Why believe infinite sets exist?

The axiom of infinity states that infinite sets exist. I will argue that this axiom lacks justification. I start by showing that the axiom is not self-evident, so it needs separate justification. Following Maddy's (1988) distinction, I argue that the axiom of infinity lacks both intrinsic and extrinsic justification. Crucial to my project is Skolem's (1922) distinction between a theory of real sets, and a theory of objects that theory calls "sets." While Dedekind's (1888) argument fails, his approach was correct: the axiom of infinity needs a justification it currently lacks. This epistemic situation is at variance with everyday mathematical practice. A dilemma ensues: should we relax epistemic standards or insist, in a skeptical vein, that a foundational problem has been ignored?

1. Infinite sets

The axiom of infinity states: *there is an infinite set*. One tends to think natural numbers form an infinite set, so do real numbers, so there has to be at least an infinite set, otherwise mathematics is left unaccounted for. This is the natural outlook I will challenge. Once we given up the project of making set theory foundational, the belief that there is an infinite set of natural numbers will seem gratuitous.

What is infinity? Natural numbers are a *simple* infinity (Dedekind 1888, par.73); there is no highest number. Dedekind tried to identify this simple infinity with *Dedekind*-infinity (idem, par.64), that of a set that can be put in a one-to-one correspondence to a strict subset of it.[1] These are two different concepts of infinity. Using either concept, the axiom of infinity can be put in many ways:

1 "a system S is said to be *infinite* when it is similar to a proper part of itself (par.32), in the contrary S is said to be a *finite* system."

(1) $(\exists I)\ (\varnothing\in I\ \&\ (\forall x)(x\in I\rightarrow \{x\}\in I))$ Zermelo (1908, p.204)

(2) $(\exists I)\ (\varnothing\in I\ \&\ (\forall x)(x\in I\rightarrow x\cup\{x\}\in I))$ von Neumann (1925, p.400)

(3) $(\exists I)\ [(\forall x)\ (x\in I\ \&\ (\forall z){\sim}(z\in x))\ \&\ (\forall x)\ (x\in I \rightarrow (\exists z)[z\in I\ \&\ (\forall w)(w\in z \leftrightarrow (w\in x\ \vee\ w= x))])]$

Boolos (1971, p.225)

(4) No finite class of individuals contains all individuals. Russell (1908, p.258)

(5) There are infinitely many individuals. Potter (2004, p.70)

(1)-(3) characterize Dedekind infinity, whereas (4)-(5) characterize simple infinity. All these are axioms of infinity because all of them are ways of saying that (at least) an infinite set exists. Do we have any reason to think the axiom of infinity is true under any of these guises?

The question might seem absurd. How could one do mathematics while denying infinite sets exist? I don't claim infinite sets don't exist. Nor do I claim they exist. I am agnostic about the truth of the axiom of infinity. Rather, I issue an epistemological challenge. What are the reasons for thinking the axiom of infinity is true? How can the axiom be justified?

This question belongs to a mature research program aimed at justifying axioms of set theory (Maddy 1988). Debates have typically surrounded the axiom of choice (Skolem 1922), and axioms positing the existence of higher cardinals (Hauser and Woodin 2014). Yet asking what justifies the claim that infinite sets exist is a peculiar question. If no justification can be found for it, nor will we be able to justify axioms for higher cardinals. With such large implications, it is surprising to see so few reasons given for believing infinite sets exist. This is no accident. So few reasons have been given because no good reasons are available. Or so I will argue.

At this point, epistemology-first philosophies often meet with an epidermic reaction from the mathematical workforce. Here is an extreme version due to Abraham Robinson (1969, p.45):

> Infinite totalities do not exist and any purported reference to them is, literally, meaningless; [yet] this should not prevent us from developing mathematics in the classical vein, involving the free use of infinitary concepts

Rhetoric and his own nominalistic leanings aside, Robinson voices an impatience with matters conceptual, and advocates business-as-usual: the free development of mathematics from within, unhindered by any philosophical thought-police.

It is hard to see what the best way of taking Robinson's point is. No one doubts the practical utility and sheer beauty of, say, Zermelo-Fraenkel set theory. Mathematicians moved by such virtues should be all the more willing to show that something as crucial to the whole enterprise as the axiom of infinity is justified. If unjustified, the axiom is just an article of faith – or a prejudice.

