Journal : SmallExtended 10670 | Article No : 466 | Pages : 24 | MS Code : 466 | Dispatch : 2-11-2021

Erkenntnis
https://doi.org/10.1007/s10670-021-00466-x

ORIGINAL RESEARCH

Counter Countermathematical Explanations

Atoosa Kasirzadeh[1,2,3]

Received: 14 April 2020 / Accepted: 3 September 2021

Abstract
Recently, there have been several attempts to generalize the counterfactual theory of causal explanations to mathematical explanations. The central idea of these attempts is to use conditionals whose antecedents express a mathematical impossibility. Such countermathematical conditionals are plugged into the explanatory scheme of the counterfactual theory and—so is the hope—capture mathematical explanations. Here, I dash the hope that countermathematical explanations simply parallel counterfactual explanations. In particular, I show that explanations based on countermathematicals are susceptible to three problems counterfactual explanations do not face. These problems seriously challenge the prospects for a counterfactual theory of explanation that is meant to cover mathematical explanations.

1 Introduction

Philosophical accounts of causal explanation in terms of counterfactuals have enjoyed popularity at least since Lewis (1973a, 1986).[1] Such counterfactual accounts, roughly, say that Suzy throwing a rock explains why the window shattered, because the counterfactual conditional *if she had not thrown the rock, the window would not have shattered* is true; that is, Suzy's throw makes a difference as to whether or not the window shatters. The prospect of extending the counterfactual accounts to mathematical explanations is appealing. If it could be done, we would be on the road to acquire a general theory of explanation in science and mathematics. Generality, some argue, is a virtue that ideally a theory of explanation should satisfy (Nickel 2010; Reutlinger et al. 2020). Moreover, the success of a counterfactual theory of mathematical explanation would have resounding impacts on the

1FL01 [1] For an overview of non-counterfactual accounts of causation, see Andreas and Günther (2021).

Atoosa Kasirzadeh
atoosa.kasirzadeh@mail.utoronto.ca

[1] University of Toronto, Toronto, Canada

[2] Australian National University, Canberra ACT 0200, Australia

[3] University of Edinburgh, South Bridge, Edinburgh, UK

Journal : SmallExtended 10670	Article No : 466	Pages : 24	MS Code : 466	Dispatch : 2-11-2021

debates about metaphysical explanation, grounding, logical explanation, artificial intelligence explanations, and non-causal explanations more generally (Schaffer, 2016; Wilson, 2018a, b; Maurin, 2019; Baron, 2020; Kasirzadeh & Smart, 2021).

Recently, there have been several attempts to liberate counterfactual accounts of explanations from their causal trappings (see for instance, Reutlinger, 2016; Baron et al., 2017; Woodward, 2018; Baron et al., 2020; Reutlinger et al., 2020). Among these, the most elaborate and systematic endeavor to extend the counterfactual theory of causal to mathematical explanations is due to Baron et al. (2017, 2020), which I will abbreviate henceforth by BCR.

BCR claim that akin to an empirical fact such as Suzy throwing a rock, a mathematical fact can also make a difference. Accordingly, they maintain that we can understand the explanatory structure of a mathematical explanation in terms of counterfactual dependency between its mathematical explanantia and its explanandum. Just like an empirical explanans, a mathematical explanans would fit into the *explanatory scheme of the counterfactual theory*[2]:

$$\begin{aligned} C \text{ explains } E \text{ if } &(1)\ \text{C and } E \text{ are true, and} \\ &(2) \text{ if } C \text{ were not true, } E \text{ would not be true.} \end{aligned}$$

Suppose C denotes a mathematical fact and E an empirical or a mathematical fact. Then we say C mathematically explains E. Furthermore, (2) becomes a *countermathematical*; that is, a conditional whose antecedent expresses a mathematical impossibility.[3] Finally, we refer to the conjunction of (1) and (2) as a countermathematical explanation.[4]

Here is a countermathematical discussed by BCR (2017):

(I) If 13 were not a prime number, then North American periodical cicadas would not have 13-year life cycles.

According to BCR, this countermathematical reveals that a mathematical fact—the primeness of 13—(partly) explains the fact that North American periodical cicadas have 13-year life cycles. BCR thus propose that a mathematical fact can make a difference just like Suzy's throw can make a difference.[5]

In this paper, I raise three problems for the current counterfactual accounts of mathematical explanations. These problems reveal the extreme difficulty one faces

2FL01 [2] C might be one member of a collection of explanantia for E. A counterfactual theory of explanation
2FL02 must be able to evaluate (2) for all the members of the collection of explanantia.

3FL01 [3] To the best of my knowledge, Lewis (1973b, p. 24) coined the term 'countermathematical'.

4FL01 [4] In this paper, I presuppose that there are genuine cases of mathematical explanations. Without this pre-
4FL02 supposition any generalization of the counterfactual theory of causal explanations to mathematical ones
4FL03 would, of course, be pointless. What I aim to establish is that mathematical explanations—presupposed
4FL04 there are any—cannot be properly analyzed by the current explanatory scheme of counterfactual theories.

5FL01 [5] If we take mathematical facts to be empirical, the antecedent of any countermathematical explanation
5FL02 will be treated similar to a causal explanation in terms of counterfactuals. The main issue is that having
5FL03 a satisfactory empiricist story about mathematics is—to say the least—very difficult. In this paper, in
5FL04 accordance with BCR, I take mathematical facts to be non-empirical.

for an appropriate evaluation of (2). Recall that evaluating (2) is absolutely central to any counterfactual account of explanation. Accordingly, I argue that varying whether or not Suzy throws a rock is entirely different from varying mathematical antecedents such as 13's primeness. Together, the three problems seriously challenge the current quests for a general counterfactual theory of explanation.

The plan of my investigation is straightforward. I outline the counterfactual approach of mathematical explanations in Sect. 2. In Sect. 3, I compare countermathematical and counterfactual explanations. Unlike counterfactual explanations, explanations based on countermathematicals are susceptible to three major problems. Firstly, there is no clear escape route from absurd contradictions when assuming a mathematical impossibility. Secondly, there is sometimes no robust space for tracing the (difference-making) ramifications of varying a mathematical antecedent. Thirdly, a countermathematical explanation provides no explanatory benefits, unlike the variation of the antecedent of a counterfactual. In Sect. 4, I briefly sketch the outline of an alternative approach for tackling some kinds of mathematical explanations. Sect. 5 concludes the paper.

It should be mentioned that BCR's account, as the most elaborate, is central to the current debate about the viability of a counterfactual account of mathematical explanations. Some of the proponents of countermathematical explanations, such as Reutlinger et al. (2020), simply presuppose the validity of BCR's account. Others, such as Woodward (2018), presumably require arguments very much along the lines of BCR to defend their accounts of countermathematical explanations—at least if a mathematical impossibility is supposed to figure in the antecedent. Hence, my counterarguments to BCR's countermathematical explanations carry over—almost unmodified—to the other similar attempts for generalizing the counterfactual theory of explanation. In what follows, I can thus focus my criticisms on BCR's account without losing much generality.

2 Countermathematical Explanations

BCR (2017, 2020) abstract away from any particular counterfactual account of explanation, such as Lewis's (1973, 1986) or the structural-equations framework (Halpern & Pearl 2005). Let us call what is common to these counterfactual accounts of explanation *the counterfactual account of explanation*. At the heart of the counterfactual account is a three-step procedure for the evaluation of conditionals:

(i) Determine the facts to be kept fixed under counterfactual variation.
(ii) Vary some facts as stated in the antecedent.
(iii) Determine the influence of the variation on the consequent.

To illustrate the evaluation procedure, let us apply it. According to the explanatory scheme of the counterfactual theory, Suzy throwing the rock explains the window's shattering if Suzy throws the rock, the window shatters, and if she had not thrown the rock, the window would not have shattered. To evaluate this counterfactual, first, we keep fixed the past up until the time Suzy throws. Second, we vary the fact that

Journal : SmallExtended 10670 | Article No : 466 | Pages : 24 | MS Code : 466 | Dispatch : 2-11-2021

Suzy throws by supposing (contrary to the facts) that she does not throw. Third, against the backdrop of the facts that are kept fixed, the influence of varying the antecedent on the consequent is established. If the consequent varies, that is the window would not have shattered, the counterfactual under consideration is evaluated to be true. If so, Suzy's throwing the rock counterfactually explains the window's shattering. With these preliminaries out of the way, I move to BCR's (2017, 2020) counterfactual account of mathematical explanations.

