN): by giving an argument for PUN that uses PUN not as a premise, but rather as a rule of inference. Purportedly, PUN would then not be helping to justify itself. Rather, it would be a rule of inference by which some other premise can justify PUN. (In the same way, unhealthy circularity is supposed to be avoided by having the fact that all F's are G is a law not help to explain why Ga, but rather help to explain why some other fact can help to explain why Ga.)

In particular, Salmon (1967, p. 13) asks us to consider the following rule of inference:

> R_2: "To argue from *Most instances of A's examined in a wide variety of conditions have been B* to (probably) *the next A to be encountered will be B*".

Salmon then asks us to consider the following argument employing R_2 as its rule of inference:

> In most instances of the use of R_2 in arguments with true premises examined in a wide variety of conditions, R_2 has been successful [i.e., the argument has a true conclusion].
>
> Hence (probably):
>
> In the next instance to be encountered of the use of R_2 in an argument with true premises, R_2 will be successful.

Salmon says that despite this argument's valiant effort to avoid circularity, this argument is nevertheless circular:

> One way in which an argument can be circular is by adopting as a premise the very conclusion that is to be proved…Another way in which an argument can be circular is by exhibiting a form whose validity is asserted by the very conclusion that is to be proved … Neither type of circular argument establishes

Footnote 11 Continued
way, we side-step questions about what, in addition to Fa and the fact that all F's are G is a law, is needed to explain Ga, since these are insufficient (as illustrated by the well-known difficulties for the D-N model of scientific explanation).

12 Compare (Lange 2013, p. 258, n. 3).

> its conclusion in any interesting fashion, for in each case the conclusiveness of the argument depends upon the assumption of the conclusion of that argument. (Salmon 1967, p. 15)

A circular argument of either kind is "completely question begging", Salmon says.[13] I agree.

As we just saw, Salmon maintains that there are two ways for an inference to be otiose because of circularity:

(i) by using p as a premise in an inferential justification of p, or
(ii) by taking the rule of inference the reliability of which p expresses and using it to underwrite the inference in which a given q serves as a premise in an inferential justification of p.

I suggest that likewise, there are two ways for a purported scientific explanation to fail because of circularity:

(i') by using p in the explanans in an explanation of p, or
(ii') by using p to help explain why (if p obtains) a given q can serve as part of the explanans in an explanation of p.

Both (i') and (ii') should be interpreted as violating the prohibition on self-explanation.

From Sects. 2–4, I conclude that various objections offered by Hicks and van Elswyk (2015), Marshall (2015), and Miller (2015) to the explanatory circularity argument require that this argument be refined in various respects. But those refinements all have independent motivations and do not ultimately blunt the argument's force against Loewer's (2012) defense of the Humean view of lawhood.

References

Achinstein, P. (1963). Circularity and induction. *Analysis*, *23*, 123–127.
Armstrong, D. (1983). *What is a law of nature*. Cambridge: Cambridge University Press.
Carroll, J. (1999). The two dams and that damned paresis. *British Journal for the Philosophy of Science*, *50*, 65–81.
Fine, K. (2012). Guide to ground. In F. Correia & B. Schneider (Eds.), *Metaphysical grounding: Understanding the structure of reality* (pp. 37–80). New York: Cambridge University Press.
Fodor, J. (1974). Special sciences (or: The disunity of science as a working hypothesis). *Synthese*, *28*, 97–115.
Garfinkel, A. (1981). *Forms of explanation*. New Haven: Yale University Press.
Hicks, M. T., & van Elswyk, P. (2015). Humean laws and circular explanation. *Philosophical Studies*, *172*, 433–443.
Jackson, F., & Pettit, P. (1992). In defence of explanatory ecumenism. *Economics and Philosophy*, *8*, 1–21.
Kitcher, P. (1984). 1953 and All that. A tale of two sciences. *Philosophical Review*, *93*, 335–373.
Lange, M. (2011). Conservation Laws in scientific explanations: Constraints or coincidences? *Philosophy of Science*, *78*, 333–352.
Lange, M. (2013). Grounding, scientific explanation, and humean laws. *Philosophical Studies*, *164*, 255–261.
Lewis, D. (1986). *Philosophical papers* (Vol. II). New York: Oxford University Press.
Lipton, P. (2004). *Inference to the best explanation* (2nd ed.). London and New York: Routledge.

[13] Achinstein (1963, p. 126) comes to the same conclusion.

Loewer, B. (2012). Two accounts of Laws and time. *Philosophical Studies*, *160*, 115–137.
Marshall, D. (2015). Humean laws and explanation. *Philosophical Studies*, *172*, 3145–3165.
Maudlin, T. (2007). *The metaphysics within physics*. New York: Oxford University Press.
Miller, E. (2015). Humean scientific explanation. *Philosophical Studies*, *172*, 1311–1332.
Mills, S. K., & Beatty, J. (1979). The propensity interpretation of fitness. *Philosophy of Science*, *46*, 263–286.
Putnam, H. (1975). Philosophy and our mental life. *Putnam, mind, language and reality* (pp. 291–303). Cambridge: Cambridge University Press.
Rosen, G. (2010). Metaphysical dependence: Grounding and reduction. In B. Hale & A. Hoffmann (Eds.), *Modality: Metaphysics, logic, and epistemology* (pp. 109–135). New York: Oxford University Press.
Salmon, W. (1967). *The foundations of scientific inference*. Pittsburgh: University of Pittsburgh Press.
Schaffer, J. (2005). Contrastive causation. *Philosophical Review*, *114*, 297–328.
Skow, B. (2016). *Reasons why*. Oxford: Oxford University Press.
van Fraassen, B. (1980). *The scientific image*. Oxford: Clarendon.