Explaining the Reliability of our Mathematical Beliefs

Phil 489 – Shakked Noy

Supervisor: Richard Joyce

Introduction

Originally formulated by Paul Benacerraf (1973) and developed by Hartry Field (1989), the Benacerraf-Field Problem highlights an important tension in the claims of mathematical realists. On the one hand, mathematical realists assert that our mathematical beliefs accurately track the (mind-independent) properties of abstract mathematical entities. On the other hand, mathematical realists concede that we acquire our mathematical beliefs without interacting with or observing those entities. Intuitively, this concession undermines the assertion that our mathematical beliefs reliably track the mathematical truth.

Although it was originally formulated as a challenge to mathematical platonism, the Benacerraf-Field Problem actually applies more broadly to any theory that claims that our mathematical beliefs accurately track mind-independent abstract facts. This includes paradigmatically anti-realist positions in the philosophy of mathematics, like modal structuralism (which translates mathematical statements into statements of modal logic[1]) or even mathematical fictionalism (which claims that mathematical knowledge consists of logical knowledge about which mathematical results follow from which axioms[2]). These theories, too, fall prey to the problem of explaining how we can acquire reliable knowledge of abstract modal or logical truths with which we can't interact.

In fact, the scope of the Benacerraf-Field Problem is even wider than that. As Clarke-Doane (2016a) argues, analogues of the Benacerraf-Field Problem threaten most of our logical, modal, and moral beliefs in addition to our mathematical beliefs. Any philosophical theory which insists that our

[1] See Hellman (1989).
[2] See Field (1989).

beliefs accurately reflect mind-independent abstract facts is vulnerable to a version of the Benacerraf-Field Problem.

My reference above to our beliefs "accurately tracking" the mathematical facts should also make it clear that the Benacerraf-Field Problem is intimately connected to genealogical debunking arguments in metaethics. Benacerraf and Field challenge mathematical realists to explain why our mathematical beliefs are reliable, given that they're acquired independently of any interaction with mathematical entities. Similarly, genealogical debunkers like Street (2006) and Joyce (2006) challenge moral realists to explain why our moral beliefs accurately track moral reality, given that our moral beliefs are the output of evolutionary, psychological, and social processes that have nothing to do with the moral facts.

Indeed, proponents of the Benacerraf-Field Problem often use language very similar to the language used by genealogical debunkers. Field (2005) writes:

> […] our belief in a theory should be undermined if the theory requires that it would be a huge coincidence if what we believed about its subject matter were correct […] it seems hard to give any account of our beliefs about these mathematical objects that doesn't make the correctness of the beliefs a huge coincidence.

Meanwhile, Street (2006) argues that:

> […] in the absence of an incredible coincidence, most of our evaluative judgements are likely to be false.

So an appropriately generalised version of the Benacerraf-Field Problem could threaten a wide array of our beliefs. Moreover, other challenges to the epistemic status of our beliefs, like genealogical debunking arguments, point to the same fundamental issue as the Benacerraf-Field Problem. As Clarke-Doane (2016a) writes,

> An important class of influential but *prima facie* independent epistemological problems are, in relevant respects, restatements of the Benacerraf Problem.

It's therefore of interest whether the Benacerraf-Field Problem can be overcome – whether we can find some way to reconcile the claim that our beliefs accurately track mind-independent abstract facts with the unfortunate reality that we can't interact with or observe those facts.

In this paper, I argue that the Benacerraf-Field Problem can be overcome in the mathematical case. I then present a general criterion for when a domain of beliefs escapes the Benacerraf-Field Problem. I (briefly) argue that our logical beliefs certainly satisfy this criterion, our modal beliefs plausibly do, and our moral beliefs probably do not. I conclude that our mathematical, logical, and possibly modal beliefs are insulated from the Benacerraf-Field Problem, but our moral beliefs are not.

In brief form, my argument is as follows. The Benacerraf-Field Problem can be viewed as a challenge to demonstrate the *counterfactual sensitivity* of our mathematical beliefs. Our beliefs about a domain D are sensitive if, had the D-truths been different and we formed our beliefs through the same process, our D-beliefs would be correspondingly different. The key insight of the Benacerraf-Field Problem is that, because our mathematical beliefs are formed independently of interaction with the mathematical facts, it's difficult to see how our mathematical beliefs could be sensitive to counterfactual variations in the mathematical facts.

So in order to escape the Benacerraf-Field Problem, we need to show that our mathematical beliefs are sensitive. Given that our mathematical beliefs (like all of our beliefs) are the output of purely physical processes, demonstrating the sensitivity of our mathematical beliefs requires demonstrating that, *had the mathematical facts been different, the physical facts would have been different in a way that produced correspondingly different mathematical beliefs.* And although it's generally accepted that the mathematical facts are causally inefficacious, it *is* true that, had the mathematical facts been different, the physical world would in many ways be correspondingly different.

Consider, for example, a physical system whose behaviour can be modelled by a set of differential equations, with the solution to those equations predicting the equilibrium behaviour of the system. If the solution to those equations were different, the equilibrium behaviour of the physical system would have to be correspondingly different. Situations like this abound in physics, biology, economics, and other descriptive sciences. Therefore, if the mathematical facts were different, the physical world would be systematically different.

Moreover, our mathematical beliefs are informed by our interactions with the physical world. Mathematics has often developed in tandem with science; mathematicians frequently use spatial intuitions which have been honed on the physical world; and the intended models of many mathematical structures are designed to resemble the physical world.

So if the mathematical facts were different, the physical facts would be correspondingly different. And since our mathematical beliefs are informed by the physical world, if the physical facts were different, our mathematical beliefs would be correspondingly different. It follows that our mathematical beliefs are sensitive, not *in spite of* but *because of* genealogical considerations.

The reason that our mathematical beliefs are sensitive is that the mathematical facts *covary counterfactually* with the physical facts. We know that the mathematical facts covary counterfactually with the physical ones because of the special role that mathematical facts play in scientific explanations of physical phenomena. More generally, our beliefs about a domain D escape the Benacerraf-Field Problem if our D-beliefs play a role in scientific explanations that is relevantly similar to the role played by mathematical facts. Our logical beliefs certainly play such a role; our modal beliefs arguably do; and I believe that our moral beliefs do not.

Importantly, my argument is *not* a traditional indispensability argument in the style of Putnam (1979) or Colyvan (2001). While indispensability arguments for mathematical realism rest on the very strong claim that mathematical facts play an *explanatorily indispensable* role in scientific explanations, my argument relies on much weaker assumptions, as I explain in Section 3.

The rest of this paper proceeds as follows. In Sections 1 and 2, I clarify exactly how the Benacerraf-Field Problem is meant to undermine our mathematical beliefs. In Sections 3-5, I consider the two most popular accounts of the role of mathematical facts in scientific explanations. I argue that both of these accounts commit us to the claim that, if the mathematical facts were different, the physical facts would have to be correspondingly different. Next, in Section 6, I discuss the genealogy of our mathematical beliefs. I combine my genealogical claims with my conclusions from Sections 3-5 to demonstrate that our mathematical beliefs are sensitive, and therefore escape the Benacerraf-Field Problem. In Section 7, I show that my arguments are robust to alternative assumptions. Subsequently, in Section 8 I consider whether analogous arguments could successfully protect our logical, modal, and moral beliefs. I argue that the answers are “yes,” “possibly,” and “no,” respectively. Finally, I conclude.

1 Reliability, Truth-Tracking, and the Modal Account

Which epistemic property of our mathematical beliefs is the Benacerraf-Field Problem meant to threaten? Field (1989) argues that

> […] we should view with suspicion any claim to know facts about a certain domain if we believe it impossible in principle to explain the reliability of our beliefs about that domain.

Meanwhile, Joyce (2016a) writes that

> […] what is really at issue [in genealogical debunking arguments] is whether moral beliefs are the result of a truth-tracking process.

Both of these statements are vague. In Field's case, it's unclear what it means for our beliefs to be "reliable," and it's unclear what would constitute an explanation of the reliability of our mathematical beliefs. In Joyce's case, it's unclear what it means for our beliefs to "track the truth." But Field and Joyce both seem to be gesturing at the same basic property of our beliefs: the property of being "appropriately responsive to" the facts with which those beliefs are concerned.[3]

So what does this property of "appropriate responsiveness" consist in? One answer, suggested by Clarke-Doane (2016a), is that our D-beliefs are reliable (or track the truth) iff they are both *sensitive* and *safe*. Our D-beliefs are sensitive if, in the closest possible world where the D-truths are different and we use the same method to acquire our D-beliefs, our D-beliefs are correspondingly different. Intuitively, our D-beliefs are sensitive if they are "counterfactually robust." By contrast, our D-beliefs are safe if they could not easily have been false. Call this account of reliability the "Modal Account."

