



Mathematics and Explanatory Generality: Nothing but Cognitive Salience

Robert Knowles¹  · Juha Saatsi¹

Received: 25 November 2018 / Accepted: 28 June 2019
© The Author(s) 2019

Abstract

We demonstrate how real progress can be made in the debate surrounding the enhanced indispensability argument. Drawing on a counterfactual theory of explanation, well-motivated independently of the debate, we provide a novel analysis of ‘explanatory generality’ and how mathematics is involved in its procurement. On our analysis, mathematics’ sole explanatory contribution to the procurement of explanatory generality is to make counterfactual information about physical dependencies easier to grasp and reason with for creatures like us. This gives precise content to key intuitions traded in the debate, regarding mathematics’ procurement of explanatory generality, and adjudicates unambiguously in favour of the nominalist, at least as far as explanatory generality is concerned.

1 Introduction

The debate surrounding the enhanced indispensability argument for mathematical platonism (EIA) has reached an impasse, descending into intuition-trading regarding purported examples of mathematical explanations of physical phenomena. Progress demands an independently motivated understanding of these explanations that favours either nominalism or platonism. Indeed, ‘[i]f there is a point of agreement in this debate, it is that we could do with a better understanding of mathematical explanation’ (Colyvan 2013: 1044). In this paper, we make significant progress towards such an understanding. Drawing on a counterfactual theory of explanation, well-motivated independently of the debate, we provide a novel analysis of ‘explanatory generality’, highlighted by advocates of EIA as a key virtue of certain mathematical explanations, and how mathematics is involved in its procurement. On our analysis, mathematics’

✉ Robert Knowles
r.knowles@leeds.ac.uk

Juha Saatsi
j.t.saatsi@leeds.ac.uk

¹ School of Philosophy, Religion, and History of Science, University of Leeds, Leeds, UK

sole explanatory contribution to the procurement of explanatory generality is to make counterfactual information about physical dependencies *cognitively salient*: easier to grasp and reason with for creatures like us. This gives precise content to the conflicting intuitions about explanatory generality traded in the debate, and adjudicates unambiguously in favour of the nominalist. Since it is independently motivated, this verdict is not question-begging, demonstrating our methodological conclusion: working with an independently well-motivated theory of explanation can push past conflicting intuitions in the EIA debate.

To be clear, we do not aim to vindicate nominalism once and for all. Our aims are more modest than that. First, we aim to provide non-question-begging reasons for thinking that the explanatory generality of certain mathematical explanations is compatible with nominalism. Second, in achieving this, we aim demonstrate the methodological conclusion mentioned above. This only addresses one virtue claimed for certain mathematical explanations, though one that has thus far been central to the platonist's case (see below). Although we regard the counterfactual theory in question capable of capturing a very broad range of explanations, it is not our aim here to argue for this account's universal superiority, nor do we commit to explanatory monism. Nevertheless, this is real progress. In light of it, if it is claimed that a purported mathematical explanation or a virtue thereof escapes our analysis, this claim had better not be supported by intuition alone. Our opponent must provide reasons, similarly independent of the EIA debate, to suppose our analysis fails to capture something of genuine explanatory worth.

Before presenting our analysis, we provide some background on the EIA debate (below), a presentation of our chosen counterfactual theory of explanation (Sect. 2), and a toy example of the kind of mathematical explanation our analysis is supposed to capture (Sect. 3).

The impasse regarding EIA stands as follows. According to EIA, scientific realists should be mathematical platonists, in light of examples of scientific explanations which seem to turn on mathematical facts. Even if alternative, nominalistic explanations can be offered, arguably such alternatives are worse explanations. Thus, a realist who infers to the *best* explanation cannot but accept commitment to whatever the mathematical facts involve. So runs the most prominent naturalistic argument for platonism (e.g. Baker 2005; Colyvan 2002, 2013; Lyon 2012). In response, nominalists can either deny that the mathematically presented explanations are better, or deny that mathematics' contribution to them is ontologically committing. There is a near-consensus that the first horn is untenable. The second horn has been popular, but here the impasse looms. While platonists take mathematics' explanatory indispensability to evidence the existence of explanatory mathematical features of reality, nominalists take mathematics to merely increase our expressive capacity, allowing us to represent the physical features of reality that are doing the real explanatory work.

On the one hand, the distinction between 'really explanatory' and 'merely expressive' can only be drawn fairly on some principled, non-question-begging grounds, and nominalists' appeal to this distinction has hitherto not convinced platonists.¹ On the

¹ For example, Saatsi (2016) draws an interesting distinction between 'thin' and 'thick' explanatory roles, arguing that a nominalist can make sense of mathematics' explanatory indispensability by maintaining

other hand, the indispensability of mathematics for *providing* an explanation is not enough in and of itself to convince the nominalist that its explanatoriness springs from correctly representing mathematical features of reality. Thus, the debate has reached a serious impasse (e.g. Baker 2017: 2; Knowles and Liggins 2015: 3403–3407), and the result is predictable: intuition trading, subtle dialectical manoeuvring, and charges of question-begging.

The claim that explanatory generality is what makes certain mathematical explanations better than any nominalistic alternative is popular among platonists. Indeed, it has been widely expressed for more than 10 years. For example, Colyvan (2002) defends EIA against Melia (2000) by noting that mathematics is indispensable for a ‘unified approach’ to presenting and solving disparate equations, and hence ‘genuinely explanatory’, since ‘unification is linked to explanatory power’ (p. 72). Alan Baker and Colyvan (2011) defend EIA by noting that any nominalised explanation of cicada periods is ‘both less general and less robust’ (p. 331). Colyvan (2013) defends EIA against Yablo (2013) by arguing that Yablo’s approach doesn’t do justice to the unifying power of mathematics (p. 1042). Plebani (2016) objects to David Liggins’ (2016) response to EIA because Liggins’ approach runs our scientific explanations ‘at the wrong level of generality’ (p. 553). Most recently and explicitly, Baker (2017) has articulated the idea that explanatory generality supports platonism. All in all, the notion of ‘explanatory generality’—together with cognate notions of ‘unification’ and ‘robustness’—has thus far been central to the platonist case.