Yet it is natural to think the verdict might be too harsh, and that a more prudent diagnosis would be to say that there is, ultimately, a mismatch between the epistemic standards we are in-principle willing to apply, and the set-theoretical practice to which we find ourselves applying it. I return to the chances of this approach in the conclusion, by re-framing the epistemological challenge as one to reform either our epistemic norms or extant mathematical practices.

2. Self-evidence

One might think the claim that infinite sets exist doesn't need to be justified simply because it is self-evident. Mere inspection should suffice to convince us that the axiom is true. Here are a few such tries. A "basal intuition of mathematics" (Brouwer 1913/1983, pp.80-81) is to reapply the successor function +1 starting with 1 up to the smallest infinite ordinal ω, delivering a (simply infinite) set of natural numbers. Similarly, Zermelo (1917, p.307) writes that infinity "must be grasped as an idea in Plato's sense." For Skolem (1922, p.299) too, "the initial foundations [should] be something immediately clear, natural, and not open to question."

Note how strongly the founding fathers of set theory disagreed. For Skolem (p.299) intuition was the province of real arithmetic, whereas Brouwer (p.80) warned that the real continuum "can never be thought of as a mere collection of units," so the real numbers don't form a set. Zermelo (p.307) found "posited" intuitiveness in set theory rather than arithmetic. What one thought intuitive, others denied. In light of this peer disagreement, rational agents should at least trust intuitions about infinity less than they did before (Kelly 2010).

Set theory, in particular, seems ill-suited for intuitions. Boolos (1971) dubs Frege's view as "the intuitive view of sets". It is supposed to be intuitive precisely because it includes Law V, which states that to every concept there corresponds a set having as members all the objects that fall under the concept. Law V entails Russell's paradox. Hence the law is false, for all its intuitiveness. There is no reason to think intuitions about whether infinite sets exist are any closer to the truth.

In part, intuitions about infinity are unappealing because they are mysterious. Are intuitions produced by a special faculty, designed to peep into Cantor's paradise? No. As Boghossian (2003, pp.230-231) reminds us:

> no-one has been able to explain, clearly enough, in what an act of rational insight could intellegibly consist… The question is whether we can be said to have some sort of *non-discursive, non-ratiocinative, insight*... an insight that would disclose immediately, and without the help of any reasoning whatsoever[.]

Boghossian suggests a more mundane source for our intuitions: our concepts. But if intuitions about infinity rely on whatever concepts of infinity each of us has,[2] they cannot justify the claim that those concepts are instantiated, on pain of circularity.

The justification challenge can't be answered by saying it's self-evident infinite sets exist.

3. "Sets" and sets

Is the axiomatic method such that the axiom of infinity needs no justification? The suggestion can be dismissed offhand by adherents to the project of justifying axioms. But formalists may question that project itself. In the wake of Hilbert's program, von Neumann (1925) could suggest it is enough to write down a set of axioms and give it a model, thereby showing the axioms are jointly consistent.[3]

No general diagnosis of the many faces of formalism is sought here. Rather, I'm interested in a

---

2 Looking into how intuitions get formed may help. Pantsar (2015) argues that our most basic conceptual metaphor regarding infinity is that of an unending process. You start counting, and could go on counting into eternity. A process which could go on indefinitely is reified into a set which is actually infinite. One moves from a possible process to an actual set. (Bear in mind that this transition is supposed to underlie the everyman's conception of infinity, acquired and entrenched ever deeper as our cognition develops.) As Pantsar explains, the transition is metaphorical. And, I add, metaphors are unfit for justificatory purposes. This quibble aside, counting indefinitely is patently not counting an infinity, a point I will belabor later. If the slippage from indefiniteness to infinity underlies the everyman's thinking about infinity, elucidating it makes for a nice debunking of any epistemological role untutored intuitions of infinity were thought to play. No wonder untutored intuitions need tutelage.

3 To the charge that "objects cannot be conjured into existence by stipulation" (made in a different connection by Potter and Smiley (2001), p.337), a straightforward reply is that consistent sets of set-theoretical axioms have non-empty models; objects aren't made up.

particular criticism brought to Hilbert's school. This criticism, leveled by Skolem in the context of discussing the axiom of choice, is that there is no guarantee that the objects in the domains of models which satisfy an axiomatic set theory are *real* sets, as opposed to objects merely being called “sets” by one's theory. He writes:

> We can, after all, ask: What does it mean for a set to exist if it can perhaps never be defined? It seems clear that this existence can be only a manner of speaking, which can lead only to purely formal propositions - perhaps made up of very beautiful *words* - about objects called *sets*.