The Modal Account runs into an immediate problem when applied to beliefs whose contents are alleged to be metaphysically necessary, like our mathematical or moral beliefs. Suppose I decide to form a belief as follows: if water is identical to H2O, I will believe that 9,161 is prime. If water is not identical to H2O, I will believe that 9,161 is not prime. Then my belief about 9,161's primeness is both safe and vacuously sensitive. It is safe because

(a) the metaphysically necessary identity between water and H2O guarantees that my belief about 9,161's primeness could not easily have been different, given that I use the same method to acquire it, and
(b) 9,161 could not easily have failed to be prime, so

[3] From now on, I'll assume that "reliability" and "truth-tracking" are referring to the same property, and I'll use the terms interchangeably.

(c) my belief about 9,161's primeness could not easily have been false.

And my belief is vacuously sensitive because the conditional "if it were false that 9,161 is prime, my belief would be correspondingly different" is vacuously true, since its antecedent is necessarily false (because the mathematical facts are generally assumed to be metaphysically necessary). Yet despite being both sensitive and safe, my belief that 9,161 is prime is clearly not reliable, nor was it formed in a way that tracks the truth.

To rescue the Reliability Account, we can claim, as Clarke-Doane (2012) suggests, that we should analyse the sensitivity of our beliefs in terms of *conceptual* possibility rather than *metaphysical* possibility. It is, we can claim, *conceptually* possible for the mathematical or moral truths to be different. If we accept this claim, then the conditional "if it were false that 9,161 is prime, my belief would be correspondingly different," is non-vacuously false, so my belief that 9,161 is prime is insensitive, and the Modal Account correctly diagnoses it as unreliable.

The introduction of conceptual possibility is not just an *ad hoc* attempt to shield the Modal Account against counterexamples: there are strong independent reasons to believe that conceptual possibility is the right modality to use. I assume we all feel intuitively that my belief that 9,161 is prime is unreliable, if formed in the manner described. If asked to explain why we feel this way, we might say that "My belief is unreliable because *even if 9,161 were not prime, I would still believe that it was*." This is a useful and illuminating explanation, even though it invokes a metaphysically impossible counterfactual. Clearly, thinking about conceptually possible counterfactuals is a helpful way to explicate the meaning of "reliability." Moreover, as Field (1989) and Clarke-Doane (2017) note, we can often non-vacuously condition on metaphysical impossibilities, especially when it comes to mathematics.

But this move to conceptual possibility creates a difficulty for the Benacerraf-Field Problem. Consider our common-sense object beliefs: beliefs about the physical conditions under which object properties (like "being a rock" or "being a restaurant") are instantiated.

Clarke-Doane (2015, 2016a, 2016b) argues that proponents of the Benacerraf-Field Problem (and genealogical debunkers) face a dilemma. If metaphysical possibility is the right modality to use, then our mathematical beliefs are vacuously sensitive, since the mathematical facts could not possibly be different. But if conceptual possibility is the right modality, then our object beliefs fail to be sensitive. As Clarke-Doane (2015) writes:

> While it may be metaphysically necessary that the conditions under which the property of being a rock is instantiated are what they are, it seems that they could have been different "as a purely conceptual matter." That they are is just what some ontologists allege. These ontologists allege that particles arranged rockwise fail to compose a rock. But had—for all that we can intelligibly imagine—this been the case, our rock beliefs would have been the same.

If Clarke-Doane is correct, then if we use conceptual possibility to evaluate the sensitivity of our beliefs, our object beliefs turn out to be insensitive. Clarke-Doane concludes that there is no formulation of the Benacerraf-Field Problem in terms of sensitivity that threatens our mathematical beliefs without also threatening many of our epistemically uncontroversial beliefs.

I will note two things about Clarke-Doane's argument. The first is that if we are conventionalists about the classification of ordinary objects,[4] the argument fails. Conventionalists claim that object truths, unlike mathematical or moral truths, are mind-dependent – the facts, for example, about what constitutes a "table" just depend on human conventions. If conventionalists are right, then the

[4] In the style of Sidelle's (1989) conventionalism about natural kinds.

closest possible world in which the object truths are different is simply one in which human conventions are different, and therefore one in which our beliefs about object truths are (presumably) correspondingly different. If this is true, then our common-sense object beliefs are sensitive, even if we use the modality of conceptual possibility.[5]

But suppose we reject conventionalism. What Clarke-Doane fails to recognize is that rather than defeating the Benacerraf-Field Problem, his argument just compromises the Modal Account (and the principle of Modal Security which he espouses[6]). Obviously, there is an intelligible sense in which our beliefs can be reliable or unreliable. And the right account of reliability will judge that my belief that 9,161 is prime is unreliable (if formed in the manner described), and that our common-sense object beliefs are reliable. Yet according to Clarke-Doane, the Modal Account cannot consistently produce both of those judgements. If Clarke-Doane is right, the Modal Account clearly fails as an account of reliability, and we ought to accept some other account instead.

I am inclined to accept conventionalism, so I don't think Clarke-Doane's argument is fatal to the Modal Account. But for those who reject conventionalism, it is worth emphasising that my arguments in this paper are not contingent on the truth of the Modal Account; I phrase my arguments in terms of the modal notions of sensitivity and safety purely for ease of exposition. As I

[5] In Clarke-Doane & Baras (forthcoming), Clarke-Doane restates this argument but with our explanatorily basic phenomenal beliefs instead of our object beliefs. Conventionalism about phenomenal experience is implausible. However, Clarke-Doane's argument presumes controversially that the laws governing the instantiation of phenomenal properties from physical ones are metaphysically necessary, which is a claim that is denied by dualists like Chalmers (1996).

[6] If Clarke-Doane is right about our object beliefs, then the "primeness of 9,161" example constitutes a counterexample to Modal Security (perhaps with some additional details specifying that historical factors made it inevitable that I would choose my belief about 9,161's primeness in this way). Clarke-Doane seems to recognize the potential for this sort of counterexample when he writes that "it might be thought that Modal Security cannot handle necessary truths that we were "bound" to believe." (Clarke-Doane 2016a, p.31) But he then offers a "machine" counterexample which fails because it has a crucial weakness that is not shared by my "primeness of 9,161" counterexample.

explain in Section 7, the basic thrust of my argument is robust to alternative definitions of "reliability."

2 Genealogy and Reliability

If we accept the Modal Account, then the Benacerraf-Field Problem can threaten our mathematical beliefs only by threatening their sensitivity or safety. As Clarke-Doane (2016a, p.24) argues, there are strong reasons to believe that our mathematical beliefs are safe. So if the Benacerraf-Field Problem threatens our mathematical beliefs, it must do so by calling into question their sensitivity.

How might the Benacerraf-Field Problem threaten the sensitivity of our mathematical beliefs? The key observation underlying the Benacerraf-Field Problem is that we acquire our mathematical beliefs without interacting with mathematical entities. This is because our mathematical beliefs are formed through purely physical processes (evolutionary, psychological, perceptual, and sociological processes), while mathematical entities are nonphysical and therefore can't play a part in those processes.

Here is an argument which leverages this key observation to argue for the conclusion that our mathematical beliefs are insensitive:

(a) Our mathematical beliefs are the products of purely physical processes. (Genealogical Premise)

(b) If the mathematical truths were different, the physical facts would remain the same. (Modal Claim)

(c) Therefore, if the mathematical truths were different, those physical processes would remain the same.

(d) Therefore, if the mathematical truths were different, our mathematical beliefs would remain the same (i.e. our mathematical beliefs are insensitive). (Conclusion)

My contention in this paper is that this sensitivity-threatening argument is unsound, because premise (b) is false. In order to show that premise (b) is false, I need to show that the mathematical facts *covary counterfactually* with the physical facts: that is, I need to show that if the mathematical facts were different, the physical facts would have to be correspondingly different.

One way to argue for the counterfactual covariance of the mathematical and physical facts would be to appeal directly to our intuitions about (conceptually) possible worlds in which the mathematical facts are different. We could try to imagine specific variations in the mathematical facts, and ask whether they'd produce corresponding variations in the physical facts. While potentially useful, there are good reasons to be suspicious of this approach. First, it's not clear that we can have reliable intuitions about metaphysically impossible worlds. Secondly, even if we did have strong and trustworthy intuitions about these cases, it still wouldn't be clear what grounds or justifies those intuitions.