The prominence of explanatory generality reveals two lacunas in the debate. First, neither side has provided a satisfactory account of explanatory generality, detailing the extent to which it is explanatory, and how exactly mathematics facilitates it.² Second, nominalists have hitherto not explicitly responded to the appeal to explanatory generality. The challenge is to explain how mathematics can make explanations better by making them more general, if not by representing more general, mathematical features of reality?

We address both lacunas by appeal to a popular, well-motivated counterfactual theory of explanation. The gist of the theory is that explanations capture systematic patterns of counterfactual dependence—causal or non-causal—that tell us how the explanandum depends on the explanans. Explanations thus provide ‘what-if-things-had-been-different’ information, indicating how the explanandum would have been different, had the explanans been different in particular ways (e.g. Woodward 2003). We will show that, from the perspective of this theory, the only kind of generality that counts as an explanatory virtue, in a sense that matters to EIA, is a kind of counterfactual generality (‘scope generality’). Although there is another prominent kind of generality (‘topic generality’), we will argue that this is not in its own right explanatory at all; it can only improve an explanation indirectly, by increasing scope generality. Furthermore, and most importantly, we will argue that increasing scope

Footnote 1 continued

that mathematics only plays a ‘thin’ role of ‘allowing us to grasp, or (re)present, whatever plays a ‘thick’ explanatory role’ (p. 12). Notwithstanding the conceptual room for Saatsi’s distinction, platonists can rightly query why the explanatory superiority of mathematical explanations is best analysed in these terms.

² Lyon (2012) has framed the ‘robustness’ of mathematical explanations in terms of the ‘program explanation’ model of Jackson and Pettit (1990); but see Saatsi (2012, 2016) for criticism.

generality can improve an explanation only up to a point that can be reached by nominalistic means (Sect. 4). Nevertheless, distinctively mathematical explanations truly can be better than any nominalistic alternatives, but we will show (Sect. 5) that this is entirely due to mathematics' *presentational* contribution. While sufficient generality can be achieved without mathematics, this cannot be done without sacrificing our ability to use the explanation, and cognitively grasp the pattern of counterfactual dependence it represents. In this sense, mathematics allows us to achieve explanatory generality *without sacrificing cognitive salience*.

2 A Counterfactual Theory of Explanation

In this section, we motivate the appeal to our chosen theory of explanation, outline it, and survey its implications for analysing comparative explanatory virtues. For convenience, we will refer to it as *the counterfactual theory*, though we recognise that there are other counterfactual theories of explanation.

In appealing to the counterfactual theory, we operate in a naturalistic (Quinean) spirit, appropriate for the EIA debate. Adopting an account that not only captures well prominent examples of mathematical explanation, but is also well-motivated by recent advances in the philosophy of explanation, we approach the debate in the light of all available evidence, including that which independently supports this account of explanation. Support for the counterfactual account involves its ability to make sense of various kinds of explanations, and judgements of comparative degrees of explanatory power, as we will discuss. It also enjoys naturalistic support from, e.g., cognitive psychology, which we will not discuss here (see e.g. Buchsbaum et al. 2012). Again, this is not to say that the counterfactual theory we use in this paper is the only game in town (cf. Sect. 1). This is just to say that the counterfactual theory enjoys a lot of support independently of the EIA debate. To the extent that the theory furthermore nicely captures what is good about mathematical explanations central to the EIA debate, its support is transferred to whichever intuitions expressed in the EIA debate it agrees with.

The counterfactual account does not prejudge the EIA-debate. It is a broadly *ontic* account, taking explanatory power to at least partly derive from representing worldly things that bear an objective, explanatory relation to the explanandum, so it does not jettison the issue of explanations' realist commitments (cf. Saatsi 2016). Further, it does not entail that successful explanations wear their realist commitments on their sleeves (Bokulich 2016; Woodward 2003; Potochnik 2018). Finally, the account does not render mathematical explanations causal. Although initially developed as an account of causal explanation, the counterfactual account has been extended to various non-causal explanations, including certain mathematical explanations (Reutlinger 2016, 2018; Saatsi 2018; Baron et al. 2017; Jansson and Saatsi 2017; French and Saatsi 2018). There is nothing in the counterfactual account itself that rules out mathematical objects or properties bearing objective, explanatory relations to physical explananda.

We can now review the counterfactual account. In short, explaining involves describing systematic patterns of counterfactual dependence. Explanatory counterfactuals are appropriately *directed* and *change-relating*, capturing objective, mind-

independent modal connections that show how the value of the explanandum variable depends on the value of the relevant explanans variable(s). These variables stand for suitably conceptualised and individuated worldly features. Explanatory counterfactuals provide ‘what-if-things-had-been-different’ information, indicating how the explanandum would have been different, had the explanans been different in particular ways. Explanation-supporting relations—nomological, causal, or mathematical—between the variables can provide this kind of modal information.

An account of explanation built on this idea allows us to capture shared intuitions regarding the comparative virtues of different explanations in terms of the counterfactual information they provide. If explaining is a matter of providing information that answers *what-if* questions, it is natural to regard as more powerful those explanations that provide *more* such answers (with respect to a given explanandum). This simple idea has rich and non-trivial consequences regarding the various ways in which explanations can be better or worse. Detailed analyses of explanatory power in this spirit have been provided (see Hitchcock and Woodward 2003; Ylikoski and Kuorikoski 2010). The latter authors identify five different aspects of explanatory power: non-sensitivity, precision, factual accuracy, degree of integration, and cognitive salience. Three are particularly relevant to our argument:

NON-SENSITIVITY The range of values that the explanans variables can take without breaking the explanatory relationship. For instance, an explanation of tides in terms of Newton’s gravitational law has a considerable degree of non-sensitivity with respect to the specific masses and locations of the sun and the moon: the explanation correctly answers a considerable range of change-relating *what-if* questions for non-actual values of these variables.

DEGREE OF INTEGRATION The connectedness of an explanation to other theoretical frameworks. From the counterfactual perspective, such integration is an explanatory virtue when it enlarges the range of *what-if* questions answerable with respect to particular explananda, or makes such questions easier to answer. One way theoretical integration can achieve this is by equipping explainers with new inferential resources (Ylikoski and Kuorikoski 2010). For example, the integration of the gravitational-dynamic theory of tides to the more general mathematical theory of Fourier analysis (by William Thomson) allowed many new *what-if* questions to be asked and answered about local tidal phenomena, increasing explanatory understanding (Cartwright 2000).