Skolem was criticizing the axiom of choice for making undefinable choice sets possible. His wider target, though, is Hilbert's axiomatic method. Reapply this to the axiom of infinity, and ask: why think the axiom of infinity is true for real sets, as opposed to true by convention for any objects that might be called “sets” by a theory? That would be to ask for a justification of the axiom of infinity.

The formalist has an easy reply. Where do we get our preferred conception of real sets? Why not get it from the axioms which, taken together, *implicitly define* what gets to count as a set? ZF, which includes an axiom of infinity, has a model in the sequence of natural numbers, then it is consistent. What more could one ask for?

It's unclear the rejoinder has a bite. The Skolemite could insist: which ones are the real sets? Which, of the many alternative set theories, is the one true theory? Quine (1960, p.259) answers: there's no fact of the matter. If the surrogates are good enough for the jobs we assign to them, which ones are real is a “don't care”.

No one doubts the attractiveness of the Quinean reply for the mathematical workforce, but it simply fails to address the ontological question. Here is an analogy. In vain do present members of the de Bourbon family call themselves “de France”, “comte de Paris”, “dauphin” etc. France is a republic, so these titles no longer carry the claim to life and limb they used to. The analogy is this. Suppose we implicitly (axiomatically) define what we decide to henceforth call “sets”. Despite a profligacy of surrogates, our decision leaves open the possibility that there are *no real sets*, much like there is no king of France however many contenders glorify the past. “Ontological seriousness” (Heil 1998) requires that we show sensitivity to the question of whether there are real sets – and whether any *of them* is infinite.[4] The further epistemological question is *why* things are as we find them in reality. But the epistemological question presupposes ontological seriousness. (None of this condones Platonism: perhaps there are no infinite sets! But the question which Platonism about sets[5] answers is a good one.)

4 As Lavers (2015) reconstrues it, central to Carnap's approach to infinity (from as early as 1934 to as late as 1950) is the use of coordinate languages. Within syntactic frameworks which mention real numbers as coordinates, the existence of infinitely many coordinates is supposedly ensured as a matter of logic. It is unclear what all this amounts to. Saying there are infinitely many real numbers is less than saying there is a set of real numbers, and that set is infinite. Moreover, if the theory couched in terms of a particular syntactic framework (a coordinate language) is consistent, then it has models. As Lavers (2015) recounts, Carnap insisted that the infinitely many positions which correspond to real numbers are not best thought of as *objects*. Why not – what should incline one to think positions aren't themselves some peculiar objects? Resnik (1981, p.530) is clear: “I view patterns and their positions as abstract entities.“ Carnap scholarship aside, it is clear we don't get an ontological free lunch of positions (and an infinite one, at that) from the mere use of coordinate languages.

5 It seems ontological seriousness compels us to distinguish between Platonists about abstract objects called “sets” (like Quine), and Platonists about sets (like Frege). This is no threat to semantic ascent. Tailoring an example from Quine (1948), the city of Naples is properly called “Naples” even if no

4. Infinite sets and the physical world

We can't evade the question of how to justify the axiom of infinity – the claim that infinite sets exist. The axiom isn't self-evident, and the axiomatic method can't help us skirt around the question of why we should believe infinite sets exist.

Whether, in working within a set theory, one actually comes to believe in the truth of the axiom of infinity as a matter of psychological fact, is an empirical question whose answer may vary from one individual to another. The *philosophical* question is on what grounds one may be justified in believing the axiom of infinity to be true. To this avail, Penelope Maddy (1988, pp.482-483) writes that some set-theoretic axioms:[6]

> follow directly from the concept of set, that they are somehow "intrinsic" to it (obvious, self-evident), while other axiom candidates are only supported by weaker, "extrinsic" (pragmatic, heuristic) justifications, stated in terms of their consequences, or inter-theoretic connections, or explanatory power

---

meanings roam the city, but a brewery could also be called "Naples" without thereby becoming the beloved city.