Instead, my strategy in Section 3 is to argue that our judgements about the *actual* relationship between mathematical facts and scientific explanations of physical phenomena *imply* that, if the mathematical facts were different, the physical facts would have to be correspondingly different. This approach avoids both of the problems I just mentioned. It shows that we are committed to believing in the counterfactual covariance of the mathematical and physical facts, without relying on our intuitions about metaphysically impossible worlds. Moreover (and this is a more speculative psychological claim), it demonstrates that any intuitions we *do* have about those metaphysically impossible worlds are grounded in our beliefs about the relationship between mathematics and science.

3 Mathematical Facts and Scientific Explanations

It is widely recognised that mathematics plays a special role in many scientific explanations, but the exact nature of this role is controversial. Two popular positions have emerged in the literature.

One position, which I will call "indispensabilism," is that mathematical facts play an indispensable explanatory role in scientific explanations. Proponents of indispensabilism include Colyvan (2001, 2010) and Baker (2005, 2009), who defend an "enhanced indispensability argument" for mathematical realism.

The other position, which I will call "easy-road nominalism,"[7] is that mathematical facts play a *useful,* but *non-explanatory*, role in scientific explanations. Proponents of easy-road nominalism include Pincock (2007), Daly & Langford (2009), Saatsi (2011), and Leng (2012).

There is a third, less popular, position on the role of mathematical facts in scientific explanation: "hard-road nominalism." Formulated most famously by Field (1980), hard-road nominalism claims that we could, in principle, formulate attractive non-mathematical versions of our scientific theories. I will not discuss hard-road nominalism, since it is much less popular than the other two positions. I believe, however, that the arguments I present in this section could be extended to cover hard-road nominalism as well.

In Sections 3.1 and 3.2, I argue that both indispensabilism and easy-road nominalism about mathematics commit us to the claim that the mathematical facts covary counterfactually with the physical ones. More precisely, they commit us to the claim that, if the conclusions of existing mathematical theorems were altered, the physical facts would have to be correspondingly different.

[7] Using the terminology coined by Colyvan (2010).

In Section 4, I show how this claim can be broadened to cover other types of variations in the mathematical facts.

3.1 Indispensabilism

Recent reformulations of the Quine-Putnam indispensability argument, called "enhanced indispensability arguments," rely on the claim that mathematical facts play an indispensable explanatory role in certain scientific explanations (Colyvan 2001). A mathematical fact plays an indispensable explanatory role in a scientific explanation if the mathematical fact does explanatory work which (a) is essential to the overall explanation, and (b) cannot be done by the physical facts alone.

It's clear, I think, that indispensabilism commits us to the counterfactual covariance of the mathematical and physical facts. Suppose that physical phenomenon P is *explained by* mathematical fact M. Then the existence of P must in some way *depend on* or *be a result of* the truth of mathematical fact M (otherwise, how would M explain it?). If this is true, then if mathematical fact M were different, the dependence of P on M guarantees that P would have to be correspondingly different.[8]

To illustrate this point more concretely, consider a paradigmatic example of the (purported) explanatory indispensability of mathematical facts. Originally discussed by Baker (2005), it concerns the lifecycles of periodical cicadas.

[8] This is the intuition underlying Sturgeon's (1997) "counterfactual test" for explanatory relevance. As Majors (2007) writes, "To say that one event is explanatorily relevant to another is intuitively to say that the former 'made a difference' with respect to the latter."

Every n years, depending on the region, adult periodical cicadas emerge *en masse* to mate, and then die shortly afterwards. Intriguingly, n is always a prime number. One explanation of this phenomenon goes as follows: it is evolutionarily advantageous for cicadas to have a life-cycle period that intersects as infrequently as possible with the life-cycle periods of periodical predator species. And a pair of number-theoretic results tell us that cycles with prime-numbered periods minimize the frequency of intersection with other periodic cycles. Baker's contention is that these number-theoretic results play an indispensable role in explaining why cicadas' life-cycle periods are always prime.

Here is Baker's (2005) explanation of why cicadas in a particular ecosystem have 17-year lifecycles, in the form of a deductive argument (premise 5 is my addition):

(1) Having a life-cycle period which minimizes intersection with other (nearby/lower) periods is evolutionarily advantageous [biological 'law']

(2) Prime periods minimize intersection (compared to non-prime periods) [number-theoretic theorem]

(3) Hence organisms with periodic life-cycles are likely to evolve periods that are prime ['mixed' biological / mathematical law]

(4) Cicadas in eco-system type, E, are limited by biological constraints to periods from 14 to 18 years [ecological constraint]

(5) 17 is the only prime number between 14 and 18

(6) Hence cicadas in eco-system type, E, are likely to evolve 17-year periods

The distinctive feature of Baker's explanation is his inclusion of the purely mathematical premise (2) (and, implicitly, (5)), which reflects his claim that mathematical facts play a key role in explaining why cicadas in eco-system type E have 17-year life-cycles. Now, suppose that the relevant mathematical facts were different – suppose, for example, that 16 was prime and 17 was not. The closest possible world in which this is true is one in which the non-mathematical assumptions about cicadas stated in premises (1) and (4) still hold.[9] By simply substituting all instances of “17” with “16” in Baker’s explanation, it follows deductively that cicadas in this counterfactual world are likely to evolve 16-year periods instead of 17-year ones. Varying the mathematical facts produces a corresponding variation in the physical facts.

Of course, not all scientific explanations are deductive like the one sketched above; some are inductive statistical explanations. But the same principle applies: as long as mathematical facts figure indispensably in the explanation of a phenomenon, varying the mathematical facts alters the predictions (whether probabilistic or deterministic) made by that explanation.

So if we accept an indispensabilist account of mathematical explanation in science, it follows that (certain) variations in the mathematical facts are necessarily accompanied by corresponding variations in the physical facts.

Note the method by which we arrived at this conclusion. In order to establish that the facts about prime numbers covary counterfactually with the facts about cicadas’ lifecycle periods, we didn’t need to rely on our intuitions about outlandish counterfactuals. That is, we didn’t need to imagine a possible world in which 16 was prime, and ask ourselves whether cicadas in this world must have 16-year lifecycles. Rather, we simply took what we believe to be *actually* true about the role of number-theoretic facts in explaining the lifecycle periods of cicadas, and demonstrated that this implies our desired modal conclusion. This strategy worked because our actual judgement about the scientific explanation has implicit modal connotations (judgements about “explanation” and “dependence” are implicitly modal).

[9] It might be objected that a world in which 16 was prime and 17 was not would be so systematically physically different from the actual world that cicadas might not exist at all. This is correct, but it concedes my desired conclusion.

3.2 Easy-road Nominalism

Most philosophers who reject mathematical indispensabilism accept some version of "easy-road nominalism" instead. Each version of easy-road nominalism has its own nuances, but all of them claim (more or less) that the role of mathematical facts in scientific explanations is to *index* or *represent* physical facts. According to easy-road nominalists, using mathematical facts to represent physical facts allows our scientific theories to be simpler, more rigorously justified, and more general. It may therefore be impossible to avoid using mathematics, if we want our scientific theories to possess those theoretical virtues. But the mathematical facts play no special *explanatory* role in scientific explanations: behind every mathematical explanation of a physical phenomenon lies a (longer, more complicated, and less general) physical explanation of that phenomenon.

Here is a representative sample of easy-road nominalist accounts:

- Pincock's (2007) "mapping account." Pincock argues that when we explain a physical phenomenon using mathematics, we "map" the relevant physical features of the situation onto a corresponding mathematical structure. Mapping to an abstract mathematical structure is useful because it makes the explanation robust to changes in the irrelevant microphysical properties of the situation.
- Leng's (2012) account is similar to Pincock's. Leng argues that (some) mathematical explanations are "structural explanations:" they gain their explanatory value by helping us abstract away from the microphysical features of the situation, to the general structural features that give rise to the phenomenon.
- Daly & Langford's (2009) and Melia's (2000) "indexing account." This account claims that scientific theories use mathematical facts to "index" physical facts, in order to achieve theoretical simplicity and economy of expression.

- Saatsi's (2011) account, which claims that mathematics can be used to represent different possible configurations of the physical facts, and can thereby aid or justify physical inferences.

Since easy-road nominalists claim that physical facts do the real explanatory work in scientific explanations, and the mathematical facts merely serve to index or represent them, it's not immediately obvious (if we accept an easy-road nominalist account) that varying the mathematical facts alters the predictions made by scientific explanations. Conceivably, if the mathematical facts were different, this would simply make them unsuitable for indexing/representing the physical facts, rather than changing any physical phenomena. Showing that easy-road nominalism commits us to the counterfactual covariance of the mathematical and physical facts therefore requires some additional work.