COGNITIVE SALIENCE ‘[T]he ease with which the reasoning behind the explanation can be followed, how easily the implications of the explanation can be seen and how easy it is to evaluate the scope of the explanation and identify possible defeaters or caveats’ (Ylikoski and Kuorikoski 2010: 215). Actual explainers are humans with limited cognitive capacities, and these limitations partly determine which explanations have more or less explanatory power by virtue of differing in their capacity to *enable* explainers (with particular training, background knowledge, etc.) to draw counterfactual inferences for different values of the explanans variables.

As we shall see, these resources furnished by the counterfactual account are highly apt for analysing mathematics’ contribution to certain mathematical explanations. We will first illustrate this with a toy example, providing an initial analysis of ‘explana-

tory generality' (Sect. 3). This analysis will be deepened in the subsequent sections. First, by showing that only one kind of generality ('scope generality') counts as an explanatory virtue, and that it does so only up to a point (Sect. 4); and then by showing that the relevant mathematical explanations' cognitive salience is what makes them superior. Thus, although we can provide a nominalistic alternative that achieves the kind of generality that matters for explanations, they are nevertheless less explanatory (Sect. 5).

3 A Toy Example

What kinds of mathematical explanations are appropriate for ascertaining mathematics' contribution to explanatory generality? In this paper, we will focus on explanations that (i) have nominalistic alternatives provided in the literature, and yet (ii) have something 'distinctively' mathematical about them (cf. Lange 2013). These conditions are met by many familiar examples from the EIA literature, such as the number-theoretic explanation of cicada periods, and the graph-theoretic explanation of the non-traversibility of Königsberg's bridges (Baker 2005; Pincock 2015). In what follows, we call explanations that meet conditions (i) and (ii) *distinctively mathematical*. For the purposes of this paper, our analysis should be understood as targeting distinctively mathematical explanations exclusively.

There two important reasons for limiting our focus in this way. First, focusing on distinctively mathematical explanations allows for an instructive contrast between mathematics-laden and mathematics-free formulations that will be invaluable for communicating our analysis. This is particularly important, since the challenge from explanatory generality turns on mathematics' increasing explanatoriness in contrast to nominalistic alternatives. Second, in focusing on distinctively mathematical explanations, we set aside explanations for which we have no nominalistic alternative. This is appropriate since the claim that mathematics is doing any explanatory work in such explanations is a moot point. Given our aims, this is dialectically advantageous. Since we take their nominalistic alternatives to be comparatively less explanatory, we concede that mathematics is explanatorily indispensable to distinctively mathematical explanations. We thus tackle the challenge from explanatory generality head-on, by arguing that the explanatory virtue that mathematics contributes to these explanations, cognitive salience, is not ontologically significant.

We recognise that limiting our scope in this way falls short of providing a full rebuttal of EIA. Nevertheless, providing non-question-begging grounds for thinking that mathematics' contribution to the explanatory generality of distinctively mathematical explanations is compatible with nominalism marks real progress in the debate, and thus demonstrates the efficacy of our methodology. Appealing to the indispensability of mathematics to e.g. phase-space explanations by itself does nothing to diminish this. Once its efficacy has been demonstrated, our methodology can be applied to make further progress in relation to more difficult examples. Although we are optimistic about extending our analysis to accommodate a broader range of explanations, including phase-space explanations (see Saatsi 2017 for a promising start), there is unfortunately not space to attempt this here.

The following toy example illustrates the key features of distinctively mathematical explanations:

A philosopher tries to divide twenty-three distinct ideas evenly among three sections, but keeps failing. Why? Because:

(A) There is no integer n such that $23/3 = n$.

This is a distinctively mathematical explanation: the explanandum is not due to contingent laws of nature (causal or otherwise), but holds with a stronger degree of (mathematical or logical) necessity (Lange 2013).

Applying the counterfactual account, we should take (A) to explain why the philosopher keeps failing because it tells us what the philosopher's failure counterfactually depends on: whether the number of ideas is divisible by three. We can now distinguish two striking kinds of generality exhibited by distinctively mathematical explanations (cf. Jansson and Saatsi 2017; Baker 2017). First, the explanation clearly has nothing to do with the nature of the things divided, apart from the stipulation that they are distinct. In this way, the explanation is extremely **topic general**. Within the counterfactual account, this is naturally understood in terms of DEGREE OF INTEGRATION. Because the explanation is integrated into the topic-neutral framework of arithmetic, it is easy to generate further explanations with radically different subject matters: (A) equally explains why mother couldn't share twenty-three strawberries between her three children, and why our parents never managed to spend all their pocket money (twenty-three pence) on three-pence gobstoppers. Despite the variety of subject matters, when described at the appropriate level of generality, it is clear that these scenarios are structurally similar: they concern attempts to divide twenty-three distinct individuals into three equinumerous collections. It is also clear that the attempts fail for the same reason. To achieve the desired level of topic generality, we must formulate the explanatory generalisation so that it just concerns individuals, instead of strawberries, gobstoppers, or any other kind of objects in particular. Arithmetic achieves this admirably.

Second, the mathematical explanation is naturally equipped to provide answers to questions concerning other actual or hypothetical situations in which the philosopher has a different number of distinct ideas to convey. In this sense, the explanation is extremely **scope general**. Within the counterfactual account, this corresponds directly to NON-SENSITIVITY (viz. the range of explanans variable values over which the explanatory relationship truthfully holds).

Consider an explanation that only captures the dependence of the value of the explanandum variable (successful-division-into-three / unsuccessful-division-into-three) on the value of the explanans variable (number of individuals) between twenty-two and twenty-four individuals. Such an explanation gets right that it is impossible to divide twenty-three things evenly into three, and that the philosopher would succeed if there were twenty-four things to begin with. But these tidbits notwithstanding, it fails to capture the broader explanatory regularity at stake. If one philosopher tries to divide twenty-three ideas into three, and another tries to divide three-hundred-and-four, they both fail for the same reason: neither set is evenly divisible by three. We do not have one explanation for one case, and another for the other; we have one expla-

nation, supported by one explanatory regularity, and two different initial conditions. This conception of scope generality naturally rides on the back of the idea—at the heart of the counterfactual account—that explanations work by presenting an explanatory relationship that connects the explanandum to the explanans, showing how the former depends on the latter by virtue of providing true *what-if* information (regarding the state of the explanandum) for at least some non-actual values of the explanans variable. The larger the range of non-actual explanans variable values that are truthfully captured by the explanatory relationship, the less sensitive the application of the explanation is to the actual values of these variables—that is, the more scope-general it is.