6 Attempted extrinsic justifications abound. Maddy (1988, p.486) invokes Cantorian finitism, for which "infinite sets are like finite ones." That is, *if* infinite sets exist, their analogy to finite ones should be maximized. This says nothing about whether infinite sets *do* exist or not. Another example is Zermelo (1908), who approaches set theory in the "anything short of contradiction" view of sets; but why should the universe of sets be maximal as opposed to minimal?

Extrinsic justifications of the axiom of infinity often follow Quine's (1951) emphasis on how central the axiom is to our knowledge. Thus, Maddy (1992, p.280) writes that

> the calculus is indispensable in physics; the set-theoretic continuum provides our best account of the calculus; indispensability thus justifies our belief in the set-theoretic continuum, and so, in the set-theoretic methods that generate it; examined and extended in mathematically justifiable ways, this yields Zermelo-Fraenkel set theory.

What Maddy seems to be suggesting is that the following chain of indispensabilities obtains. The axiom of infinity is indispensable to any workable set theory (say, ZFC). Any such theory is indispensable for other mathematical theories (like the calculus). The latter, in turn, are indispensable to mathematical physics. Mathematical physics is indispensable to our knowledge of many particular physical facts. What are we to make of these claimed indispensabilities?

Suppose the chain of indispensabilities is, in fact, a chain of bottom-up *confirmation.* Some particular facts obtain, and this confirms a physical theory (say, Newtonian mechanics in one of its guises). Confirming this empirical theory confirms the mathematical backbone without which the physical theory wouldn't be possible (the calculus). Confirming this confirms whatever foundations the calculus can't do without (inter alia, set theory).[7]

The problem is that the chain of confirmation breaks down as soon as we move from physics to mathematics. Unlike empirical claims, mathematical claims are either necessarily true, or necessarily false. They are true in all possible worlds or in none. Added confirmation in our world may lack or be

7 How confirmation relates to indispensability is treated more in-depth by Sober (1993) and Vineberg (1996). But the point I make in the text is quite general, and doesn't depend on the fine print in a theory of confirmation.

reversed in a different possible world.[8] So it is hard to see how physical theories could help confirm mathematical theories.

Suppose, alternatively, that indispensable to a truth is not what that truth confirms, but what *best explains* that truth (Baker 2005). Here we run into problems immediately. To revert to Maddy's example, set theory *doesn't* best explain calculus. Indeed, explaining calculus doesn't seem to be a well-defined task. So in what sense is set theory, and the axiom of infinity in particular, indispensable to calculus?

An easy answer might be: metaphysical *dependence.* Calculus couldn't be true unless infinite sets exist. Is this conditional true? Zermelo (1908) thought so because he thought the province of set theory is the whole of pure mathematics.[9] Russell (1919) thought the opposite.[10] At the very least, the claim that all mathematics is nothing but set theory is unmotivated. As Maddy (2011, p.33) remarks, problems in other areas of mathematics can *usefully* be *represented* in set-theoretic terms without

8 If one takes the radical step of saying laws of nature are themselves metaphysically necessary, this would put them in the same position as mathematical propositions. Then it is hard to see how laws of nature themselves could ever be confirmed.

9 Zermelo (1908, p.200) writes: “Set theory is that branch of mathematics whose task is to investigate mathematically the fundamental notions “number”, “order”, and “function,” taking them in their pristine, simple form, and to develop thereby the logical foundations of all of arithmetic and analysis.”

10 Russell notes (1919, pp. 132-133) that, in the absence of the axiom of infinity, set theory “cannot deal with infinite integers or with irrationals. Thus the theory of the transfinite and the theory of real numbers fails us”. Presumably, asking if real numbers can be represented as sets, Russell isn't asking whether they can be represented *as themselves.* So he must think we have reason to treat numbers and sets as distinct, unless we are convinced otherwise.

making any assumption about whether numbers or functions *really are* just sets.

Now reconsider the claim that calculus couldn't be true unless infinite sets exist. On Zermelo's outlook, this claim says that calculus can't be developed with the means of finite set theory. On Russell's outlook, it says calculus can't be represented in finite set theory. Both outlooks share a presupposition. Why think calculus needs to be representable in set theory at all?[11] No reason has been given to think that *real* sets *must*, of necessity, tally up to real numbers and functions. Suppose the presupposition fails, and real sets cannot represent real numbers and functions. This is conceivable. Nothing rules it out. If true, it goes to show that Maddy's claim that calculus couldn't be true unless infinite sets exist is false.