First, what does it mean to say that mathematical facts are used to "index" or "represent" physical facts? Leng (2012) writes:

> We can think of a mathematical structure as characterized by axioms. A physical system instantiating that structure is one where those axioms are true when interpreted as about that physical system.

On this view, mathematical facts are theorems: conditional statements about what follows from a list of mathematical axioms. When we apply a mathematical theorem to a physical situation, we show that the situation satisfies the "physical interpretation" of the theorem's axioms.

To illustrate, consider Colyvan's (2001) example of the Borsuk-Ulam Theorem (I will return to this example throughout this section). The Theorem is a result in algebraic topology which helps demonstrate that, at any given moment, there must exist two antipodal points on the Earth's equator with equal temperature. The Theorem states that:

> Let f be a continuous function from S (the circle) to $\mathbb{R}$ (the real numbers). Then there exist antipodal points x and $-x$ in S such that $f(x) = f(-x)$.

If we think of temperature as a function from points on the equator to real numbers which represent degrees Celsius, then the physical interpretation of the axioms of the Borsuk-Ulam Theorem is as follows:

- The equator is roughly circular (i.e., the domain of the temperature function is a physical instantiation of the circle S).
- At any given moment, there exists no spatial point x on the equator such that, for some positive t, any segment of the equator centred on x will contain a spatial point with temperature at least t degrees different than x's temperature (i.e., the temperature function satisfies the physical interpretation of continuity).

And the physical interpretation of the conclusion of the theorem is:

- At any given moment, there exist two antipodal points on the equator with equal temperature.

Now, why is it advantageous to model physical situations using mathematics in the way I just described? A range of answers are available into the literature. Saatsi (2011) writes:

> Mathematics can be knowledge conferring by virtue of providing justification for a hypothesis regarding a physical fact without explaining that fact. It can do this by demonstrating that the fact in question 'follows' from other physical facts regarding which we have better justified beliefs.

According to Saatsi, a mathematical theorem can help justify our inference from a set of physical premises to a physical conclusion. To return to the example I discussed above, once we've established that the equator is roughly circular and that temperature is continuously distributed along the Earth's surface, the Borsuk-Ulam Theorem allows us to infer that there must exist two antipodal points on the equator with equal temperature.

Pincock (2007) and Leng (2012) offer an alternative perspective: they argue that mathematizing scientific explanations helps make those explanations more general. Pincock gives as an example the explanation of why it's impossible to walk across each of the bridges of Königsberg exactly once and return to the same starting position. Modelling the bridges as a graph (a set of vertices and edges) allows us to explain this fact using Euler's Theorem, which states that a graph has a walk of the type I described iff every vertex of the graph has an even valence. All of the edges in the "graph" of the bridges of Königsberg have an *odd* valence, so such a walk does not exist.

Why is this graph-theoretic explanation preferable to a physical explanation? Pincock (2007) writes:

> Abstract explanations are useful to scientists because they are successful even when the microphysical configuration of a system changes. A microphysical explanation of why we could not walk a certain path on the bridge might fail if the microphysics of the bridges was sufficiently altered, e.g. the bridges were turned into gold. The abstract explanation seems superior because it gets at the root cause of why walking a certain path is impossible by focusing on the abstract structure of system.

The graph-theoretic explanation is more general not just because it's robust to changes in the *microphysical* properties of the bridges; it's also robust to changes in the *macrophysical* setup of the bridge system. It doesn't matter, for the graph-theoretic explanation, that there are seven bridges in Königsberg, or that they are arranged exactly as they are. All that matters is that at least one of the vertices in the "graph" of the bridges has an odd valence. The graph-theoretic explanation works for any set of bridges that has this property.

As this example should make clear, when we make a scientific explanation more general, what we're really doing is narrowing down the set of assumptions that we need to explain or derive the explanandum. The graph-theoretic explanation is more general than the (hypothetical) physical one because it requires a much smaller set of assumptions about the bridge system.

So although Pincock and Leng talk about *generality* rather than *justification*, they're implicitly claiming that mathematical inferences help us justify physical inferences. More specifically, they're claiming that a mathematical inference can help us identify the minimal set of assumptions required to (justifiably) make an analogous physical inference. In Pincock's example, a mathematical inference (Euler's Theorem) helped us identify the minimal properties a system of physical bridges must possess in order for us to infer that an Eulerian walk along those bridges is impossible.

The upshot of this discussion is that I believe that most easy-road nominalists accept that there's a close connection between *mathematical inferences* and *physical inferences*: valid mathematical inferences can be "translated" into valid physical inferences. Suppose we believe this. How does that commit us to the counterfactual covariance of the mathematical and physical facts?

To see how, let's return to the example of the Borsuk-Ulam Theorem. Recall that the Theorem states that:

> Let f be a continuous function from S (the circle) to $\mathbb{R}$. Then there exist antipodal points x and $-x$ in S such that $f(x) = f(-x)$.

If we assume that the physical interpretations of the theorem's axioms hold (the equator is roughly circular and temperature is continuously distributed), the Borsuk-Ulam Theorem guarantees that the physical interpretation of the conclusion (there exist antipodal points on the equator with equal temperature) must hold.

Now, suppose that the Borsuk-Ulam Theorem had a different conclusion: suppose, for example, that it implied that there *cannot* be antipodal points x and $-x$ such that $f(x) = f(-x)$. I claim that this is consistent with two possible configurations of the physical facts. The first possibility is that the physical situation no longer satisfies the physical interpretation of the Theorem's axioms: perhaps temperature is discontinuous at certain points, or the equator is relevantly different from a circle. In this case, obviously, the physical facts must be different.

The second possibility is that the situation still satisfies the physical interpretations of the Theorem's axioms. In this case, given the Theorem's truth, it must be the case that (the physical interpretation of) its new conclusion holds. So, again, the physical facts must be different.

But *why* is it the case that, if the physical situation continues to satisfy the Theorem's axioms, it must satisfy its new conclusion? The brute assertion that it *must* is unsatisfying. It may even be defensible to deny that, if a physical situation instantiates the axioms of a mathematical theorem, then it must instantiate the theorem's conclusion. Clarke-Doane (2012) suggests something along those lines while responding to Joyce's (2006) claim that mathematical beliefs are fitness-enhancing only if they are true.

Consider the fact that, if one lion goes behind a bush and then another lion goes behind the same bush, there are now two lions behind that bush. Clarke-Doane denies that this fact is a consequence of (or has anything to do with) the mathematical fact that 1+1=2; rather, he claims that it is a consequence of the first-order logical fact that *if there is an x behind the bush such that x is a lion, and a y behind the bush such that y is a lion, and $x \neq y$, then there are two lions behind the bush.*[10]

Clarke-Doane even suggests that it is conceptually possible for the facts about arithmetic to be different while the first-order logical facts about lions remain the same (and vice versa). For

[10] Field (1989, p.8) makes a similar point.

example, he claims, it is conceptually possible to imagine that 1+1=0 but that one lion behind the bush and another lion behind the bush still makes two lions behind the bush. He writes:

> Realistically construed, the claim that 1+1=0 speaks of numbers. It says, roughly, that the number 1 bears the plus relation to itself and to 0. What if we imagine that the number 1 bears the plus relation to itself and to 0 and maintain the (first-order) logical truth that if there is exactly one lion behind bush A, and there is exactly one lion behind bush B, and no lion behind bush A is a lion behind bush B, then there are exactly two lions behind bush A or B?

Clarke-Doane's arguments suggest the following view: the mathematical facts (realistically construed)[11] "float free" of the physical (and first-order logical) facts. Arithmetic facts are facts about *numbers*, and have nothing to do with the physical behaviour of lions behind bushes. The fact that 1+1=2 does not imply that, if there is one lion behind the bush and another lion behind the bush, there must be two lions behind the bush. More broadly, the fact that a physical situation instantiates the axioms of a mathematical theorem does not guarantee that it also instantiates its conclusion.

The view I just outlined seems minimally plausible to me, which is why it's unsatisfactory to simply deny it. However, I believe that if we accept the easy-road nominalist claims I discussed above, we are committed to rejecting this view. Suppose we accept Saatsi's claim that mathematical inferences help justify physical inferences, or Pincock's and Leng's claim that mathematical inferences help us discover the minimal assumptions required to make analogous physical inferences. Then we are committed to claiming that mathematical inferences *carry over* to physical settings. That is, we are committed to accepting that if a certain inference is valid in a mathematical setting, it must also be

[11] I'm not assuming mathematical realism, but I do want my argument to be robust to it.

valid in an analogous physical setting. This implies that the mathematical facts do not "float free" of the physical facts in the way I described above.