As noted, platonists have variously appealed to explanatory generality in support of EIA. For such an appeal to be decisive, we need to know the extent to which the kind of generality at stake is explanatory, and how exactly mathematics contributes to it. Such questions are best answered from the perspective of a well-motivated theory of explanation, which neither platonists nor nominalists have offered. With our theory in hand, we will now answer these questions in a principled manner.

4 The Explanatory Limits of Generality

On the counterfactual account, explaining is a matter of delineating a pattern of counterfactual dependence by providing information that correctly answers relevant *what-if* questions. Explaining—providing explanatory information—is a human activity, the goal of which is the provision of explanatory *understanding*, which is naturally construed as an agent's *ability* to correctly answer *what-if* questions (e.g. Ylikoski and Kuorikoski 2010). This introduces epistemic and pragmatic aspects to the analysis of 'best explanation', since explaining in the best possible way involves weighing up considerations which sometimes pull in different directions. With this in mind, we will now consider both scope generality and topic generality in detail, to see the extent to which these different kinds of generality can be said to make explanations better *qua* explanations. We will first argue that although increasing scope generality can improve an explanation, it does so only up to a point. We will then argue that topic generality in its own right is not explanatory at all; it can only improve an explanation by increasing scope generality.

Scope Generality Referring to our toy example, (A) allows for a *maximally* scope-general explanation: there is no limit to the number of individuals it can take as the value for the explanans variable. Similarly, the number-theoretic explanation of cicada periods places no limit on possible alternative durations of these periods, and the graph-theoretic explanation of non-traversibility of Königsberg's bridges places no limit on possible alternative bridge-configurations. While we can form nominalistic counterparts by just listing what these explanations say about some finite range of explanans values (§5), one might think that the maximisation of scope generality is in and of itself a reason to prefer the mathematical version. There are good naturalistic grounds to resist this thought, however.

On the counterfactual account, explaining is a matter of providing information that correctly answers *what-if* questions pertinent to the explanandum at stake. Which

questions are pertinent? We will now argue that we do not judge one explanation to be better than another on the basis of its answering what-if questions *without limit*. As we illustrate below, some explanations underwrite counterfactual scenarios that completely transcend our science-based grasp of the possibilities in question. We should not take seriously such ‘extreme’ counterfactuals. No scientist would or should consider these what-if questions pertinent to the explanandum in question, in judging one explanation to be better than another. Doing so would demonstrate an alarming, almost fetishistic confidence in the explanatory, mathematised theory to reliably tell us about counterfactual situations far-removed from the theoretical context in which the theory was developed. Moreover, evaluating such counterfactuals involves super-empirical issues of modal semantics that should not be decided so easily by a scientific theory developed for completely different purposes. In short, such extreme counterfactuals are irrelevant to the explanatory practice of science, and thus make no naturalistically recognisable contribution to our explanatory understanding. Rather, explanatory content is exhausted by counterfactuals that scientists would and should take seriously, given the empirical evidence at stake and the theoretical context in question.

For illustration, consider the much-discussed case of North-American periodical cicadas, which lie dormant for 13 or 17 years (depending on the subspecies). Why the prime numbers? Because an organism which lies dormant for a prime number of years minimises the frequency of overlaps between their emergence and the emergence of nearby periodical predators (assuming these nearby predators have life-cycles of between 2 and 12 years).³ Here, the operative mathematical generalisation can be presented as follows:

- (B)** If p is any natural number, then p is coprime with every $q < p$ (and hence maximises its LCM with every $q < p$) iff p is prime.

In the counterfactual framework, this explanation turns on grasping how the fitness-maximising cicada life-cycles—construed as a variable that can take different values—depends on other biologically relevant variables, such as predator species’ life-cycles. Roughly speaking, the explanatory counterfactuals capture the dependence of long-run evolutionary outcomes, regarding specific life-cycle periods, on the existence of predator species with life-cycles of nearby periods, as well as ‘ecological constraints’ that appropriately limit the range of viable possibilities.

(B) maximises the scope generality of the explanation: there is no upper limit to the number it can take as the value for the explanans variable p , representing the cicadas’ life-cycle period in years. However, most natural numbers—most of the possible p values—correspond to extreme counterfactuals that we cannot evaluate in anything but an abstract mathematical sense, at least not without modal presuppositions that are not part of the naturalistic context in which (B) is judged to provide the best explanation. For example, consider all p values according to which the cicadas have life-cycle periods greater than the age of the universe. We maintain that we cannot take the answers that (B) entails regarding such extreme counterfactuals to correctly answer what-if questions concerning *cicadas*, since we cannot evaluate such counterfactuals

³ This follows Baker (2005); see Wakil and Justus (2017) for worries, which are irrelevant for our argument and analysis.

without taking a stand on issues regarding modality and modal semantics that go well beyond the naturalistic commitments of evolutionary biology. The limitless scope generality offered by (B) does not seem to underwrite or account for the explanatory superiority of this distinctively mathematical explanation.

In the same spirit, consider our toy example again. We can use (A) to answer questions such as: ‘What if things were different, such that the philosopher attempted to divide $3 \times 10^{82} + 1$ ideas between three sections?’, where 3×10^{82} is of the order of the number of atoms in the observable universe. The answer is obvious, as far as arithmetic is concerned. But it is not at all clear why one should expect this answer to be part of explaining why the philosopher cannot divide her ideas equally between three sections. Although mathematically clear-cut, reflecting on the counterfactual in question reveals it to be fantastically far-removed from the explanandum and its theoretical context. It is clearly psychologically and practically absurd that a philosopher could have that many ideas to express in a paper. Furthermore, for all we know, it may be nomologically impossible for any organism, in any possible universe, to come to have that many distinct ideas (assuming any kind of physicalism about the mind). Metaphysical and cosmological speculation aside, we maintain that there is nothing at stake with such extreme explanans variable values for the humdrum explanandum in question. One can employ (A) to its full explanatory effect regarding the explanandum at stake, without taking a stand on the murky modal issues concerning such extreme counterfactuals. And in judging (A) to provide the best explanation, we do not, and should not, take a stand on subtle issues in modal semantics upon which evaluating such extreme counterfactuals turns.