Generally, mathematical explanations come in two flavors: extra- and intra-mathematical (Colyvan, 2012, Ch. 5; Colyvan et al., 2018). Extra-mathematical explanations explain empirical facts, in part, by mathematical facts (Baker, 2005; Lange, 2013; Lyon & Colyvan, 2008). Intra- mathematical explanations explain mathematical facts such as an explanatory proof for why a mathematical theorem should be accepted (Mancosu, 2008; D'Alessandro, 2020; Lange, 2018). An extra- or intra-mathematical explanation can be expressed in the form of a counterfactual. Recall the explanatory scheme of the counterfactual theory. If C and E are mathematical facts, we have a countermathematical that is obtained from an intra-mathematical explanation. By contrast, if C is a mathematical fact and E an empirical fact, we have an extra-mathematical explanation.[6]

Let me review an extra- and an intra-mathematical explanation in order to fix intuitions about the respective types of countermathematicals.

2.1 Extra-Mathematical Explanations

Perhaps the most familiar example of an extra-mathematical explanation in the philosophical literature is Baker's (2005) case of the North American periodical cicadas. A simplified version of this explanation is as follows. Two sub-species of North American periodical cicadas have life cycles of 13 and 17 years, respectively. Why these two lengths? The explanatory response appeals to two mathematical facts (a), (b), and two empirical facts (c), (d):

(a) 13 and 17 are prime numbers.
(b) Prime numbers maximize their lowest common multiple relative to all lower numbers; that is, they minimize the intersection of periods.
(c) Ecological conditions restrict the life cycle of cicadas to 12–18 years.
(d) The predators of the cicadas have periodical life cycles.

Under the paradigm of evolutionary biology that successful organisms evolve in an optimal way, (a)—(d) explain why North American cicadas have 13-year and 17-year life cycles: prime-numbered life cycles minimize the frequency of co-occurrence with periodical predators with life cycles that are strictly less than the cicada's life-cycle length. This is because the lowest common multiple of two numbers is

[6] BCR (2017) explore the prospects of a counterfactual theory of extra-mathematical explanations. BCR (2020) examine how a counterfactual account of intra-mathematical explanations work. The two accounts are very closely tied to each other.

Journal : SmallExtended 10670 | Article No : 466 | Pages : 24 | MS Code : 466 | Dispatch : 2-11-2021

maximal if and only if the two numbers are coprime. A cicada having a 15-year life cycle overlaps periodically with predators having 1-, 3-, 5-, and 15-year life cycles. A cicada with a 13-year (17-year) life cycle, by contrast, overlaps only with predators of 1- and 13-year (17-year) life cycles.

BCR's (2017) counterfactual account of extra-mathematical explanation holds that the variation of the mathematical fact (a) makes a difference to the optimal life-cycle length of the cicadas. If 13 were not a prime number, *ceteris paribus*, North American periodical cicadas would not have evolved to have 13-year life cycles. Why? If 13 were not prime, it would have factors in addition to 1 and 13, and thus the 13-year life cycle would overlap with more than two life cycles. Hence, 13-year life cycles would not be optimal any more to avoid predators. According to BCR, the truth of a conditional such as (II) establishes why the optimal life-cycle length of 13 years is explained by 13's primeness:

(II) If, in addition to 13 and 1, 13 had the factors 2 and 6, North American periodical cicadas would not have 13-year life cycles.

On BCR's (2017, p. 4) account, mathematical facts are necessarily true. Varied mathematical facts are thus impossible. The antecedent of (II) expresses an impossibility, making the conditional a so-called 'counterpossible'. A countermathematical hence is a counterpossible whose impossible antecedent is mathematical.

There are two general approaches to the evaluation of counterpossibles: vacusim and non-vacuism.[7] In accordance with BCR, in this paper, I assume that a non-vacuist evaluation procedure is the sensible route to adopt when evaluating countermathematicals.

BCR's (2017, p 7) proposal for evaluating countermathematicals keeps classical logic fixed when varying mathematical facts. After all, the ordinary cases of mathematical explanations, including the instances discussed by BCR (2017, 2020), Reutlinger (2016), Reutlinger et al. (2020), and Woodward (2018), are generated from the mathematical facts of classical logic. I will discuss what happens to their account if we move to a domain of mathematics based on a contradiction-tolerant logic in the next section. For now, I would like to emphasize that their commitment to classical logic forces us to prevent any absurd contradictions. BCR's (2017) alleged solution is to keep fixed as much of mathematics as possible without engendering contradiction. In BCR's (2017, p. 7) terms:

> Here's our suggestion: work backwards from the desired twiddle. First, twiddle 13 and hold some portion of the number theory structure fixed. Does a contradiction result? If yes, then relax the amount you've held fixed

7FL01 [7] On the one hand, vacuists such as Williamson (2018) claim that all counterpossibles with impossible
7FL02 antecedents are true. On this view, all countermathematical explanations are true. This gives too many
7FL03 countermathematical explanations. On the other hand, non-vacuists such as Nolan (1997), Berto et al.
7FL04 (2017), and BCR (2017, 2020) argue that some counterpossibles are false and some are true. On this
7FL05 view, a mathematical impossibility may or may not explain another fact depending on whether the cor-
7FL06 responding countermathematical comes out true or false.

Journal : SmallExtended 10670 | Article No : 466 | Pages : 24 | MS Code : 466 | Dispatch : 2-11-2021

> and re-twiddle. Does a contradiction result? If yes, then relax the amount you've held fixed and re-twiddle. Does a contradiction result? If yes ... And so on. Stop when you get to the maximal amount you can hold fixed within mathematics without inducing a contradiction.

How does this solution apply to the case of the cicadas? BCR (2017, p. 7) claim that we can provide a 'surgical strike' on the primeness of 13: one 'can hold all of number theory fixed *except* for the twiddles to 13 if one is prepared to change the way multiplication works'. Their reason is that there can be a varied version of the multiplication operator, namely multiplication*, which works exactly like multiplication, except that it maps the inputs 2 and 6 to 13. As I will take issue with their claim in the next section and do not want to misrepresent their remarks on multiplication*, I will quote them at length:

> Multiplication* will preserve the same theorems as multiplication, and imbue the natural numbers with the same structure, except for whatever disruption is involved in changing the factors of 13; [...] Moreover, the structure will be consistent just if multiplication* does not take one set of numbers as input and map those same numbers onto two different outputs. Because functions are so easy to come by, we can be assured that there is some function that behaves exactly this way, and so no contradictions will arise by twiddling multiplication so that it matches multiplication*.

BCR (2017) adapt the abstract three-step procedure (i)—(iii) of Sect. 2 to evaluate a countermathematical that figures in an extra-mathematical explanation as follows:

- (i′) Keep fixed as much as possible about mathematics and the empirical world under countermathematical variation.
- (ii′) Vary the mathematical facts in the antecedent while respecting (i′); that is, keep fixed as much of mathematics as possible consistent with the variation.
- (iii′) Determine the influence of this variation on the empirical consequent.

Let us apply BCR's evaluation recipe to the countermathematical (II). First, keep fixed a structural morphism between number theory and the empirical domain of the cicada life cycles, in particular how the structure of natural numbers map on the structure of life-cycle lengths in years. Second, vary some facts of number theory such that 13 has the factors 1, 2, 6, and 13 while keeping as much of mathematics fixed as is possible in a consistent way; for instance, change the multiplication operator to multiplication*. This results in a number theory* that is as much as possible like ordinary number theory except that 13 is not prime. Third, because the morphism between number theory and the empirical domain is kept fixed, the variation of number theory to number theory* implies that a cicada with 13-year life cycle overlaps with predators having 2-year and 6-year life cycles. As a result, a 13-year life cycle is not optimal to avoid predators. The countermathematical (II) thus comes out true, or so argue BCR (2017).

Journal : SmallExtended 10670 | Article No : 466 | Pages : 24 | MS Code : 466 | Dispatch : 2-11-2021

Let me briefly review an instance of intra-mathematical explanation, before I move to examining three major problems with countermathematical explanations.

2.2 Intra-Mathematical Explanations

Consider the following number-theoretic fact: (Γ) The product of any three non-zero, consecutive natural numbers is divisible by 6. Why? The explanation appeals to two mathematical facts α and β (e.g., Lange, 2014):

(α) For any three consecutive nonzero natural numbers, at least one of those numbers is even and therefore divisible by 2.
(β) For any three consecutive nonzero natural numbers, exactly one is divisible by 3.