It is customary to say the axiom of infinity is so central to other areas of mathematics and the purer reaches of empirical science that it receives bottom-up support. It is, in a catchy phrase, "too big to fail" - too central to the business of mathematics, too well-entrenched to doubt it without mutilating current mathematical developments. Such metaphors are attractive because they are familiar. I fear they cannot be given the precise content needed to justify the axiom of infinity. Do attempts to justify the axiom that are somehow intrinsic to the concept of a set fare any better? To this I now turn.

5. Concepts and sets

Intrinsic justifications of the axiom of infinity proceed either via attempts to unpack the concept of sethood such that it will turn out there are infinite sets, or via direct argument. First things first.

---

11 True, we might devise theories about objects *called* "sets". One constraint on such theories might be that real arithmetic and functional analysis should be representable in them. But looking into the axiom of infinity is ontologically serious business. We need to talk about real sets (if any there be).

Merely having a concept of infinity doesn't guarantee it has a non-empty extension.[12] Infinite sets cannot be secured this way. But perhaps specific concepts of set are such that they guarantee the existence of infinite sets. This would do justice to Maddy's claim that *intrinsic* justifications are possible for axioms (here, for the axiom of infinity).

Boolos (1971, pp. 218-219) justifies the *iterative* conception by suggesting it is the next best thing from the failed Fregean conception. In presenting it, Boolos takes the claim that there is an ω-stage in the hierarchy as an *axiom* (axiom VI, p.223) constitutive of the iterative conception of sets and goes on to prove (3) above (stated and proven pp.225-226) on its basis. Boolos does provide a justification for thinking there are infinite sets, but his procedure seems circular. He assumes the existence of a simply infinite set (a set isomorphic to the set of natural numbers), and goes on to prove, on that basis, that a Dedekind infinite set exists (a set isomorphic to a proper subset of itself).

Maddy (1988, p.486) seems to suggest that the existence of infinite sets is somehow intrinsic to the iterative conception of a set. Yet there is nothing *incoherent* in an iterative hierarchy of *finite* sets. Finite arithmetic (Cohen 1966) presupposes one such hierarchy.[13] One can deny (3) (numbering cf.

---

12 Given that we have (more than) one concept of infinite set, how could there be no infinite set? The question presupposes an instance of Frege's Basic Law V (1888, Part I): $(\forall F)(\forall G)(\forall x)((Fx \leftrightarrow Gx) \leftrightarrow (\varepsilon x.Fx = \varepsilon x.Gx))$. Taking F as G, this implies any concept has an extension. Law V leads to Russell's paradox.

13 Cohen (1966, pp. 22-23) shows how a *finite* first-order arithmetic can be developed without the assumption, *or the result*, that there is a simply infinite set. Cohen (1966, p.23) develops a system of finite set-theoretical arithmetic having axioms for extensionality, the existence of an empty set, pairing,and union. In terms of these axioms, Cohen (p.24) takes the successor of an integer x to be $x \cup \{x\}$ such that any two integers are comparable (either one is a member of another, or they are identical), and integers are transitive (members of members of an integer *x* are members of *x*). The

section 1) yet retain an iterative conception of finite sets. Nothing in the very concept of an iterative hierarchy implies that infinite sets exist.

6. Dedekind's possible objects of thought

Unable to deduce the existence of an infinite set from its concept, we should look to direct arguments supporting such existence. While straightforward, this line has been largely ignored in recent debate. The most promising attempts date back to Dedekind's "The nature and meaning of numbers" (1888/2007). Dedekind doesn't include an axiom of infinity. While fully aware of how important it is whether infinite sets exist, Dedekind thought he can *prove* their existence. He gave two arguments purporting to prove it. The first argument is the most famous one; it appeals to the alleged (Dedekind) infinity of my possible objects of thought and seeks to prove that there is an infinite set. The second argument is less explicit, but more to the point: there are (simply) infinitely many natural numbers, and surely natural numbers form a set, therefore there is an infinite set. Both arguments fail unless additional assumptions are made, as I will now try to show.