In general, we should say that an increase in scope generality is explanatorily valuable only to the extent that it increases the provision of correct answers to what-if questions that are pertinent and reasonable from the point of view of empirically well-supported scientific theories. It is a challenge to pin down *exactly* what counts as a reasonable what-if question, but we do not need to meet it here. It suffices for us to have argued that there is some point after which the implicated counterfactuals are not reasonable in the relevant sense. We can conclude from this that, in as far as a distinctively mathematical explanation scopes beyond any nominalistic alternative, the difference trades only in extreme counterfactuals, so we have good naturalistic grounds to deny that this explanation’s maximal generality in and of itself renders the explanation better. Given that we can offer an alternative nominalist account of distinctively mathematical explanations’ superiority within the counterfactual framework (§5), this will suffice to rebut the challenge from explanatory generality, as far as scope generality is concerned.

Topic Generality From the counterfactual point of view, topic generality, whether or not achieved with mathematics, only contributes to a *particular* explanation if it enables the answering of more of the relevant *what-if* questions *concerning that explanation’s explanandum*.

To illustrate, suppose we have hitherto been unable to explain why mother cannot divide twenty-three strawberries between her three children. By learning the explanatory core of our toy-example, (A), an explanation for why mother fails presents itself, putting us in a position to answer a range of relevant *what-if* questions regarding

mother's predicament. We have integrated two disparate structurally similar phenomena to the arithmetical background theory. This is a good thing. But the new *what-if* questions we can now answer (with respect to mother) do not yield any further explanatory understanding of why the philosopher cannot divide twenty-three ideas equally between three sections. We can of course answer further what-if questions regarding the philosopher by first answering the corresponding questions about mother, and then translating. But these questions would be answerable via (A) directly, so all the topic generality has done here is open up a new, indirect means of achieving the same understanding of our original explanandum. It has not *increased* our understanding. From a more global perspective, our stock of answerable *what-if* questions has increased; but this is only because we have formulated a new explanation with its own associated *what-if* questions. In cases such as these, topic generality is clearly very useful, but it does not itself increase the explanatoriness of any of the particular explanations.⁴

Turning again to the cicada case, consider an example from Baker (2017). Why is it that in fixed-gear bicycles with 14-tooth rear cogs and front cogs with either 47, 48, or 49 teeth, those with 47 in front minimise the wear on their rear tire? The best explanation involves the same mathematics as the cicadas case: of (14, 47), (14, 48), and (14, 49), only the first is a *coprime* pair, and coprime integers maximise their LCM with all lower integers. This means that bikes with 47-tooth front cogs will maximise the number of full pedal-turns required to make the rear cog (and therefore the rear tire) return to its original position, and so are less likely to stop on the same patch of tire. Since this explanation is structurally similar to the cicada case, we can give a topic general explanation that covers both. Baker achieves this by expanding the explanatory core of the cicada explanation (adding some further premises which needn't concern us here). However, there is clearly something odd in thinking that, for example, biologists in possession of the number-theoretic explanation of cicada periods could understand *cicadas* better by reflecting on fixed-gear bicycles. Assuming it is a good one, the biologists' explanation is as powerful as it gets, as far as cicadas are concerned, and no amount of structurally similar applications of number theory enhances it. Thus, from the viewpoint of the counterfactual account, topic generality itself is not an explanatory virtue. It contributes to explanatory power only if it brings with it an explanatory increase in scope generality.

At this point one might object: so much worse for the counterfactual account, if it is unable to capture the significance of *unification* as an explanatory virtue. Granted, rendering an explanation more topic-general doesn't help us to understand the original explanandum better; it does, however, result in a new explanation of a broader class of phenomena that includes both the original and the new explanandum. So, for example, by expanding the core of the cicada example, Baker (2017) has perhaps formulated a new explanation of *both* cicada life-cycles *and* the features of certain fixed-gear bicycle cogs. We now have a choice as to whether to adopt this more general explanation of this more general phenomenon, or keep the relevant explanations separate. We should, the objection goes, prefer a more unified theory of the world, *ceteris paribus*; so, we ought to adopt the more topic general explanation.

⁴ From the perspective of the counterfactual account this is far from being an idiosyncratic feature of mathematical explanations. See, e.g. Ylikoski and Kuorikoski (2010: pp. 214–215), who offer a general discussion of the virtues of formal unification in science without an increase in explanatory information.

This line of reasoning is problematic for two reasons. First, we suspect that, lurking behind this objection is a Quinean way of thinking that is inappropriate in this context: that ontological questions should be settled at the level of our simplest, most unified theory of the world. This way of thinking presupposes the dialectic of the traditional indispensability argument, on which EIA is supposed to improve by avoiding resting the case for platonism on the widely-contested principle of confirmational holism. While unifying two explanations may make for an overall more virtuous (e.g. simpler, more elegant) theory, we fail to see how it can increase our understanding of the respective explananda, for all the reasons presented above.⁵ Nevertheless, we are open to an objection based on an appeal to unification, but more work would need to be done to mount one, as we argue below.

This brings us to our second reason. The above objection fails to appreciate what we achieve by appealing to a theory of explanation that enjoys considerable support independently of the EIA debate. We have shown in this section that, from the perspective of the counterfactual theory, distinctively mathematical explanations exhibit a high degree of scope generality, and in virtue of this are very good explanations. In the following section, we use this same theory to explain why the inclusion of mathematics in these explanations makes them better than they would be without it. This arguably does everything an analysis of a kind of mathematical explanation should do: it accounts for what makes them good explanations, and accounts for the role of mathematics in generating this explanatory goodness. Given this, the fact that our analysis does not recognise unification as an explanatory virtue of distinctively mathematical explanations provides good, non-question-begging reason for thinking that unification is not an explanatory virtue of distinctively mathematical explanations. To respond to this by pointing out that our analysis does not capture unification, and is for that reason inadequate, is to cling to one's intuitions in spite of the evidence we have provided against them.