α and β entail (Γ): the product of any three non-zero, consecutive natural numbers is divisible by $2 \times 3 = 6$. This explanation is used to illustrate the basic idea behind BCR's (2020) counterfactual account of intra-mathematical explanation: the variation of the mathematical explanans (α) makes a difference to the mathematical explanandum (Γ). To show this, BCR (2020) claim that we must first evaluate the following countermathematical expressing the explanatory structure between (α) and (Γ):

(III) If it were not the case that for any three consecutive nonzero natural numbers, at least one of those numbers is even (and therefore divisible by 2), then it would not be the case that the product of any three non-zero, consecutive natural numbers is divisible by 6.[8]

How to evaluate (III)? BCR (2020, p 26) suggest that we can adjust the recipe (i)—(iii) to evaluate the intra-mathematical explanations figuring in countermathematicals as follows:

(i″) Keep fixed as much as possible about mathematical facts and their intrinsic properties under counterfactual variation.
(ii″) Vary the mathematical facts in the antecedent while respecting (i″); that is, keep fixed so much of ‘upstream mathematics’ as possible consistent with the variation.
(iii″) Determine the influence of this variation on the mathematical consequent.

[8] (β) also explains (Γ). To establish this, another countermathematical must be evaluated: (IV) If it were not the case that for any three consecutive nonzero natural numbers, exactly one is divisible by 3, then it would not be the case that the product of any three non-zero, consecutive natural numbers is divisible by 6. The recipe for the evaluation of (IV) is very similar to that of (III). For simplicity, I only focus on the evaluation of (III).

Journal : SmallExtended 10670 | Article No : 466 | Pages : 24 | MS Code : 466 | Dispatch : 2-11-2021

What are the intrinsic properties of mathematical facts? BCR (2020) suggest a notion similar to Lewis's (1983) duplication-based conception of intrinsicality. According to this notion, a property is intrinsic if and only if, for any two duplicate things, either both have the property or neither does (Lewis, 1983, pp. 355–356).

How to apply this notion of intrinsicality to the realm of abstract mathematics? BCR (2020) propose the following:

> In the mathematical case, as in the non-mathematical case, this means holding fixed as much as we can concerning the intrinsic properties of whatever mathematical features are mentioned in the antecedent of a given counterfactual, compatible with realising the antecedent itself. The less we hold fixed about the intrinsic properties of whatever we are interested in, the less confident we should be in the outcome of the evaluation procedure. That is because the counterfactual situation we end up considering may bear little resemblance to the actual scenario at issue in relevant respects (i.e., respects of intrinsic similarity).

In practice, however, this proposal remains utterly unclear as BCR (2020) do not sketch at all how to make sense of 'two duplicate numbers', or how we should even start thinking about the intrinsic properties of numbers. For the sake of argument, let us assume we can fix some intrinsic properties of numbers, whatever they are. In other words, let us assume that somehow (i$''$) is obtained. The evaluation recipe for (III) goes then as follows. Consider three nonzero, consecutive natural numbers such as 503, 504, and 505. Tweak the natural numbers by making them such that none of 503, 504, 505 is even. We get to step (iii$'''$). The product of any two non-zero, natural numbers is even if and only if at least one of the numbers is. The product of these numbers is $(503 \times 504) \times 505$. According to the tweak, 505 is not even. So, we should turn to 503 and 504. Again, according to the tweak, neither is even. Therefore, $(503 \times 504) \times 505$ is not even. A requirement for divisibility by 6 is that the number is even. $(503 \times 504) \times 505$ is not even, hence $(503 \times 504) \times 505$ is not divisible by 6. Therefore, (III) is true, or so BCR (2020, p 26) argue.

3 Counter Countermathematical Explanations

In this section, I develop three arguments against the current counterfactual accounts of mathematical explanation. In particular, I question the plausibility of a principled procedure for evaluating explanatory countermathematicals. The first argument points out that there is no clear escape route from absurd contradictions when assuming a mathematical impossibility. This argument questions whether the mathematical explanans figuring in the antecedent can be meaningfully varied in the context of mathematical explanations. The second argument says that sometimes there is no robust space for tracing the (difference-making) ramifications of the twiddled mathematical fact to the consequent. This questions whether the influence of the impossible variation of the antecedent can be robustly and meaningfully determined. The third argument shows that a countermathematical explanation provides no explanatory benefits.

3.1 No Clear Escape Route from Absurd Contradictions

To evaluate a countermathematical, BCR suggest that we should check whether varying mathematical facts results in (absurd) contradiction. If so, we should relax the fixed portion of pure mathematics and vary again, and we continue this procedure until the minimal amount of change without introducing (absurd) contradiction is achieved.

In the rest of this section, I will argue that BCR offer no satisfactory route for preventing absurd contradictions when we vary the antecedent of an ordinary countermathematical explanation. In Sects. 3.1.1 and 3.1.2, I make the argument in relation to mathematics based on classical and contradiction-tolerant logics, respectively. Let us keep in mind that in classical logic, *all* contradictions are unacceptable and hence absurd; that is, there is no difference between acceptable and unacceptable contradictions. However, a contradiction-tolerant logic distinguishes between acceptable and absurd contradictions. While it avoids absurd contradictions, a contradiction-tolerant logic searches for sorting out what acceptable contradictions are.

3.1.1 Mathematics Based on Classical Logic

Let us scrutinize BCR's (2017, pp. 7–8) evaluation recipe for the countermathematical (II): If, in addition to 13 and 1, 13 had the factors 2 and 6, North American periodical cicadas would not have 13-year life cycles. For evaluating this countermathematical, they vary multiplication to multiplication*. On this proposal, multiplication* maps the inputs 2 and 6 to 13, that is (1) $2 \times^* 6 = 13$. Moreover, multiplication* is meant to behave like ordinary multiplication, 'except for whatever disruption is involved in changing the factors of 13'. In particular, multiplication* 'takes all of the same inputs and yields all of the same outputs as multiplication except in one special case of 13'. For instance, (2) $2 \times^* 3 = 6$, (4) $3 \times^* 4 = 12$ and (5) $2 \times^* 2 = 4$. Substituting (2) in (1) we obtain (3) $2 \times^* 2 \times^* 3 = 13$. Substituting (5) in (4) we obtain (6) $2 \times^* 2 \times^* 3 = 12$. In a few steps, we obtain an absurd contradiction, as multiplication* maps the same inputs to 12 and 13; either we must take $12 = 13$, or we must assume that 12 does not belong to the set of natural numbers.

Since BCR's (2017) account requires avoiding contradiction, we need an additional 'disruption', contrary to their strong claim that multiplication* yields the same outputs as multiplication except for the inputs 2 and 6. But which one?

If we disallow substitution of equal parts, mathematics loses considerably in force and usefulness. This undermines BCR's own proposal. For instance, BCR (2017, pp. 7–8) setup the multiplication* operation as follows: 'Whereas multiplication never takes in 2 and 6 and yields 13, multiplication* does exactly that. Moreover, whereas multiplication takes in 2 and 6 and yields 12, multiplication* does not'.

In mathematical terms, this means that $2 \times 6 \neq 13$, $2 \times^{*} 6 = 13$, $2 \times 6 = 12$, and $2 \times^{*} 6 \neq 12$. If we are not allowed to use the substitution of equal parts (substituting 12 for 2×6 in $2 \times 6 \neq 13$), we will not obtain $12 \neq 13$. This result is, of course, needed for BCR's proposal when they set up a distinction between the functioning of multiplication and multiplication*.

If we omit 12 from the set of integers, we violate the recursive nature of the natural numbers. This omission will have some significantly undesirable consequences for BCR's own proposal. Consider the just cited quote above. If there is no 12 in the set of natural numbers, this proposal is void of meaning.

If we deny one of (2) or (5), the sequence of integers will look quite different to the extent that this sequence will not really be a part of actual mathematics based on classical logic. Mathematical facts are 'tightly integrated', as BCR (2017, p. 3) also acknowledge. Hence, varying a mathematical fact propagates through the whole of mathematics, and results in a mathematics that is far from the mathematics applied in the generation of explanations. Why should we accept this distant mathematics as relevant to the counterfactual analysis? BCR offer no answer.