Thus, if we accept an easy-road nominalist account of the application of mathematics in science, we must accept that, if the conclusion of the Borsuk-Ulam Theorem were different, the physical facts would have to be correspondingly different. More generally, consider a theorem which is invoked in a scientific explanation. If the conclusion of the theorem were different, one of two possibilities must obtain. Either the physical situation must no longer instantiate the axioms of the theorem, or the physical situation must instantiate the new conclusion of the theorem. In either case, the physical facts must be different.

Strictly speaking, there is a third possibility. Recall that when analysing the sensitivity of our moral and mathematical beliefs, we must use *conceptual* possibility rather than ordinary metaphysical possibility. And even if easy-road nominalism commits us to the existence of a modally robust connection between mathematical inferences and physical inferences, it's not clear that this connection is *conceptually* necessary. It does seem, as Clarke-Doane claims, *conceptually* possible for a physical situation to instantiate the axioms of a theorem without instantiating its conclusion; at least, this is not any less intelligible than imagining that 17 is not prime. It is therefore *conceptually* possible for the conclusion of a theorem to be different while the physical facts remain identical.

This third possibility is not, however, a problem, since it is far more remote than the first two possibilities. Remember that our D-beliefs are sensitive if, *in the closest possible world* in which the D-truths are different and we use the same method to acquire our D-beliefs, our D-beliefs are correspondingly different. This "closest possible world" requirement protects the definition of sensitivity from collapsing into triviality. Obviously, for any belief X, there is *a* conceptually possible world where X is false, and we form our belief in the same way, but we nevertheless believe that X. It does not follow from this that all of our beliefs are trivially insensitive. Clearly, it is only *nearby* possible worlds that matter in determining the sensitivity of our beliefs. Intuitively, this is because we are interested in whether our beliefs *would be* correspondingly different if the relevant facts were different, not whether our beliefs *could fail to be* correspondingly different if the relevant facts were different.

Moreover, the third possibility is substantially more remote than the first two. Generally speaking, metaphysically impossible worlds are much further away from the actual world than metaphysically possible worlds. The first and second possibilities above involve *one* metaphysical impossibility: that a mathematical theorem has a different conclusion. The third possibility requires *two* metaphysical impossibilities: that a mathematical theorem has a different conclusion, *and* that the Earth instantiates that theorem's axioms without instantiating its (new) conclusion. Naturally, then, the third possibility is more remote than the first two, so it is not relevant when analysing the sensitivity of our beliefs.

It is therefore true that easy-road nominalism commits us to the claim that *certain* variations in the mathematical facts imply corresponding variations in the physical facts. I say "certain" because I've only considered one type of variation in the mathematical facts: namely, the conclusions of existing mathematical theorems being altered. I haven't considered other types of variations in the mathematical facts. I now turn to this point.

4 What if There Were No Mathematical Facts?

Our mathematical beliefs are sensitive if they're robust to *arbitrary* variations in the mathematical facts. In Section 3, I considered one type of variation in the mathematical facts: what if the conclusions of existing mathematical theorems were altered, while the truth of those theorems was preserved? I argued that this type of variation would have to be accompanied by corresponding variations in the physical facts.

My arguments from Section 3 could easily be extended to cover another possibility: what if the mathematical facts were different because entirely new mathematical theorems were true? If that were the case, then the new theorems would either combine with physical premises to imply new physical conclusions in the way I described in Section 3.1, or they would impose constraints on physical inferences in the way I described in Section 3.2. Either way, the physical facts would have to be correspondingly different.

There is a third way in which the mathematical facts could be different: what if there were no mathematical facts at all? In genealogical debunking arguments against our moral beliefs, this is the counterfactual that is most often invoked by debunkers. Joyce (2001) writes:

> Now imagine instead that the actual world contains no such [categorical] requirements at all – nothing to make moral discourse true. In such a world, humans will still be disposed to make these judgments […]

It is not immediately clear that, if there were no mathematical facts, the physical facts would have to be any different. Consider again the Borsuk-Ulam Theorem:

> Let f be a continuous function from S to $\mathbb{R}$. Then there exist antipodal points x and $-x$ in S such that $f(x) = f(-x)$.

Suppose that the Borsuk-Ulam Theorem were false (its conclusion failed to follow from its axioms). It could still be true that, at any given moment, there exist two antipodal points on the Earth's equator with the same temperature. This could be true for a different reason (for example, because a physical law dictated that it must be the case), or simply by coincidence. The world could remain physically identical, even if the Borsuk-Ulam Theorem were false.

More generally, mathematical theorems entail or help justify physical conclusions; but even if those theorems were false, those conclusions could still obtain. Thus, if there were no mathematical facts, the physical facts would not *necessarily* have to be any different. However, as I will explain, I don't believe that this fact threatens the sensitivity of our mathematical beliefs.

I argued in Section 3 that our beliefs about the role of mathematics in scientific explanations commit us to the claim that mathematical inferences "carry over" to analogous physical settings.[12] Since an inference is really just a law ("if fact X obtains, then fact Y must obtain"), a different way of stating my conclusion is to say that the physical laws must be consistent with the mathematical laws: physical laws cannot contradict their mathematical analogues. If it were possible for the physical and mathematical laws to be inconsistent with each other, mathematics would not be useful in science, since mathematical inferences couldn't be used to justify or refine physical inferences.

My arguments in Section 3 therefore imply that the mathematical facts impose *constraints* on the physical world by limiting the possible range of physical laws: only physical laws that are consistent with the mathematical facts can exist.[13] [14] For example, since it is true that 2+2=4, there cannot be a physical law which says that two strawberries together with another two strawberries yields five strawberries.

Now, imagine a world W in which there are no mathematical facts. The absence of mathematical facts means that there are no constraints on the possible physical laws in W, and, therefore, no constraints on the possible configurations of the physical facts. However, as I explained above, W *could* be physically identical to the actual world. It could be physically identical either purely by

---

[12] I was discussing the beliefs of easy-road nominalists, but presumably indispensabilists believe an even stronger version of this.
[13] This is very similar to Lange's (2013) characterization of the relationship between mathematical and physical laws.
[14] I therefore, like Melia (2006), deny the conservativeness of "applied mathematics," which in this context is pure mathematics along with bridge laws linking mathematical inferences to physical ones.

coincidence (i.e. all the physical particles in W happened to end up in the same arrangement as in the actual world), or because W happened to have exactly the same physical laws as the actual world.

So it is possible that, if there were no mathematical facts, the world would remain physically identical. Does this possibility undermine the sensitivity of our mathematical beliefs? I don't think it does. This is because, if there were no mathematical facts, it would require a huge coincidence for the world to remain physically identical. Among *all* the possible configurations of the physical laws in the absence of mathematical constraints, the world would have to retain *exactly the same* physical laws. As I'll explain, I don't think this sort of coincidence can undermine our beliefs.

A mathematical debunker might object as follows:

> Recall that what matters for the sensitivity of our mathematical beliefs is what happens in the *closest possible world* in which there are no mathematical facts. The closest possible world in which there are no mathematical facts is a world in which all other facts are maximally preserved. So the closest possible world in which there are no mathematical facts is a world that is physically identical to our own. This world is the relevant counterfactual for assessing the sensitivity of our mathematical beliefs; it doesn't matter that it requires a "huge coincidence."

To illustrate the problem with this objection, consider the following example. Imagine that there exists an important lottery, the outcome of which has major implications for the way society will be structured. Imagine that I am aware of the existence of an unscrupulous group who may try to rig the lottery in their favour. The conditional probabilities work out such that, as a good Bayesian, I form the following belief: "if the unscrupulous group wins the lottery, then they must have rigged it."

The day of the lottery rolls around, and the unscrupulous group wins, with their victory radically altering the structure of society. Based on their victory, I form the belief that the unscrupulous group rigged the lottery. We can, I think, agree that my belief is epistemically perfectly sound.

But let's try to analyse the sensitivity of my belief that the unscrupulous group rigged the lottery (supposing that it is actually true). My belief is sensitive if, in the closest possible world in which they did *not* rig the lottery, I do *not* believe that they rigged it. Since I form the belief that the unscrupulous group rigged the lottery iff they win the lottery, the crucial question is: in the closest possible world in which the unscrupulous group didn't *rig* the lottery, do they still *win* it?

Here is one answer: if the unscrupulous group had not rigged the lottery, they would have been extremely unlikely to win it. So the closest possible world in which they didn't rig the lottery is one in which they don't win it. Therefore, the closest possible world in which they didn't win the lottery is one in which I don't believe that they rigged it. My belief is therefore sensitive.