A more appropriate means of objecting in this vein would be to provide an alternative analysis of distinctively mathematical explanations that enjoys at least as much independent support as our own, and show that, according to it, unification is of genuine explanatory value. The prospects for this appear dim. While there is good evidence that scientists prefer explanations that are more scope general, there is no corresponding evidence that evolutionary biologists would prefer, or have any reason to prefer, an explanation that is topic general enough to apply to both the cicadas and the fixed-gear bicycles. We suspect that evolutionary biologists would have little interest in the more general explanation, qua evolutionary biologists. If anything, there is reason to think that biologists would prefer a less general explanation that is more grounded in the specific phenomenon they are studying. After all, the more general explanation may fudge important distinctions between the respective domains. In general, making explanations more abstract in order to increase the DEGREE OF INTEGRATION between disparate areas of inquiry sometimes results in a loss of important domain-specific detail, resulting in a loss of other explanatory virtues, such as factual accuracy, or cognitive salience. The desirability of this trade-off depends on contextual, prag-

⁵ Although, see Sect. 6 for the related notion that appealing to more general facts can add 'explanatory depth'.

matic factors that determine how the value of different what-if questions is weighed (cf. Ylikoski and Kuorikoski 2010: 213–214).

We have, of course, only provided good reason for thinking that unification is not an explanatory virtue of distinctively mathematical explanations. It is open to our opponents to appeal to other kinds of mathematical explanation and argue that their topic generality is explanatorily valuable in and of itself, and thus that our analysis will not adequately capture them.

5 The Explanatory Contribution of Mathematics

We have now traced the explanatory limits of two kinds of generality to which mathematics contributes. We have seen that, from the perspective of the counterfactual theory of explanation, topic generality improves explanations only to the extent that it results in an explanatory increase in scope generality, and that explanations can be improved through scope generality only up to a point. Given this, the fact that mathematics *maximises* either kind of generality does not ipso facto amount to an explanatory contribution. What, then, accounts for the judgement that distinctively mathematical explanations truly are better than any nominalistic alternative? We will now argue that mathematics makes a genuine explanatory contribution to distinctively mathematical explanations, by maximising these explanations' cognitive salience.

Note first that, for distinctively mathematical explanations, mathematics is not indispensable for generating a desirable level of scope generality. After all, we can in principle state without mathematics the explanatory dependence between the values of explanandum and explanans variables up to the desired point, however scope-general this point is deemed to be. With reference to our toy example, suppose we do this for all possible numbers of ideas up to one exceeding the number of atoms in the observable universe. This is illustrated by the following, which can be written, mathematics-free, with first-order logic plus identity:

- (A') If there is exactly one idea, then the philosopher will not be able to divide it evenly among three sections **and**₍₂₎ if there are exactly two ideas, then the philosopher will not be able to divide them evenly among three sections... **and**_(3×10^{84}) if there are exactly 3×10^{84} ideas, then the philosopher will be able to divide them evenly among three sections.

Even though the explanation based on (A') scopes way beyond any reasonable scenario involving a philosopher writing a paper, and achieves the desired level of scope generality, we agree with the platonists that the explanation turning on (A) is really more explanatory. But not because it (A) further increases or maximises scope generality. Rather, the (A)-explanation is better because it achieves the desired level of scope generality *without compromising cognitive salience*.

To see this, note how long and cumbersome the nominalistic explanation would be. Presenting and using such an explanation to answer *what-if* questions concerning particular counterfactual situations would be unduly difficult. By contrast, (A) throws the same explanatory information into sharp relief, and affords a simple means of answering the desired range of *what-if* questions: just plug in the appropriate numeral

and calculate. The explanation turning on (A) also avoids the feeling that, though there is no explanatorily relevant information to be gained by using (A) to answer *what-if* questions concerning numbers of ideas greater than the number of atoms in the observable universe, there is something arbitrary about stopping the explanation there (or anywhere else). Indeed, we may have no idea at what precise point the answers to the relevant *what-if* questions cease to be explanatorily relevant. But note that none of these virtues that make the (A)-explanation more explanatory are due to (A) containing explanatorily relevant information that cannot be in principle presented without mathematics. Rather, they are due to its presenting the relevant information in a better way. Hence, the indispensable contribution that mathematics makes in relation to scope generality is a matter of improving—indeed maximising—cognitive salience.⁶

The same line of reasoning applies to the cicada example. Consider the following two generalisations, each increasing in their scope:

- (B') Of time periods 12 to 18 years long, 13- and 17-year periods minimise intersection with all periods shorter than 12 years.
- (B'') Of time periods 12 to 18 years long, 13- and 17-year periods uniquely minimise intersection with all periods shorter than 12 years, and, of time periods 16 to 22 years long, 17- and 19-year periods uniquely minimise their intersection with all periods shorter than 16 years.

Explanatory generalisations (B') and (B'') are nominalistic. We think (along with the platonists) that the explanation turning on the earlier mathematical generalisation (B) is truly more explanatory than any of those turning on such finite nominalistic generalisations. Why? Note first that mathematics is not necessary for increasing scope generality here: the explanation turning on (B'') is more scope-general than the explanation turning on (B'), and we can further increase scope generality with explanatory generalisations of the same (nominalistic) kind ad infinitum. Thus, at whatever point the answers to relevant *what-if* questions cease to be explanatorily relevant, we can generate the corresponding level of scope generality without mathematics. Clearly, (B) *maximises* scope generality: there is no limit to the range of possible explanans variables it can take, and so no limit to the range of *what-if* questions it can answer. But this is not why the explanation turning on (B) is better. We can increase the scope of the generalisation non-mathematically well beyond its capacity to answer *what-if* questions which are relevant and reasonable from the point of view of empirically well-supported scientific theories. For example, a mathematics-free explanation of cicadas' life-cycles that accommodates every duration length up to the age of the universe overshoots in this way.