Worse, multiplication* will not preserve the same theorems of number theory as multiplication. Consider, for instance, the fundamental theorem of number theory:

> Every integer greater than 1 either is a prime number itself or can be represented as the product of primes. Moreover, each integer has one and exactly one prime factorisation.

Now, $2 \times^{*} 2 \times^{*} 3 = 12$ and $2 \times^{*} 2 \times^{*} 3 = 13$ outrightly violate the fundamental theorem of number theory. If we change this fundamental theorem, we would radically change actual mathematics. For instance, what becomes of Goldbach's conjecture that every even integer greater than 2 can be expressed as the sum of two primes? In this case, BCR's (2017, p 3) surgical strike on 13's primeness became in no time a doomsday attack on number theory. It is at best misleading to say that multiplication* 'will preserve the same theorems as multiplication and imbue the natural numbers with the same structure, except for whatever disruption is involved in changing the factors of 13'. As noted earlier, any kind of contradiction in classical logic is absurd. BCR's (2017) cited claim remains hollow as long as they do not delineate precisely which theorems and which structures are preserved, which disruptions are required, and most crucially how the preservation and disruption are possible.

One potential fix, as BCR (pp. 8–9, 2020) presume, is to divide between the 'upstream' and 'downstream' facts of mathematics. Let us assume that relative to a fact F_m of a mathematical structure, we have a procedure to divide the upstream and the downstream facts. The 'upstream mathematical facts' are those within a mathematical structure on which F_m depends. The 'downstream facts' from a given mathematical structure are those that depend upon F_m.

On BCR's (2020) account, when evaluating a countermathematical explanation, we hold fixed as many general, upstream mathematical principles as possible. Those mathematical principles that are downstream to the tweaked mathematical fact are not hold fixed, as much as possible, so that the tweak has enough conceptual space to ramify properly. Although this proposal might seem theoretically promising, unfortunately, in practice it does not resolve any of the issues

I raised above. As illustrated in six simple steps (1)—(6), an absurd contradiction obtains in a very small neighboring region of natural numbers. These steps only rely on a very small vicinity of natural numbers composed of 2, 3, 4, 12, 13, and multiplication*. 2, 3, and 4 appear in the sequence of natural numbers in an upstream way: the recursive definition of natural numbers starts from 2, 3, 4 and only after it arrives at 12 and 13. So, even if we have a procedure to distinguish between the relevant upstream and downstream mathematical facts, still the uprising of absurd contradictions occurs in this local neighborhood (or we do not know how to treat them), according to the discussion above. Classical logic simply does not allow for distinguishing between local and absurd contradictions: all contradictions are absurd in classical logic.

Very similar worries apply to the cases of intra-mathematical explanations. Recall the following countermathematical:

(III) If it were not the case that for any three consecutive nonzero natural numbers, at least one of those numbers is even and therefore divisible by 2, then it would not be the case that the product of any three non-zero, consecutive natural numbers is divisible by 6.

Let us apply BCR's (2017) procedure for the evaluation of (III). Consider 503, 504, and 505 as three nonzero, consecutive natural numbers. Now, tweak the natural numbers such that none of 503, 504, 505 is even, and keep everything else in the immediate vicinity of the tweak fixed. In a mathematics based on classical logic, either a natural number is divisible by two or is not divisible by two. As a result of the tweak, 504 is not even, and so it means that it is not divisible by two. According to a plausible interpretation of upstream facts, the following mathematical fact about 504 is upstream: $504 = 252 \times 2$. As a result of the tweak, 504 is not even, which means that 252×2 is not even, and therefore not divisible by 2. This means that neither 252 nor 2 is divisible by 2. We get to an absurd contradiction: 2 is not divisible by 2. One escape route might be to change the multiplication operator to another operator such as multiplication**. Let's say multiplication** behaves just like multiplication except that $504 \neq 252 \times 2$. Fair enough, but what is 504 equal to? We must run into very similar problems as with multiplication* outlined above.

One might object to this argument that the factors of 504 are irrelevant to its evenness. I think this is false because 'even' means divisible by 2, or having the factor 2. If the factors are irrelevant to whatever 504 means, what is remaining of this number's meaning? The proponents of countermathematical explanations owe us an answer.

There might be a different way to interpret BCR's proposal. This way requires to identify propositions with sets of possibilities, and then interpret possibilities not in a mathematical way, but in terms of what *you* consider to be possible or impossible (see Huber (2021, Ch. 6) for a formal sketch of this treatment). This interpretation remains open to full investigation, and I do not tackle it in details here. However, I see a potential problem for applying this agent-relative interpretation to the counterfactual analysis of mathematical explanations. A reasonable counterfactual analysis of extra- and intra-mathematical explanation should be in search of correct truth values of some sort, rather than an individual belief about the truth or falsity of a countermathematical explanation. After all, we want a counterfactual analysis of mathematical and scientific explanations to be robust enough in delivering what such explanations are, and not coming out true for person *A*, false for person *B*, and indeterminate for person *C* depending on and sensitive to different individual's interpretations of possibility or impossibility. BCR (2017, p. 6), for example, assert that 'we're just going to assume that these counterfactuals are true and then give a way of evaluating these counterfactuals that yields their correct truth-values'.

To evaluate a countermathematical in terms of what an individual considers possible or impossible about mathematical facts can yield interesting results for the acceptability of or the belief in countermathematicals, but not for their truth. For the truth value of a countermathematical it seems that what a person considers possible is too subjective. And since mathematical explanations require a true counterfactual, mere acceptability or belief in this countermathematical is not enough to establish mathematical explanations. That is, the belief of an individual about the possibility or impossibility of a mathematical fact, as noted earlier, is not in the business of establishing countermathematical explanations. A correct truth value, rather than a purely subjective belief about a counterfactual, seems to be the plausible robust constraint for extending the counterfactual accounts to intra- and extra-mathematical explanations in science and mathematics.

So far, I have established that varying mathematical facts of the antecedent of a countermathematical either leads to inevitable absurd contradictions or provides serious challenges to BCR's proposal for two examples. It is easy to see how these examples generalize for other instances. As a result, given BCR's account, we do not really know how to make the 'surgical strikes' on mathematical facts of interest. It follows that the second step of BCR's evaluation recipe for counterpossibles can easily fail. (ii$'$) and (ii$''$) require to vary the antecedent of countermathematicals while keeping mathematics consistent with the variation. I have shown that BCR (2017) cannot even uphold their own example. For instance, varying the number-theoretic fact of 13's primeness by changing multiplication to multiplication* violates the fundamental theorem of number theory, and makes it extremely difficult, if not practically impossible, to know what to make of this distant mathematics with this new number theory. The same is true of varying the evenness of 504. Hence, given BCR's (2017, 2020) proposal, the variation of a mathematical operation, such as the one of multiplication to multiplication*, does not tell us in any clear way how mathematics would change.

Up to this point, I have focused the discussion within the domain of classical logic. On this assumption, varying mathematical facts leads to unavoidable absurd contradictions. Relaxing this assumption and considering a contradiction-tolerant

Journal : SmallExtended 10670 | Article No : 466 | Pages : 24 | MS Code : 466 | Dispatch : 2-11-2021

logic as the basis of mathematics, in which *some* contradictions are allowed and managing some contradictions is possible, might seem to be a solution to save countermathematical explanations. This relaxation, however, invites another set of serious problems.

3.1.2 Mathematics Based on Contradiction-Tolerant Logics

What if the worlds comply with some contradiction-tolerant logic, such as a Priestian paraconsistent logic (Priest, 2002)?[9]In such worlds, *some acceptable* inconsistencies and contradictions might be true. In contrast to classical logic that does not distinguish between the two notions of contradiction and absurdity, one main challenge of a contradiction-tolerant logic is to sort out acceptable contradictions (i.e., contradictions without explosion) from the absurd ones (i.e., contradictions with explosion). From the fact that mathematics can be based on contradiction-tolerant logics, it does not follow that any kind of contradiction is permissible. Proper reasons and proofs must be developed to show that the contradictions such as 13 not being prime (given that it is proven to be prime) or 504 not being even (given that it is proven to be even) are acceptable and not absurd. Allowing for some contradictions does not imply that in any given area of mathematics we can suppose that there are contradictions.

Paraconsistent logics originally came to be in order to deal with some classical self-reference paradoxes such as Russell's paradox, the Liar paradox, and more generally paradoxes that came about in the foundational considerations of mathematics (see Priest, 2002). From a parconsistent perspective, a localised contradiction such as the truth and falsity of the Liar sentence 'This sentence is false.' does not lead to absurd contradictions that trivializes a mathematical theory. So far so good.