Here is an alternative answer: the closest possible world in which the unscrupulous group didn't rig the lottery is a world where all non-rigging facts are maximally preserved. So the closest possible world in they didn't rig the lottery is one in which they still won the lottery, so that all the social consequences of the lottery's result remain the same. Therefore, the closest possible world in which they didn't win the lottery is one in which I still believe that they rigged it. My belief is therefore insensitive.

These answers contradict each other because they use different notions of modal "closeness." The first answer interprets "closest possible world in which X is true" as "the world that is most likely to occur, given that X is true,"[15] and the second answer interprets it as "the world in which X is true and all non-X facts are maximally preserved."

It seems clear to me that the first answer is the correct one when it comes to analysing the sensitivity of our beliefs. When we ask whether a belief A is sensitive, what we're interested in is *the likelihood that, if A were false, we would still believe that A*. The first answer correctly states that my belief about the unscrupulous group is sensitive because, *if the group had not rigged the lottery, it would have been extremely unlikely for me to end up believing that they had rigged it.*[16]

To return to the mathematical context: the reason our mathematical beliefs are sensitive is that, if there were no mathematical facts, it would be extremely unlikely[17] that we'd end up in a world with precisely the same physical laws as the actual world. So even though it's true that, if there were no mathematical facts, the world *could* remain physically identical, this fact does not undermine our mathematical beliefs.

5 How Much of Mathematics is Indispensable to Science?

Field (1989) anticipates the argument I advance in this paper. He writes:

[15] This is loose and imprecisely formulated. Technically, any *particular* world in which they don't win the lottery is just as likely as the world where they win the lottery, since any particular outcome of the lottery is as likely as any other particular outcome. But you get the idea.
[16] More broadly, this suggests that we shouldn't be interested in what happens in the single probabilistically closest possible world, but in some average of what happens in a bunch of possible worlds, weighted by their probabilistic closeness.
[17] By "unlikely," I mean either "unlikely" in a non-probabilistic sense (like the non-probabilistic sense of "coincidence" that Street 2006 uses), or in a probabilistic sense (it's not inconceivable to imagine defining a probability measure on the set of all possible worlds where there are no mathematical facts).

> One could argue, for instance, that if mathematics is indispensable to the laws of empirical science*, then if the mathematical facts were different, different empirical consequences could be derived from the same laws of (mathematicized) physics*. (emphasis in original)

Field objects that

> […] the amount of mathematics that gets applied in empirical science […] is relatively small. This means that only the reliability of a small part of our mathematical beliefs could be directly explained by [this argument] […]

Field is right that only a limited segment of our mathematical beliefs are explicitly invoked in scientific explanations. But this fact does not threaten my argument. Consider an abstract mathematical fact M which is not directly invoked in any scientific explanations. If M were false, various other mathematical facts would have to be different as well. For example, any fact which implies M would also have to be false; and any facts which M is used to derive might also end up being false. So the falsity of M would create a "ripple effect" that would eventually extend to the applied mathematical facts. The resulting changes in the applied mathematical facts would imply changes in the physical facts, and therefore, if M were false, the physical facts would have to be different.

Why should we expect the ripple created by changing an abstract mathematical fact to eventually reach applied mathematical facts? We should expect this because the practical mathematical methods that we apply in scientific investigations are nested within, and rely on, more abstract mathematical frameworks. Changes in the abstract facts that define and characterize those frameworks will imply changes in the practical mathematical facts that we apply in science.

For example, calculus, which is used pervasively in science, depends on the axioms of real analysis. Elementary results in optimisation, like the Extreme Value Theorem, rely on abstract axioms like the

Least Upper Bound Axiom. Real analysis, in turn, depends on the axioms of set theory. Many proofs in analysis rely on some weak version of the Axiom of Choice: for example, proving the equivalence of the sequential-convergence and epsilon-delta characterisations of the continuity of a function requires countable Choice.

Similarly, modern probability theory (which provides the foundations for statistics) is formulated in terms of the language and structures of measure theory. Measure theory also depends on the axioms of set theory: for example, the nontriviality of the Lebesgue Measure depends on the Axiom of Choice.

Moreover, as Wigner (1960) classically describes, many fundamental objects in physics are directly modelled as elements of abstract mathematical structures. For example, possible states in quantum mechanics are modelled as elements of a Hilbert space. (Dirac, 1958).

Finally, virtually every result in applied mathematics depends on the less controversial axioms of set theory, like the Axioms of Union or Extensionality.

This brief discussion should make it clear that changes in the axioms or essential theorems of analysis, set theory, or many other abstract domains of mathematics would imply changes in applied mathematics, and therefore changes in the physical facts. So even very abstract mathematical facts covary counterfactually with the physical facts, and thus (if my arguments are successful) our abstract mathematical beliefs are also sensitive, contrary to Field’s objection.

Of course, there may be certain mathematical domains that are so deeply segregated from applied mathematics that changes in their constitutive facts would not lead to any changes whatsoever in the physical facts. As Field (1989) writes,

> […] too many different answers to questions about, say, large cardinals or the continuum hypothesis or even the axiom of choice work well enough at giving us the lower level mathematics needed in science and elsewhere.

But, as Clarke-Doane (2016a) argues, it is not clear why it is important to defend the reliability of our beliefs about extremely segregated mathematical domains. Indeed, it *is* quite mysterious how we could acquire reliable knowledge about mathematical domains that are so thoroughly isolated from the physical world.

6 A Genealogical Vindication of Our Mathematical Beliefs

Recall that the sensitivity-threatening argument I sketched in Section 2 proceeded as follows:

(a) Our mathematical beliefs are the products of purely physical processes. (Genealogical Premise)

(b) If the mathematical truths were different, the physical facts would remain the same. (Modal Claim)

(c) Therefore, if the mathematical truths were different, those physical processes would remain the same.

(d) Therefore, if the mathematical truths were different, our mathematical beliefs would remain the same. (Conclusion)

In Sections 3-5, I argued that any major variation in the mathematical facts would produce corresponding variations in the physical facts. If my arguments are correct, then this sensitivity-threatening argument fails because premise (b) is false. However, defeating this particular formulation of the sensitivity-threatening argument does not quite put the nail in the coffin on the Benacerraf-Field Problem. For example, a proponent of the Problem could alter premise (b) to read

(b) If the moral/mathematical facts were different, the physical facts would not be different *in a way that would produce correspondingly different moral/mathematical beliefs.* (Modal Claim*)

This proponent could argue that even if changes in the mathematical facts imply *some* changes in the physical facts, these changes would not be enough to correspondingly alter our mathematical beliefs. This modified version of the argument would require a much more detailed account of the genealogy of our mathematical beliefs than any account which has previously been on offer in the literature, but such an account could in principle be given.

So even if my arguments from the previous sections are successful, they leave our mathematical beliefs in an epistemically unstable position: simple formulations of the Problem have been defeated, but more sophisticated formulations could potentially still undermine our mathematical beliefs.

There is, however, a way to preempt any potential reformulations of the sensitivity-threatening argument. My arguments in Sections 3-5 suggest that, if the structure of the universe of mathematical facts were to change, then the structure of the physical universe would have to change correspondingly. So if I can argue that we acquire our mathematical beliefs in a way that is appropriately informed by the structure of the physical world, then I can *positively demonstrate* the sensitivity of our mathematical beliefs. If changes in the mathematical universe imply corresponding changes in the physical universe, and if our mathematical beliefs are appropriately informed by the physical universe, then changes in the mathematical universe must imply corresponding changes in our mathematical beliefs.

Giving a full account of the genealogy of our mathematical beliefs is beyond the scope of this paper (and, plausibly, beyond the scope of philosophy as a discipline). But I will present two arguments which strongly suggest that our mathematical beliefs are appropriately informed by the physical world.

First, mathematics has often developed in tandem with science, and mathematical theories have frequently been developed in order to meet the needs of scientists. As Field writes:

> The theory of real numbers, and the theory of differentiation etc. of functions of real numbers, was developed precisely in order to deal with physical space and physical time and various theories in which space and/or time played an important role. (Field 1980, p.34)

Of course, more abstract mathematical theories were not formulated for use in science. But they *were* often formulated in order to provide rigorous foundations for informal methods that were widely applied in science. For example, the development of analysis in the 19th and 20th centuries was motivated in part by the need to establish plausible axiomatized foundations for calculus.

Thus, informal applied methods grew out of scientific practice; and rigorous abstract theories grew out of those informal applied methods. So much of mathematics has its roots in observations of the physical world.