The (B)-explanation is not more explanatory because it maximises scope generality. It is more explanatory because it achieves the desired level of scope generality *without*

⁶ From the perspective of the counterfactual account this is unsurprising, and far from being an idiosyncratic feature of mathematical explanations. See, e.g. Ylikoski and Kuorikoski (2010: 214–215), who argue quite generally that an explanation can improve in terms of its cognitive salience without an increase in explanatory information.

compromising cognitive salience. A nominalistic explanation that achieves the desired level of scope generality would be extremely long and unwieldy, so extracting from it answers to *what-if* questions would be difficult, time-consuming, and cognitively opaque to us. In contrast, the number-theoretic rule makes generating answers to *what-if* questions very easy and cognitively transparent. It also avoids our having to answer the extremely difficult and somewhat tangential question of at what precise point the relevant *what-if* question cease to be explanatorily relevant. None of these benefits are due to (B) presenting explanatory information that cannot be presented nominalistically. They are due to its presenting the relevant information in a better way.

We conclude that mathematics' role in the procurement of explanatory for distinctively mathematical explanations is a matter of improving cognitive salience. This analysis is not specific to a number-theoretic explanation, and it applies, *mutatis mutandis*, to other well-known examples of distinctively mathematical explanations, such as Euler's graph-theoretic explanation of the non-traversibility of Königsberg's bridges. (See Jansson and Saatsi (2017) for a starting point for this latter example.)

6 Objection and Reply

Before we wrap-up, we need to consider one more potential objection. Perhaps we have unduly focused on explanatory generality in terms of an explanation's scope of application (what we might call 'explanatory breadth'), and ignored a no less important form of understanding offered by mathematics: that of picking out a more general feature of reality whose instantiation is responsible for the dependencies at stake (what we might call 'explanatory depth'). Take the cicada example. Any nominalistic generalisation stating which durations minimise intersections up to some finite limit presents a series of facts about time, revealing a pattern of dependence between certain time periods and the obtaining of an intersection-minimisation relation between them. It does not explicitly tell us *why* these facts obtain, or why the pattern emerges.

In contrast, the objection goes, (B) arguably tells us why this pattern holds, by showing that each of these particular matters of fact (that p_1 maximises its LCM with each $q < p_1$, that p_2 maximises its LCM with each $q < p_2$, etc.) obtain in virtue of a single, more general fact: that all and only prime numbers are coprime with all lower integers. If this is a genuine explanation, it is a mathematical explanation of a mathematical fact—an *intra*-mathematical explanation; but, it could be argued, it furnishes a deeper understanding of the associated facts about time. Again, we have a pattern of particular matters of fact—that each of a certain set of time periods minimises intersection with each time period shorter than itself—and the idea is that (c') reveals that each of these particular facts obtains in virtue of a single, more general fact, and thus provides a deeper explanation than any nominalistic alternative. This kind of platonist story can be found in the literature.⁷

⁷ Baker (2017: 199) appears to have this in mind when he says that the number-theoretic explanation is *deeper* than mathematics-free alternatives, because it tells us why there are unique intersection-minimising periods within some ranges of periods and not others. Pincock (2015) argues that certain mathematical explanations explain by identifying physical facts as particular instances of more general, abstract facts.

Our response is as follows. One should not be misled by the way in which the nominalistic generalisations are presented. While they are in a sense a conjunction of facts, these facts are not disunified. The aim of explaining on the counterfactual account is to characterise an actually obtaining relation of dependence between the explanans and the explanandum, by revealing how changes in the former would change the latter. What is revealed in doing this is not just a bunch of disjoint facts that cry out for unification in terms of something else; it is how the actual dependence relation holding between the explanans and the explanandum manifests itself across various relevant counterfactual scenarios. So, while the platonist may want to say that the pattern of dependence holds in *this* way because of *that* mathematical fact, the nominalist will just say that the pattern of dependence holds in this way because that is how the dependence relation represented by the explanation manifests itself. Thus, from the perspective of the counterfactual account, the charge that any nominalistic generalisation would make for a comparatively shallow explanation is misplaced.

This is unlikely to convince someone strongly under the sway of the intuition that the mathematically-presented explanations are somehow ‘deeper’. However, as things stand, that is all it is: an intuition. As we have stated, our aim is to demonstrate how to push past the trading of intuitions, and this is where our appeal to a theory of scientific explanation that is motivated independently of the debate really pays off. Our analysis explains why distinctively mathematical explanations are to be preferred over their nominalistic alternatives independently of their generating depth in the above sense, and it does so by folding it into a general and independently well-supported theory of scientific explanation. On our story, the place of distinctively mathematical explanations in a broader story about the preferences and standards operative in scientific practice is clear, and this sets a very high bar. Anyone wishing to argue for platonism by appeal to mathematics’ procurement of depth (or any other ‘virtue’ our analysis leaves out) will have to develop an analysis that rivals our own in its naturalistic credentials.⁸ Thus, whether our opponents find the above response convincing or not, they are forced to do further work in the methodological spirit of this paper to mount a counter-reply. Resting on their intuition about the relative depth will not suffice. This marks precisely the kind of progress we have aimed to achieve in this paper.

7 Conclusion

There is a critical, but seldom-recognised assumption shared by both sides in the deadlocked debate about EIA: that the source of an explanation’s explanatory power is purely a matter of its correctly representing explanatory features of reality. On the basis

Footnote 7 continued

Others highlight cases where the mathematics used to explain what look like very different physical phenomena can be unified under a more general mathematical method, and argue that this suggests there is a deeper, mathematical reason lying behind the relevant physical phenomena (e.g. Colyvan 2002 and Baron et al. 2017).

⁸ For example, Baron et al. (2017) develop a counterfactual analysis of mathematical explanation with an eye to unifying intra- and extra-mathematical explanations, so their account may offer a means of spelling out the above kind of depth counterfactually. They hold back from endorsing the analysis; but if they were to endorse it, and use it in support of EIA, they would have to show that their account genuinely captures something of scientific explanatory value that ours misses.

of this assumption, platonists can challenge nominalists by reference to mathematical explanations that are judged to be more explanatory than nominalistic alternatives: the difference in their explanatoriness must be due to the mathematics representing explanatory features of reality that their nominalistic alternatives do not, and nominalists are thereby challenged to say what those features are, if not mathematical. This critical assumption lies behind the challenge from explanatory generality, based on the popular idea that mathematics makes certain explanations better than nominalistic alternatives by making them *more general*.