Now, the assumption that there are different kinds of contradiction-tolerant logics might sound appealing for the purpose of tweaking the antecedent of a countermathematical explanation. The tempting idea is that mathematics based on a contradiction-tolerant logic can function as a haven safe from absurd contradictions because such mathematics tolerates some acceptable contradictions. For instance, it is tempting to think that tweaking mathematical facts such as the primeness of 13 comes at no serious cost in a variant of mathematics based on a contradiction-tolerant logic.

Before I explore the success of this proposal and to avoid any confusion, let me explicitly reiterate the specific kind of countermathematicals of interest to any generalized counterfactual account of explanation. We are interested in the evaluation of those countermathematicals which have the following form: their antecedent expresses the negation of a mathematical explanans and their consequent is equivalent to the negation of the explanandum. The main question that the defenders of a counterfactual theory of mathematical explanation need to answer is this: does tweaking a mathematical explanans reveal the explanatory structure in terms of the

[9] In addition to Priestian paraconsistent logic, there are other variants of paraconsistent logic as defended by, for example, Batens (1990), Da Costa (1997), and Meheus (2003). For a survey exploring these variants, see Tanaka (2003).

Journal : SmallExtended 10670 | Article No : 466 | Pages : 24 | MS Code : 466 | Dispatch : 2-11-2021

counterfactual dependence between the mathematical explanans and the explanandum, and if so how?

Recall the explanation of the life cycles of North American periodical cicadas. Let us call the explanandum of this explanation *Ex*. Recall (a) 13 and 17 are prime numbers. These numbers are prime in mathematics based on classical logic, and their primeness (rather than their non-primeness) makes the mathematical facts explanatory in the first place. Let us denote mathematics based on classical logic by $\mathfrak{M}$. According to $\mathfrak{M}$, it is either true that 13 is a prime number or it is false. There is no third option. If we choose a mathematical fact from $\mathfrak{M}$, only the two options of truth or falsity are available. That is, the explanans (a) is either true or false when we use $\mathfrak{M}$ to explain *Ex*, and exactly because (a) is true, it becomes an explanans and acquires the explanatory relevance to *Ex*. Only *after* we assume $\mathfrak{M}$ as our reasoning scheme, we agree about what prime numbers, odd numbers, and even numbers are. Therefore, adopting $\mathfrak{M}$ as a reasoning scheme is required to assume the truth of (a) and (b). Recall (b) Prime numbers maximize their lowest common multiple relative to all lower numbers. Only after accepting $\mathfrak{M}$ as our reasoning scheme and the truth of the empirical facts (c) and (d), we could build the mathematical explanation with the explanandum *Ex*. Formally speaking, where A, B, C, and D denote the propositions (a), (b), (c), and (d), respectively, and $\Box$ stands for necessity in mathematics based on classical logic:

$$\mathfrak{M} \models \Box A \text{ and } \mathfrak{M} \models \Box B$$
$$(\Box A \wedge \Box B \wedge C \wedge D) \rightarrow Ex$$

Now, I agree that we might be able to shift the underlying reasoning scheme to mathematics based on a contradiction-tolerant logic $\mathfrak{M}^*$. However, the change in the reasoning scheme does not guarantee that contradictory suppositions such as the non-primeness of 13 and non-evenness of 504 are non-absurd and hence allowed. It might be that for avoiding absurd contradictions, the mathematical facts of interest to ordinary mathematical explanations remain the same; that is, the part of mathematics that incorporates explanatory mathematical facts are bounded with classical logic, because the set of acceptable contradictions is empty (i.e., all contradictions have absurd consequences). Moreover, there is no reason to accept that the world tolerating the impossible mathematics is the relevant world for the countermathematical analysis—i.e., closest to the actual world in which the mathematical explanation of interest holds. As long as BCR's proposal, or any other working proposal along the lines of BCR, does not provide a principled procedure to distinguish between the absurd and acceptable contradictions, any such proposal remains on shaky foundations.

For instance, it might be that in the impossible world of interest in which tweaking mathematical facts is allowed, the relevant bits of mathematics to explanatory reasoning (primeness of 13 or non-evenness of 504) remain untouched by the exotic property of *some well-justified and relevant non-absurd contradictions are acceptable*. In such a world, these bits of mathematics relevant to explanations would stay out of the scope of the acceptable contradictions. As a result, it remains a viable option that a contradiction-tolerant mathematics does not allow for supposing the non-primeness of 13 or non-evenness of 504, even though it allows for other well-justified and acceptable

Journal : SmallExtended 10670	Article No : 466	Pages : 24	MS Code : 466	Dispatch : 2-11-2021

contradictions. We just do not know. For the sake of the argument, let us assume that $A \wedge \neg A$ is allowed in $\mathfrak{M}^*$:

$$\mathfrak{M}^* \vDash A \wedge \neg A$$

Another problem intrudes. By definition, the (members of the collection of) mathematical explanantia for *Ex* should be true (or should hold) in the actual world. That is, truth (or mathematical adequacy) is needed to call a mathematical fact an explanans in the actual world. Now, let us assume that $\mathfrak{M}^* \vDash A \wedge \neg A$. How should we settle the truth value of A in this world? As soon as we suppose $\mathfrak{M}^*$, we move from the actual world to a world where a different mathematics holds. In such a world, it remains unclear what is true or not, especially in relation to the antecedent and the consequent of countermathematicals. And this is simply because some basic rules and laws cannot hold anymore in this distant world.[10] The supposed relation between antecedent and consequent might get lost in translation, so to speak, when moving from a world to another, in which some of the most basic laws and rules do not hold. It could just be that the notion of truth we employ for explanations does not apply to such worlds.

Let me clarify a point before I go further. While I am sympathetic to the non-vacuist proposal that *some* counterpossible conditionals are true and some are false, I disagree that, based on BCR's account or any account along their line, we are able to evaluate that some *explanatory* countermathematicals are true and some are false. For example, I can agree with Berto et al. (2017) that the counterpossible conditional 'If Hobbes had (secretly) squared the circle, all sick children in the mountains of South America at the time would have cared.' is false; whereas 'If Hobbes had (secretly) squared the circle, all sick children in the mountains of South America at the time would not have cared.' is true. However, the arguments in this section support my doubt that we can make such judgments in the case of countermathematical explanations, for which an explanatory relation between the mathematical explanans and the explanandum must hold.

[10] To make this point more concrete, I would briefly describe a case in which an exemplar of inconsistent mathematics, the early infinitesimal calculus, has been used in physics. The early calculus posited that infinitesimals are quantities with zero values in some cases, and non-zero values in other cases within the very same proof (Berkeley, 1734). For instance, consider $f(x) = x^2$. Its derivative, according to early infinitesimal calculus is: $f'(x) = \frac{(x+\delta)^2 - \delta}{\delta}$. On the one hand, the infinitesimal δ must be nonzero, because it appears in the denominator. On the other hand, by simplifying $f'(x)$, we obtain $f'(x) = 2x + \delta$. By taking $\delta = 0$, we get $f'(x) = 2x$. Here, we carry out the reasoning by relying on some global contradictory information: $\delta \neq 0$ and $\delta = 0$. Using these pieces of inconsistent mathematics with care within a particular reasoning scope has resulted in mathematicians doing reasoning with inconsistent mathematics without running into mathematical absurdities such as 2 is not divisible by 2. The set of information by which one could reason at a given time, however, was consistent (McCullough-Benner, 2020). Hence, the fact that sometimes inconsistent mathematics is used to explain or represent an empirical phenomenon, does not mean that in general any kind of inconsistent mathematics can be used to explain any empirical phenomenon, and more relatedly that the explanatory structure between the mathematical explanans and the explanandum of a mathematical explanation can be cashed out by the current theories of counterfactual analysis.

In the next section, I will observe another problem with the current counterfactual accounts of mathematical explanation.

3.2 No Robust Space for Ramifications

The fact that an explanans contributes to the generation of the explanandum guarantees a relevance relation between the antecedent and the consequent of a countermathematical explanation. BCR examine this relevance relation in terms of ramifications of mathematical twiddles through a morphism fixed between the mathematical structure occurring in the explanans and the physical or mathematical structure in the explanandum. In this section, I argue that steps (iii′) or (iii″)—determining the influence of the varied explanans on the consequent of a countermathematical—are susceptible to two issues: sometimes following the ramification procedure does not deliver the truth value of a countermathematical one might intuitively expect, and sometimes the delivered truth value is trivial. I explain these two points in the remainder of this section.