Secondly, mathematicians frequently make use of their spatial intuitions – intuitions which have developed through exposure to the physical world.[18] Anyone who has studied analysis or topology knows that intuitions about one- or three-dimensional space play a key role in understanding properties like connectedness or boundedness, and that proofs of theorems about connectedness or boundedness are often based on simple physical arguments. It is unsurprising that physical intuitions play such an important role in many mathematical domains, because many mathematical structures are just abstract generalisations of a Euclidean space equipped with a certain function. For example,

[18] It's fabled that LEJ Brouwer was inspired to prove his famous fixed point theorem by observing the swirling of coffee in his mug (Bocklandt, 2017).

the concept of a "metric space" is just a generalisation of n-dimensional Euclidean space with the standard distance function.

So our mathematical beliefs are, in many ways, appropriately informed by the structure of the physical world. As I explained above, this allows us to positively demonstrate the sensitivity of our mathematical beliefs. It allows us to show that our mathematical beliefs are sensitive *because of* (not despite) their physical genealogy.

7 Alternative Accounts of "Truth-Tracking"

Recall that the Benacerraf-Field Problem, as originally formulated, threatens the "reliability" of our mathematical beliefs. Showing that our mathematical beliefs are *sensitive* is useful only if the Modal Account correctly captures what it means for a belief to be reliable. Do the arguments I've presented hold up even if we reject the Modal Account, and accept some alternative account of "reliability" instead?

I believe that they do. I will briefly consider three alternative accounts: Field's preferred account of "reliability," and two accounts of "truth-tracking" from the genealogical debunking literature.

Field (1989) writes

> […] there was no obvious reason to put the challenge of explaining the reliability of our mathematical beliefs in modal or counterfactual terms in the first place. […] there is still the problem of explaining the *actual* correlation between our believing "p" and its being the case that p.

Field suggests that in order to explain the reliability of our mathematical beliefs, we need to explain the actual correlation between "mathematicians believing M" and "M being true."

If the arguments I've presented are correct, then it's easy to explain this correlation: mathematicians hold their beliefs because of their interactions with the physical world, and beliefs formed in this way tend to match the mathematical facts because the structures of the physical world mirror the structures of the mathematical universe. If (as I've argued) this explanation is successful, then our mathematical beliefs are reliable, even if we accept Field's account.

As I explained in Section 1, the challenge to show that our beliefs are reliable is very similar (or even identical) to the challenge from the genealogical debunking literature to show that our beliefs track the truth. In the debunking literature, there are two primary accounts of what it means for our beliefs to "track the truth." These accounts are suggested by Street (2006), and by Joyce (2016b) and Harman (1986).

Street (2006) claims that our beliefs about a domain D *fail to* track the truth if it would require an "incredible coincidence" for our D-beliefs to be true. Suppose we accept the conclusions for which I've argued in this paper. Then the truth of our mathematical beliefs is no coincidence: as I argued in Section 4, it would in fact require a huge coincidence for our mathematical beliefs to be *false* while the physical world remains as it is.[19] Our mathematical beliefs are therefore reliable, even if we accept Street's account.

By contrast, Joyce (2016b) and Harman (1986) claim that our beliefs about a domain D track the truth if the best explanation of how we came to acquire our D-beliefs presupposes or implies the truth of those beliefs. Suppose we believe that, as I've argued, our mathematical beliefs are informed by the structure of the physical world. Then the best explanation of how we came to acquire our mathematical beliefs proceeds something like this:

1) The physical world has features X.

[19] Since that would require that no mathematical facts exist *and* the world remains physically identical.

2) Based on features X, we (for psychological/evolutionary reasons) formed analogous mathematical beliefs.

Does this explanation presuppose or imply the truth of our mathematical beliefs? Strictly speaking, the answer is no. In Section 3, I argued that the structure of the physical world must be compatible with the structure of the mathematical universe. The fact that the physical world has features X *suggests* that the mathematical universe also has features X, since it implies that the mathematical universe could not have any features which contradict X. But it does not *imply* that the mathematical universe has features X, since it could be the case, for example, that there are no mathematical facts at all (as I explained in Section 4).

However, as I argued in Section 4, the fact that the physical world is *compatible with* the nonexistence of any mathematical facts does not undermine our mathematical beliefs, since, if no mathematical facts existed, it would require a huge coincidence for the physical world to have all the structures that it actually does. So even if we accept Joyce and Harman's account, our mathematical beliefs are not undermined by the Benacerraf-Field Problem.

Notably, my responses to all three of these alternative accounts of "reliability" were quite similar to each other: in each case, I used the broad conclusions for which I've argued to show that our beliefs are reliable, even on this alternative account. This should hopefully illustrate that the arguments I've presented in this paper aren't just a technical, dialectic-specific response to the modal formulation of the Benacerraf-Field Problem. Rather, my arguments point to general features of our mathematical beliefs which show, intuitively, that our mathematical beliefs are "reliable," whatever technical meaning we attribute to "reliability."

7 Logic, Modality, and Morality

The sensitivity-threatening argument that I've discussed in this paper could easily be generalised to threaten our beliefs in a variety of domains:

(a) Our D-beliefs are the products of purely physical processes. (Genealogical Premise)

(b) If the D-truths were different, the physical facts would remain the same. (Modal Claim)

(c) Therefore, if the D-truths were different, those physical processes would remain the same.

(d) Therefore, if the D-truths were different, our mathematical beliefs would remain the same. (Conclusion)

Premise (a) is plausibly true of all of our beliefs. So whether this argument succeeds in a particular domain depends on whether premise (b) is true for that domain. In particular, it's of interest whether premise (b) is true for the logical, modal, and moral facts.

If the arguments I've presented in this paper are correct, then premise (b) is certainly false for the logical facts. The logical facts are even more deeply embedded in scientific explanations than the mathematical facts are, both because all scientific inferences rely on logical assumptions, and because the mathematical facts themselves depend on the logical facts.[20] So if the logical facts were different, their role in scientific explanations implies that the physical facts would have to be correspondingly different. Premise (b) is therefore false for the logical facts, and as a result our logical beliefs are not vulnerable to the Benacerraf-Field Problem.

Whether premise (b) is true for our modal beliefs is much harder to evaluate. To begin with, it's unclear whether it's even intelligible to talk about the sensitivity of our modal beliefs. Since the modal facts in a particular world are *about* counterfactual worlds, what would it mean to imagine a counterfactual world in which the modal facts are different?

[20] The logical facts determine which mathematical conclusions follow from which mathematical axioms – for example, many theorems rely on the Law of Excluded Middle.

We might be able to get around this issue by thinking about different "universes" of possible worlds. Rather than asking ourselves about the status of our modal beliefs in nearby possible worlds *within* our modal universe, we could ask about the status of our modal beliefs in a similar world that inhabits a *separate* modal universe.

If we go down this route, it seems possible to argue that scientific explanations rely on modal facts in much the same way that they rely on mathematical and logical facts. After all, statements about counterfactuals and possibilities are endemic in scientific discourse, and some scientific theories are defined directly in modal terms (for example, the Rubin causal model in statistics and econometrics is defined in terms of "potential outcomes"[21]). It therefore seems like modal facts play a crucial role in scientific explanations. But whether this sort of argument would succeed is a topic that requires a much more extensive discussion than the brief one I've offered, so I'll suspend judgement on the issue.

Finally, I believe that premise (b) is *true* for our (non-naturalistically construed) moral beliefs, and our moral beliefs are therefore vulnerable to the Benacerraf-Field Problem and genealogical debunking arguments.

More precisely, suppose we accept moral non-naturalism (premise (b) is obviously false for moral naturalists).[22] Consider our "higher-order" moral beliefs: our beliefs about the conditions under which moral properties (like "goodness") are instantiated. For example, "an action is morally right iff it increases the net amount of pleasure in the universe" is a higher-order moral belief.

It's clear to me that, if the higher-order moral truths were different, the physical world would remain exactly the same. For example, suppose it was instead true that "an action is morally right iff

[21] See Rubin (1974).
[22] By "moral naturalism," I mean the position that the moral facts are just natural facts. By "moral non-naturalism," I mean the position that irreducible, non-physical normative truths exist.

it increases the net amount of *pain* in the world." Then none of the physical facts would have to be any different. A different class of actions would be "morally good," but this is a change in the moral facts and not the physical ones.

A moral realist might object as follows:

> There is one class of cases where a change in the higher-order moral facts would lead to a corresponding change in the physical facts. Suppose I believe that pleasure is morally good because it is actually true that pleasure is morally good. Then if it were instead true that pain is morally good, I would alter my belief in order to believe that pain is morally good.

The obvious problem with this objection is that it begs the question against the Benacerraf-Field Problem, by presupposing that my moral beliefs are sensitive. So this objection is not available in this context.