Our analysis reveals that nominalists can reject this critical assumption: explanations can become more explanatory without accurately representing further explanatory features of reality. According to the counterfactual account, explanatory power is a measure of how much modal explanatory information it provides to an explainer. This is a function both of the amount of such information the explanation provides about the relevant worldly features, *and* of the way in which this information is presented. We take it as obvious that only increases in information about the explanatorily relevant worldly features—objective explanatory dependences—as opposed to changes in how such information is presented, stand a chance of securing further ontological commitments. Our analysis of some of the key exemplars of distinctively mathematical explanation, displaying striking scope and topic generality, shows that mathematics does not yield an increase in this kind of ontologically committing explanatory information. Its role is rather that of improving cognitive salience. From the perspective of the counterfactual account, we thus get a clear verdict in favour of nominalism: mathematics' indispensable role in distinctively mathematical explanations provides no reason to believe in mathematical objects.

There is some scope to resist our analysis of mathematics' contribution, e.g. by arguing that there is a further, scientifically kosher dimension of explanatory power that our analysis doesn't capture. Any such claim needs to be accompanied by evidence that avoids the kind of intuition-trading and foot-stomping that we have avoided by working within a well-founded theory of scientific explanation. As far as we can see, the only way to do this would be to motivate an alternative account of mathematics' contribution to distinctively mathematical explanations within an equally well-supported theory of explanation.

Acknowledgements We would like to thank anonymous reviewers and Steven French for detailed comments on earlier drafts, as well as John Wigglesworth and other members of the Philosophy of Science reading group at University of Leeds. Thanks also to audiences at University of Leeds, University of Vienna, and University of Milan, where this work has been presented. Robert Knowles gratefully recognises the British Academy for partly funding this research (Grant No. 107581).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Baker, A. (2005). Are there genuine mathematical explanations of physical phenomena. *Mind*, *114*(454), 223–238.
- Baker, A. (2017). Mathematics and explanatory generality. *Philosophia Mathematica*, *25*(2), 194–209.
- Baker, A., & Colyvan, M. (2011). Indexing and mathematical explanation. *Philosophia Mathematica*, *19*(3), 323–334.
- Baron, S., Colyvan, M., & Ripley, D. (2017). How mathematics can make a difference. *Philosophers' Imprint*, *17*, 1–19.
- Bokulich, A. (2016). Fiction as a vehicle for truth: Moving beyond the ontic conception. *The Monist*, *99*(3), 260–279.
- Buchsbaum, D., Bridgers, S., Skolnick Weisberg, D., & Gopnik, A. (2012). The power of possibility: Causal learning, counterfactual reasoning, and pretend play. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1599), 2202–2212.
- Cartwright, D. E. (2000). *Tides: A scientific history*. Cambridge: Cambridge University Press.
- Colyvan, M. (2002). Mathematics and aesthetic considerations in science. *Mind*, *111*, 69–74.
- Colyvan, M. (2013). Road work ahead: Heavy machinery on the easy road. *Mind*, *121*(484), 1031–1046.
- French, S., & Saatsi, J. (2018). Symmetries and explanatory dependencies in physics. In A. Reutlinger & J. Saatsi (Eds.), *Explanation beyond causation: Philosophical perspectives on non-causal explanations*. Oxford: Oxford University Press.
- Hitchcock, C., & Woodward, J. (2003). Explanatory generalizations, part II: Plumbing explanatory depth. *Nous*, *37*, 181–199.
- Jackson, F., & Pettit, P. (1990). Program explanation: A general perspective. *Analysis*, *50*, 107–117.
- Jansson, L., & Saatsi, J. (2017). Explanatory abstractions. *The British Journal for the Philosophy of Science*, *70*(3), 817–844. <https://doi.org/10.1093/bjps/axx016>.
- Knowles, R., & Liggins, D. (2015). Good weasel hunting. *Synthese*, *192*(10), 3397–3412.
- Lange, M. (2013). What makes a scientific explanation distinctively mathematical. *The British Journal for the Philosophy of Science*, *64*(3), 485–511.
- Liggins, D. (2016). Grounding and the indispensability argument. *Synthese*, *193*(2), 531–548.
- Lyon, A. (2012). Mathematical explanations of empirical facts, and mathematical realism. *Australasian Journal of Philosophy*, *90*(3), 559–578.
- Melia, J. (2000). Weaseling away the indispensability argument. *Mind*, *109*(435), 455–480.
- Pincock, C. (2015). Abstract explanations in science. *The British Journal for the Philosophy of Science*, *66*(4), 857–882.
- Plebani, M. (2016). Nominalistic content, grounding, and covering generalizations: Reply to 'grounding and the indispensability argument'. *Synthese*, *193*(2), 549–558.
- Potochnik, A. (2018). Eight other questions about explanation. In A. Reutlinger & J. Saatsi (Eds.), *Explanation beyond causation: Philosophical perspectives on non-causal explanations*. Oxford: Oxford University Press.
- Reutlinger, A. (2016). Is there a monist theory of causal and noncausal explanations? The counterfactual theory of scientific explanation. *Philosophy of Science*, *83*(5), 733–745.
- Reutlinger, A. (2018). Extending the counterfactual theory of explanation. In A. Reutlinger and J. Saatsi (Eds.), *Explanation beyond causation*. Oxford: Oxford University Press.
- Saatsi, J. (2012). Mathematics and program explanations. *Australasian Journal of Philosophy*, *90*(3), 579–584.
- Saatsi, J. (2016). On the 'indispensable explanatory role' of mathematics. *Mind*, *125*, 1045–1070.
- Saatsi, J. (2017). Dynamical systems theory and explanatory indispensability. *Philosophy of Science*, *84*(5), 892–904.
- Saatsi, J. (2018). On explanations from geometry of motion. *The British Journal for the Philosophy of Science*, *69*, 253–273.
- Wakil, S., & Justus, J. (2017). Mathematical explanation and the optimization fallacy. *Philosophy of Science*, *84*(5), 916–930.
- Woodward, J. (2003). Experimentation, causal inference, and instrumental realism. In H. Radder (Ed.), *The philosophy of scientific experimentation* (pp. 87–118). Pittsburgh: University of Pittsburgh Press.
- Yablo, S. (2013). Explanation, extrapolation, and existence. *Mind*, *121*(484), 1007–1029.

Ylikoski, P., & Kuorikoski, J. (2010). Dissecting explanatory power. *Philosophical Studies*, 148(2), 201–219.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.