Recall step (iii′) for evaluating a countermathematical conditional (Baron et al., 2017, p. 2): '... consider the downstream implications for the facts that we are not holding fixed of letting the antecedent vary: we see how the twiddle 'ramifies' through these facts'. In the case of cicadas, the fixed mathematics is the mathematical structure of natural numbers, and the relevant physical structure is time measured in years (Baron et al., 2017, pp. 10–11). BCR claim that, as a result of considering the downstream ramifications of twiddled mathematical facts, a conditional such as (V) should be recovered false.

(V) If, in addition to 13 and 1, 13 had the factors 19 and 23, North American periodical cicadas would not have 13-year life cycles.

(V) is false according to BCR for the following reason:

> Hold fixed the morphism. Now make the counterfactual change to the mathematics. The world keeps up its end of the bargain, and so a 13-year lifespan is now divisible into 19- and 23-year intervals. The cicadas don't budge: 13 remains the optimal strategy for avoiding predation by organisms that have life cycles up to 18 years. Of course, if there are 19- or 23-year predators, then 13 is no longer optimal. However, there are ecological constraints on the cicada case that rule out these predators.

As BCR (2017) point out, if there are 19- or 23-year predators, then 13 is no longer optimal. This means that twiddling the primeness of 13, irrespective of empirical constraints, should make the countermathematical (V) true. After all, they aim to deliver an evaluative procedure for finding the correct truth value of countermathematical explanations through the dependency between the countermathematical antecedent and its consequent. BCR think that (V) comes out false as a result of the ecological constraint. Given that the ecological constraint is itself an empirical explanans, and hence external to finding an explanatory dependence between the antecedent and the consequent of (V), the truth value of the countermathematical—being false—is not really determined by tracing the ramifications of the varied antecedent on the consequent of the countermathematical. The falsity of (V) is rather determined by supposing that an ecological constraint that holds in the actual world also holds in the relevant countermathematical world(s), where 13 is divisible by 19 and 23.

However, suppose there is a world where 13 is divisible by 19 and 23 and there are 13-year predators. Well then the ecological constraint could be plausibly different: it could be that the constraint extends to 13 which is at least as great as 23. After all, if cicadas can become 13 years old, then in virtue of the fixed morphism between the mathematical and empirical structure, they can also become at least 23 years old. So BCR have a choice to make here. Either they do not allow that the varied antecedent ramifies via the fixed morphism to the physical structure, and so determines the truth value of the countermathematical. But then they owe us an answer why the ramification I have just presented is invalid for establishing the truth value of a countermathematical. Or else they need to admit that (V) comes out true. But then varying 13's primeness in this way mathematically explains why cicadas have 13-year life cycles. And this is what BCR explicitly deny.

We have just seen a case in which the expected truth value of a countermathematical is not obtained by merely evaluating the countermathematical given the fixed morphism between mathematical structures in its antecedent and physical or mathematical structures in its consequent. Rather, it is obtained by what an empirical constraint dictates. This shows that, sometimes, the last step of the evaluative procedure of countermathematicals does not deliver the truth value one might intuitively expect.

Moreover, consider the following countermathematical:

(VI) If 13 had only the factor 1, North American periodical cicadas would not have 13-year life cycles.

If 13 had only the factor 1, 13 would not be prime. This variation of 13's primeness does not change its optimality (if anything it makes it even more optimal). After all, less factors of the number representing the cicada's life cycles result in more optimality. Hence, under this variation of 13's primeness, the cicadas would have 13-year life cycles, and so (VI) comes out false. This shows that varying the primeness of 13 may not explain the 13-year life cycles.

Now, let us turn to an example in which the delivered truth values are trivial, simply because there is no robust space for exploring the ramification of the counterfactual twiddle. In the case of intra-mathematical explanations, the evaluation of countermathematicals may result in obtaining trivial truth values. Recall countermathematical (III) If it were not the case that for any three consecutive nonzero natural numbers, at least one of those numbers is even (and therefore divisible by 2), then it would not be the case that the product of any three non- zero, consecutive natural numbers is divisible by 6. For the purpose of the argument, let us assume that 503, 504, and 505 are not even. How does this tweak ramify to the consequent? BCR's (2020) strategy is to rely on the following fact: the product of two natural numbers is even, only if at least one of them is. Under the twiddle, 505 is not even. So, for $(503 \times 504) \times 505$, we need to see whether (503×504) is even. However, neither of 503 or 504 is even by the twiddle.

Therefore, the product of 503, 504, and 505 is not divisible by 6. The problem here is that the antecedent and the consequent are not really mathematically distinct. The assumptions about the properties of numbers in the antecedent appear right away in the consequent. This makes the step (iii'') of the evaluation recipe of countermathematicals, for determining the influence of varying the antecedent on the consequent, idle. There is no space left to track down the ramifications running from the mathematical twiddle to the consequent, because the main components of the consequent trivially change as soon as the assumptions about the evenness of the numbers change in the antecedent. To see this, there is no intermediate ramification step between the antecedent and the consequent. There is simply no space for any ramification. The twiddle that all of 503, 504, and 505 are not even immediately affects the consequent. But this effect is too immediate to count as a ramification.[11]

3.3 No Explanatory Benefits

What are the *explanatory* benefits of varying mathematical facts? How would the world look like if $2 + 2$ were not equal to 4? Well, the honest answer is that we just do not know. This question provides no insight. Similarly, under BCR's (2017, 2020) assumptions, the variation of mathematical facts is uninformative. How the impossible 'perturbation' to the antecedent of a countermathematical is supposed to influence its consequent is fully left to the reader's intuitions, and these intuitions can be deeply fallible. The countermathematical (II) If 13 were not a prime number, then North American periodical cicadas would not have 13-year life cycles, for instance, does not provide any insight on its own. What we need in the case of countermathematical explanations is some sort of explanatory benefit. After all, the whole project of extending the counterfactual account of causal explanations is the exploration of the idea that the explanatory dependence between the mathematical explanans and the explanandum of any mathematical explanation is analyzable in terms of counterfactual dependence.

What makes us understand the explanation of the cicada example is that a life cycle is optimal only when it minimizes the overlap with the periodical predators; and this is given within the ecological constraints, by necessity, only when the life cycle is prime-numbered. I will say more about the significance of explanation by necessary constraints as an approach to mathematical explanations in the next section (Lange, 2013, 2016).

To be more precise, recall the optimality model of Sect. 3: (a) 13 and 17 are prime, (b) prime numbers maximize their lowest common multiple relative to all lower numbers; that is, they minimize the intersection of periods, and (c) the ecological constraints entail that prime-numbered life cycles minimize the chance of co-occurrence with predators that have similar life-cycle lengths. Given the ecological

[11] The situation is entirely different in evaluating a non-explanatory countermathematical such as 'If Hobbes had (secretly) squared the circle, all sick children in the mountains of South America at the time would have cared'. Here, there is no such shared structure at work between the antecedent and the consequent.

restriction of the life span of cicadas to 12–18 years, and the explanatory assumption of evolutionary biology that successful organisms must have evolved in an optimal way, there is just no other possibility left than 13-year and 17-year life cycles. In brief, if we respect scientific practice, 13's primeness explains the life-cycle length of certain cicadas because, if cicadas have evolved in an optimal way, *it cannot be otherwise*. The claim that the primeness of 13 and 17 is explanatory thus derives from the fact that the optimal numbers *must be* prime (provided the interval is restricted to between 12 and 18). The claim that primeness is explanatory does not derive from a counterfactual, or better countermathematical, variation of primeness and its propagated influence. As shown, it is extremely difficult (if not practically impossible) to analyze it on such grounds.