I'm not aware of any other plausible way to argue that changes in the higher-order moral facts would imply corresponding changes in the physical facts. There is a literature that argues that moral facts are indispensable to certain scientific explanations, so perhaps arguments analogous to the ones I've advanced could be made on behalf of our moral beliefs. But this literature on "moral indispensability arguments" mostly focuses on examples involving lower-order moral facts (facts about specific instantiations of moral properties). For example:

- That Hitler was morally depraved explains why he murdered the Jews, in Sturgeon (1997).
- That Bob is honest explains why he brought forward self-incriminating evidence, in Railton (1998).
- That Donald behaved rudely explains why the people around him were embarrassed and annoyed, in Roberts (2016).

A lower-order moral fact like “Hitler was morally depraved” can be decomposed into a higher-order moral component (“agents with X physical characteristics are morally depraved”) and a purely physical component (“Hitler had X physical characteristics”). We can vary this lower-order fact either by varying its higher-order moral component or by varying its physical component. Obviously, if we vary its physical component (by imagining that Hitler did not have physical characteristics X) , then the physical facts would have to be different. So moral indispensabilists are right that variations in the lower-order moral facts can produce corresponding variations in the physical facts. But it’s equally clear that if we vary the higher-order component (by imagining that the facts governing the instantiation of “morally depraved” were different), then the physical facts would not have to be any different; only the moral facts would be different. So variations in the higher-order moral facts aren’t accompanied by corresponding variations in the physical facts.

Thus, I believe that premise (b) is true of our (non-naturalistically construed, higher-order) moral beliefs, and therefore our (non-naturalistically construed, higher-order) moral beliefs are undermined by the sensitivity-threatening argument.

Conclusion

The Benacerraf-Field Problem highlights an important tension in our beliefs. On the one hand, we wish to claim that our mathematical beliefs accurately track mind-independent abstract facts. On the other hand, we accept the common-sense claim that all of our beliefs are the output of purely physical processes. These two claims are in conflict.

In this paper, I’ve argued that they can nevertheless be reconciled. I’ve argued that our beliefs about the role of mathematics in scientific explanations commit us to the claim that the physical world reflects the structure of mathematical reality. Moreover, I’ve argued that our mathematical beliefs are informed by the characteristics of the physical world. I've concluded that our mathematical beliefs accurately track mathematical reality, because they're informed by physical structures which reflect the structure of mathematical reality.

I've also shown that an analogous argument could be made on behalf of our logical beliefs, and possibly also on behalf of our modal beliefs. By contrast, an analogous defence of our moral beliefs would not succeed. Our moral beliefs *are* undermined by the fact that we acquire them without any interaction with the moral facts. Our moral and mathematical beliefs are therefore not epistemically on a par, at least with respect to genealogical considerations, contrary to the claims of some authors.

References

Baker, Alan. 2005. “Are there Genuine Mathematical Explanations of Physical Phenomena?” *Mind* 114, no.454: 223-237

Baker, Alan. 2009. “Mathematical Explanations in Science.” *British Journal for the Philosophy of Science* 60, no.3: 611-633

Benacerraf, Paul. 1973. “Mathematical Truth.” *The Journal of Philosophy* 70, no.19: p.661-679

Bocklandt, Raf. 2016. “Reflections in a cup of coffee.” *Indagationes Mathematicae.* 10.1016/j.indag.2017.10.002.

Chalmers, David. 1996. *The Conscious Mind.* Oxford University Press

Clarke-Doane, Justin and Dan Baras. (Forthcoming). “Modal Security.” *Philosophy and Phenomenological Research*

Clarke-Doane, Justin. 2012. “Morality and Mathematics: the Evolutionary Challenge.” *Ethics* 122, no.2: p.313-340

Clarke-Doane, Justin. 2015. “Justification and Explanation in Mathematics and Morality.” In *Oxford Studies in Metaethics, Volume 10*, edited by Russ Shafer-Landau, p.81-101. Oxford University Press

Clarke-Doane, Justin. 2016a. “What is the Benacerraf Problem?” In *New Perspectives on the Philosophy of Paul Benacerraf: Truth, Objects, Infinity,* edited by Fabrice Pataut, p.17-37. Springer International Publishing Switzerland

Clarke-Doane, Justin. 2016b. “Debunking and Dispensability.” In *Explanation in Ethics and Mathematics: Debunking and Dispensability,* edited by Uri D. Leibowitz and Neil Sinclair, p.23-36. Oxford University Press

Clarke-Doane, Justin. 2017. “Modal Objectivity.” *Noûs* 53, no.2: 266-295

Colyvan, Mark. 2001. *The Indispensability of Mathematics.* Oxford University Press

Colyvan, Mark. 2010. “There is no Easy Road to Nominalism.” *Mind* 119, no.474: 285-306

Daly, Chris and Simon Langford. 2009. “Mathematical Explanation and Indispensability Arguments.” *The Philosophical Quarterly* 59, no.237: 641-658

Dirac, Paul. 1958. *The Principles of Quantum Mechanics.* Oxford University Press

Field, Hartry. 1980. *Science Without Numbers*. Oxford University Press

Field, Hartry. 1989. *Realism, Mathematics, and Modality.* Basil Blackwell

Field, Hartry. 2005. “Recent Debates about the A Priori.” In *Oxford Studies in Epistemology Volume 1*, edited by Tamar Szabo Gendler and John Hawthorne, p.69-89. Oxford University Press

Harman, Gilbert. 1986. “Moral Explanations of Natural Facts – Can Moral Claims Be Tested Against Moral Reality?” *The Southern Journal of Philosophy* 24: p.57-67

Hellman, Geoffrey. 1989. *Mathematics Without Numbers: Towards a Modal-Structural Interpretation.* Oxford: Clarendon Press

Joyce, Richard. 2001. *The Myth of Morality.* Cambridge University Press

Joyce, Richard. 2006. *The Evolution of Morality.* MIT Press

Joyce, Richard. 2016a. “Evolution, Truth-Tracking, and Moral Skepticism.” In *Essays in Moral Skepticism,* edited by Richard Joyce, p.142-159. Oxford University Press

Joyce, Richard. 2016b. “Confessions of a Modest Debunker.” In *Explanation in Ethics and Mathematics: Debunking and Dispensability,* edited by Uri D. Leibowitz and Neil Sinclair, p.124-145. Oxford University Press

Lange, Marc. 2013. “What Makes a Scientific Explanation Distinctively Mathematical?” *British Journal for the Philosophy of Science* 64, no.3: 485-511

Leng, Mary. “Taking it Easy: a Response to Colyvan.” *Mind* 121, no.484: 983-994

Majors, Brad. 2007. “Moral Explanation.” *Philosophy Compass* 2, no.1: p.1-15

Melia, Joseph. 2000. “Weaselling Away the Indispensability Argument.” *Mind* 109, no.435: 455-478

Melia, Joseph. 2006. “The Conservativeness of Mathematics.” *Analysis* 66, no.291: 202-208

Pincock, Christopher. 2007. “A Role for Mathematics in the Physical Sciences.” *Noûs* 41, no.2: 253-275

Putnam, Hilary. 1979. “What is Mathematical Truth.” In *Mathematics Matter and Method: Philosophical Papers, Volume 1,* edited by Hilary Putnam, p.60-78. Cambridge University Press

Railton, Peter. 1998. “Moral Explanation and Moral Objectivity.” *Philosophy and Phenomenological Research* 58, no.1: p.175-182

Roberts, Debbie. 2016. “Explanatory Indispensability Arguments in Ethics and Philosophy of Mathematics.” In *Explanation in Ethics and Mathematics: Debunking and Dispensability,* edited by Uri D. Leibowitz and Neil Sinclair, p.185-202. Oxford University Press

Rubin, Donald. 1974. “Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies.” *Journal of Educational Psychology* 66, no.5: p.688-701

Saatsi, Juha. 2011. “The Enhanced Indispensability Argument: Representational versus Explanatory Role of Mathematics in Science.” *British Journal for the Philosophy of Science* 62, no.1: 143-154

Sidelle, Alan. 1989. *Necessity, Essence, and Individuation: a Defence of Conventionalism.* Cornell University Press

Street, Sharon. 2006. “A Darwinian Dilemma for Realist Theories of Value.” *Philosophical Studies* 127, no.1: 109-166

Sturgeon, Nicholas. 1997. “Moral Explanations.” In *Ethical Theory 1: the Question of Objectivity*, edited by James Rachels. Oxford University Press

Wigner, Eugene. 1960. “The Unreasonable Effectiveness of Mathematics in the Natural Sciences.” *Communications on Pure and Applied Mathematics* 8, p.1-14s