BCR assume that the primeness of 13 explains that North American periodical cicadas have 13-year life cycles. Hence, the countermathematical (II) must come out true on their account: If 13 were not prime, North American periodical cicadas would not have 13-year life cycles. Here is how they apply their account to this countermathematical. To vary the antecedent, BCR suggest using a new specifically designed operator, multiplication*, which takes 2 and 6 as input and outputs 13. This operator takes us to the closest world(s) to the actual world in which 13 is not prime, thanks to multiplication*. However, there are many other worlds in which 13 is not a prime number, but the non-primeness of 13 is obtained differently. For example, consider a world in which 13 is not prime because it only has the factor 1. Isn't this world closer to the actual world compared to a world in which 13 in addition to 1 and 13 has the factors 2 and 6? BCR (2017) do not provide any answer in their account for extra-mathematical explanations, and it is hard to see how a principled procedure would look like. If the criterion for choosing the closest world(s), as BCR (2020) suggest, is the world with the minimum changes to the intrinsic properties of the primeness of 13, the answer would be positive. The number of the violations to the intrinsic properties of the primeness of 13 is one if we go to a world in which 13 is only divisible by 1 (rather than being divisible by 1 and 13). The number of the violations to the intrinsic properties of 13 is two if we go to a world in which 13, in addition to 13 and 1, is also divisible by 2 and 6 (13 obtains two new factors). When we rely on a similarity order between worlds based on intrinsic properties, it seems that the multiplication*-world is less similar to the actual world than a world where 13 has only the factor 1. This poses the question why we should choose the multiplication*-world if it is not for 'finding' a presumed countermathematical dependence?

Here is a rough characterization of BCR's account in action. To be able to assess the applicability of their countermathematical account, they need to compare it to our intuitive background knowledge about what explains what for some examples. This means that, for the given examples, first we have an intuitive idea about the implication of varying the mathematical explanans. We also intuitively know what the truth value for a countermathematical will be, if an extension of a counterfactual theory of causal explanations is to be successful. Now, given this knowledge, BCR first set the desired result of the mathematical variation (for example, that 13 is not prime); second, they work backwards and pick some world(s) in which a specifically designed mathematical operator (such as multiplication*) is introduced; third, they consider the ramification of this variation; however, I suspect that this world is chosen in a way that it delivers the truth value for the countermathematical such that the desired countermathematical explanation becomes true. I suspect that if BCR wanted to obtain the opposite truth value for the countermathematical conditional (II), they could have designed another mathematical operation which takes us to another strange world which would serve their purpose.

To be clear, I am not claiming the BCR propose that every time we evaluate a countermathematical we must first start with a desired truth value for the countermathematical and then make twiddles until it has that truth value. However, their choice of the relevant impossible world(s) for the examples for which we intuitively know what the countermathematical truth value should be seems rather ad hoc. I suspect that the closest world with multiplication* is chosen such that it delivers a desired truth value that we expect for this example. Why should we not consider, for instance, the impossible world in which 13 has only the factor 1? In such a world the countermathematical could easily turn out to be false.

This procedure stands in contrast to the epistemic or practical benefits we acquire from counterfactual causal explanations. There, we have an epistemic space for exploration of the consequences of Suzy not throwing the rock. We are not bound to a similar non-explanatory procedure in this explorative space such that the agent engaged in the counterfactual analysis can play around with different situations in which the window does not shatter, and thereby attains understanding.

Epistemic benefits such as exploring epistemic space for the purpose of understanding are not the only explanatory benefit that we expect to acquire from a counterfactual analysis of explanation. In the cases of counterfactual causal explanations, variation of an empirical antecedent can provide some instrumentalist insights that are frequently used to deliberate, to predict, or to control outcomes of the variation. We can control and deliberate on some nearly perfect duplicates of empirical facts, for instance, by running agent-based simulations on similar scenarios, or more abstractly by some thought experiments that are set in the context of causal explanations. To vary mathematical facts does not allow us to do this. Entertaining a mathematical impossibility such as the non-primeness of 13, viz. a varied 'mathematical fact', does not provide us with such potential benefits.

This problem of no explanatory benefit, in particular, questions whether the steps (iii') and (iii'') of the evaluation recipe for countermathematicals, namely the determination of the influence of the variation of antecedent on the consequent, are attainable.

So far, I have proposed three challenges to the current proposals for understanding mathematical explanations based on counterfactuals. But what should we make of mathematical explanations? I suggest an easy answer for extra-mathematical explanations in the next section. An answer for intra-mathematical explanations is more intricate and must be given elsewhere.

Journal : SmallExtended 10670 | Article No : 466 | Pages : 24 | MS Code : 466 | Dispatch : 2-11-2021

4 Towards an Alternative Approach

One promising account for analyzing extra-mathematical explanations is a hybrid, integrated account of causal counterfactuals and constraint-based explanations. Recall that extra-mathematical explanations include some mathematical facts and some empirical facts in the collection of their explanantia. The empirical facts can be twiddled according to causal counterfactual accounts. On the other hand, as Lange (2013, 2016) defends this in relation to various examples, the mathematical facts can be taken as necessary constraints, having a modal force stronger than the laws of nature. These facts dictate what can be and cannot be otherwise. For instance, 13's primeness explains the life-cycle length of certain cicadas because, if cicadas have evolved in an optimal way (and given all empirical facts driven from evolutionary biology), *it cannot be otherwise*.

On standard counterfactual accounts of causal explanation, we can vary an empirical fact while keeping mathematical facts fixed. Consider the counterfactual:

(VII) If ecological constraints restricted the life-cycle length of cicadas to 14–16 years, the cicadas would not have 13-year life cycles.

Here, the empirical fact about the ecological constraints is varied. This possibility gives us what BCR hope to achieve by varying 13's primeness. For instance, if the possible fact that the cicadas have 14-year life cycles were true, they would overlap with predators having 1-, 2-, 7-, and 14-year life cycles. In this case, the cicadas would not avoid predators optimally. The antecedent of (VII) suggests an empirical variation from the actual biological constraint to a merely possible biological constraint that does the job in counterfactual thinking. Variation of the empirical fact, viz. that ecological constraints restrict the life-cycle length of cicadas to 14–16 years, can be done while keeping the mathematical facts fixed: that 13 and 17 are prime numbers. Hence, the empirical variation can be done by assuming that they would have, for instance, 14-year, 15-year, or 16-year life cycles. But this is very different from assuming that 13 were 12, or the like.

In light of the counterfactual explanation of (VII), it appears questionable whether the counterfactual account of causal explanations needs any extension to accommodate extra-mathematical explanations. BCR aim to explain certain empirical facts by varying mathematical facts. However, their enterprise to cover extra-mathematical explanations is redundant if the empirical facts can be counterfactually explained without varying mathematical facts. On this picture, mathematical explanantia just play a constraining role.[12]

Acquiring a general account for analyzing intra-mathematical explanations, however, is more challenging, and it seems to me to remain an open question. In the literature, some general accounts for accommodating (some) intra-mathematical explanations are already offered. Two prominent accounts are Steiner's (1978) explanatory proofs and Kitcher's (1989) explanatory unification. The scope and

12FL01 [12] For details about the modal characteristics of mathematical facts as compared to empirical facts, see
12FL02 Lange (2016).

Journal : SmallExtended 10670	Article No : 466	Pages : 24	MS Code : 466	Dispatch : 2-11-2021

validity of these accounts remains a matter of criticism. For instance, drawing on mathematical practice, Lehet (2021) and D'Alessandro (2020) discuss examples of intra-mathematical explanations that go beyond explanatory proofs. In addition, Mancosu and Hafner (2008) show that Kitcher's model makes predictions about explanatoriness that go against specific cases in mathematical practice.

In my view, a more promising proposal is a 'bottom up' approach (Mancosu, 2008) which requires investigating several case studies that are deemed explanatory in mathematical practice. When enough case studies across various areas of mathematics are done, we might be able to provide a general account for intra-mathematical explanations. This remains a task to be done elsewhere.

5 Conclusion

Recently, several philosophers such as Reutlinger (2016), Baron et al. (2017), Woodward (2018), Baron et al. (2020), and Reutlinger et al. (2020) have attempted to extend the counterfactual theory of causal explanations to mathematical explanation. These attempts have had resounding impacts on theorizing about scientific explanation, metaphysical explanation, metaphysical causation, and logical explanation. According to these attempts, roughly, we can apply a standard way of thinking about causal counterfactuals to countermathematicals.

Among these, Baron et al. (2017, 2020) offer the most elaborate and influential endeavor. Some defenders of countermathematical explanations simply presuppose the validity of BCR's account. For example, Reutlinger et al. (2020) argue for a necessary condition—called a dependency condition—common to counterfactual theories of explanation. This condition states that 'The explanandum counterfactually depends on certain possible changes in the conditions described by the explanans (i.e. if the explanans conditions were different, then the explanandum would be different as well)'. How to evaluate such countermathematical conditionals? Reutlinger et al. (2020) take the semantic procedure for the evaluation of countermathematical explanations as proposed by Baron et al. (2020) as a premise of their account. Others require arguments very much along the lines of BCR to defend their account of countermathematical explanations. For instance, a difference-making account of countermathematical explanation such as one along the lines of Woodward (2018) requires answering to what-if-things-had-been-different questions in the sense that if the conditions in the explanans had been different, what the explanandum expresses would have been different. If these conditions are purely mathematical (in the case of intra-mathematical explanations), then we run into very similar issues as the ones that trouble BCR.

By providing a detailed criticism of BCR's account, I have argued against those contemporary attempts which claim that the features of the counterfactual account of causal explanation carry over to mathematical explanations.

I have discussed three main problems which pertain to the current counterfactual theories of mathematical explanations. In light of these problems, I have shown that the steps of the common recipe for evaluating countermathematical explanations are not satisfactory enough. As a result, we do not have—as of yet—a plausible counterfactual theory for mathematical explanations. I agree that some theoretical virtues might dictate searching for a general theory of explanation. However, I have shown that the current accounts fail in providing a satisfactory counterfactual account of mathematical explanations. Only a significant modification of the current accounts might rescue the search for a fully general counterfactual theory of explanation. Hence, without resolving the issues raised in this paper, the current proposals ought to be rejected. I have also claimed that a hybrid account integrating the virtues of the causal counterfactual approach and constraint-based approach to mathematical explanations can accommodate extra-mathematical explanations. A general account for intra-mathematical explanations, however, remains to be developed.

Finally, I would like to point to a potential approach which would resist (some of) the criticisms I raised in this paper. This approach requires to identify propositions with sets of possibilities, and then interpret possibilities not in a mathematical way, but in terms of what *you* consider to be possible or impossible (Huber, 2021, Ch. 6). This interpretation remains open to full investigation, and I do not tackle it in details here. If so, any exploration of this account, however, requires to justify the ways in which an individual's belief about the truth or falsity of a countermathematical statement gives rise to countermathematical explanations. It seems (to me) that explanations require more than mere belief: they require truth.

Acknowledgements I am grateful to Jim Brown, Franz Huber, Mario Günther, Alan Hájek, Mike Miller, Michael Strevens, and two anonymous referees for *Erkenntnis* for critical comments and feedback on earlier drafts of this paper. I also thank Koji Tanaka, Alan Baker, audiences at the Workshop on Metaphysical Explanation in Science in Birmingham, participants in the workshop on Non-Causal Explanations: Logical, Linguistic and Philosophical Perspectives in Ghent, and audiences at the American Philosophical Association Eastern Division for valuable discussions and suggestions.

References

Andreas, H., & Günther, M. (2021). Regularity and inferential theories of causation. In E. N. Zalta & Metaphysics Research Lab (Eds.), *The Stanford encyclopedia of philosophy*. Stanford University (**Fall 2021 edn**).

Baker, A. (2005). Are there genuine mathematical explanations of physical phenomena? *Mind, 114*(454), 223–238.

Baron, S. (2020). Counterfactual scheming. *Mind*, *129*(514), 535–562.

Baron, S., Colyvan, M., & Ripley, D. (2017). How mathematics can make a difference. *Philosopher's Imprint, 17*(3), 1–29.

Baron, S., Colyvan, M., & Ripley, D. (2020). A counterfactual approach to explanation in mathematics. *Philosophia Mathematica, 28*(1), 1–34.

Batens, D. (1990). Against global paraconsistency. *Studies in Soviet Thought, 39*(3–4), 209–229.

Berkeley, G. (1734). The analyst: A discourse addressed to an infidel mathematician [online]. http://www.maths.tcd.ie/pub/HistMath/People/Berkeley/Analyst.

Berto, F., French, R., Priest, G., & Ripley, D. (2017). Williamson on counterpossibles. *Journal of Philosophical Logic, 47*(4), 693–713.

Colyvan, M. (2012). *An introduction to the philosophy of mathematics*. Cambridge University Press.

Colyvan, M., Cusbert, J., & McQueen, K. (2018). Two flavours of mathematical explanation. In A. Reutlinger & J. Saatsi (Eds.), *Explanation beyond causation* (pp. 231–249). Oxford University Press.

Da Costa, N. C. (1997). *Logiques classiques et non classiques: Essai sur les fondements de la logique*. Masson.

D'Alessandro, W. (2020). Mathematical explanation beyond explanatory proof. *The British Journal for the Philosophy of Science, 71*(2), 581–603.

Halpern, J. Y., & Pearl, J. (2005). Causes and explanations: A structural-model approach. Part I: Causes. *The British Journal for the Philosophy of Science, 56*(4), 843–887.

Huber, F. (2021). *Belief and counterfactuals: A study in means-end philosophy.* Oxford University Press.

Kasirzadeh, A. and Smart, A. (2021). The use and misuse of counterfactuals in ethical machine learning. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 228–236.

Kitcher, P. (1989). Explanatory unification and the causal structure of the world. In P. Kitcher & W. Salmon (Eds.), *Scientific explanation, (Minnesota Studies in the Philosophy of Science, Volume XIII)* (pp. 410–505). University of Minnesota Press.

Lange, M. (2013). What makes a scientific explanation distinctively mathematical? *The British Journal for the Philosophy of Science, 64*(3), 485–511.

Lange, M. (2014). Aspects of mathematical explanation: Symmetry, unity, and salience. *Philosophical Review, 123*(4), 485–531.

Lange, M. (2016). *Because without cause*. Oxford University Press.

Lange, M. (2018). Mathematical explanations that are not proofs. *Erkenntnis, 83*(6), 1285–1302.

Lehet, E. (2021). Induction and explanatory definitions in mathematics. *Synthese, 198*(2), 1161–1175.

Lewis, D. (1973). Causation. *The Journal of Philosophy, 70*(17), 556–567.

Lewis, D. (1973). *Counterfactuals*. Blackwell Publishing.

Lewis, D. (1983). New work for a theory of universals. *Australasian Journal of Philosophy, 61*(4), 343–377.

Lewis, D. (1986). Causal explanation. *Philosophical papers* (Vol. II, pp. 214–240). Oxford University Press.

Lyon, A., & Colyvan, M. (2008). The explanatory power of phase spaces. *Philosophia Mathematica, 16*(2), 227–243.

Mancosu, P. (2008). Mathematical explanation: Why it matters. In P. Mancosu (Ed.), *The philosophy of mathematical practice* (pp. 130–150). Oxford University Press.

Mancosu, P., & Hafner, J. (2008). Beyond unification. In P. Mancosu (Ed.), *The philosophy of mathematical practice* (pp. 151–178). Oxford University Press.

Maurin, A.-S. (2019). Grounding and metaphysical explanation: it's complicated. *Philosophical Studies, 176*(6), 1573–1594.

McCullough-Benner, C. (2020). Representing the world with inconsistent mathematics. *The British Journal for the Philosophy of Science, 71*(4), 1331–1358.

Meheus, J. (2003). Paraconsistent compatibility. *Logique et Analyse*, 251–287.

Nickel, B. (2010). How general do theories of explanation need to be? *Noûs, 44*(2), 305–328.

Nolan, D. (1997). Impossible worlds: A modest approach. *Notre Dame Journal of Formal Logic, 38*(4), 535–572.

Priest, G. (2002). *Paraconsistent logic* (pp. 287–393). Springer.

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., Colyvan, M., and Krzyzanowska, K. (2020). The prospects for a monist theory of non-causal explanation in science and mathematics. *Erkenntnis*, 1–21.

Schaffer, J. (2016). Grounding in the image of causation. *Philosophical Studies, 173*(1), 49–100.

Steiner, M. (1978). Mathematical explanation. *Philosophical Studies, 34*(2), 135–151.

Tanaka, K. (2003). Three schools of paraconsistency. *The Australasian Journal of Logic, 1,*, 28–42.

Williamson, T. (2018). Counterpossibles. *Topoi, 37*(3), 357–368.

Wilson, A. (2018). Grounding entails counterpossible non-triviality. *Philosophy and phenomenological research, 96*(3), 716–728.

Wilson, A. (2018). Metaphysical causation. *Noûs, 52*(4), 723–751.

Woodward, J. (2018). Some Varieties of Non-Causal explanations. In A. Reutlinger & J. Saatsi (Eds.), *Explanation beyond causation* (pp. 117–137). Oxford University Press.