Here is Dedekind's first argument (1888, par.66):

> My own realm of thoughts, i. e., the totality *S* of all things, which can be objects of my thought, is infinite. For if *s* signifies an element of *S*, then is the thought *s*', that *s* can be object of my thought, itself an element of *S*. If we regard this as transform φ(*s*) of the element s then has the transformation φ of *S*, thus determined, the property that the transform *S*' is part of *S*; and *S*' is certainly proper part of *S*, because there are elements

point of finite arithmetic (p.22) is that "all of traditional elementary number theory can be formulated in it [even though] in this system the elements are to be thought of as finite sets."

> in *S* (e. g., my own ego) which are different from such thought *s*' and therefore are not contained in *S*'. Finally it is clear that if *a*, *b* are different elements of *S*, their transforms *a*', *b*' are also different, that therefore the transformation φ is a distinct (similar) transformation (par.26). Hence *S* is infinite, which was to be proved.

Dedekind's line of thought is: let us find a possible object of thought. Suppose I think a thought *s* (say, a thought that 2+2=4, or a thought about the number 4). I could also think the thought *that I could think s,* or *think about thinking about s.* I could also think the thought *that I could think that I could think s,* or *thinking about thinking about thinking about s*. And so on. There seems to be no limit to the possible objects of my thought. Does this entail that the set of my possible objects of thought is *infinite*?

If we mean the simple infinity of natural numbers, then I would have to iterate the *I could think* operator $\aleph_0$-many times. There is no reason to believe (nor to deny) a thought or an object of thought so complex *could* ever exist. So the assumption that such a complex thought or object of thought *could* exist is unjustified.

If we mean Dedekind infinity, then we get the conclusion *only if* we take the operator *I could think* as being closed over the set of my possible objects of thought. Otherwise, the *etc.* above might fall short of reaching infinity. There is no reason to affirm (or to deny) that the *I could think* operator is closed over the totality of my possible objects of thought. So the assumption that it is closed is unjustified (where the possibility is metaphysical, as above). Moreover, in order for an operator to be closed on a set, that set *has to* exist, otherwise the closure of the operator is not defined. And Dedekind has provided no reason to believe my possible objects of thought *do* form a set (as Zermelo 1908, fn.8 observes).[14]

14 Indeed, Dedekind uses a variety of terms in roughly equivalent ways: system, collection, class. Some others might be added: plurality, set, aggregate, etc. Belonging to the pre-Russell era of set

Neither the simple nor the Dedekind concept of infinity warrant saying my possible objects of thought form an infinite set; Dedekind's argument in paragraph 66 fails.

7. Numbers and sets

A first question one might ask (Zermelo 1917, p.300) about Dedekind's argument from "the realm of thought" is why it is necessary at all. Isn't the thought that 0, 1, 2, etc. form an infinite set of natural numbers enough?

It depends on how the thought is fleshed out. Take the following obvious route. For each natural number *n* we can form its singleton, by applying the axioms of pairing and extensionality. We can then do the union of all these singletons. Obviously, the union is not a simply finite set. (For suppose it was: then there would be a simply finite number *m* of elements in the union, we could build $\{m\}$, and the union would now have $m+1$ elements. Contradiction.) Therefore, the union is a simply infinite set. In conversation, this is the reasoning most set theory students think of first. It is short, clear, and persuasive.

And yet there are several problems with this line of thought. First, suppose there indeed were a set which is not simply finite. Does it follow it is simply infinite? Dedekind writes (par.160): "A system Σ is either finite or infinite, according as there does or does not exist a system Z_n similar to it." Dedekind's statement in par.160 is too quick. Here is why.

We can *define* the concept of a Dedekind finite set as the concept of a set for which there is no one-to-one correspondence between that set and a proper subset of it. We can *define* the concept of a simply finite set as the concept of a set for which there is a one-to-one correspondence with a simply

---

theory, Dedekind could not suspect these terms would subsequently be used differently. But in the post-Russell usage, it is not clear why my possible objects of thought form a set.

finite integral domain Z_n (Dedekind 1888, par.119). From these definitions, it follows that a set is either Dedekind finite or Dedekind infinite, and that a set is either simply finite or simply infinite. But nothing can be brought into existence by definition *alone*. From the fact that there is no finite set that satisfies condition ψ, it does *not* follow that there is an infinite set that does satisfy condition ψ. We should not confuse *sentential negation* and *set complementation.* From the statement that it is not the case that the set of all natural numbers is simply finite, it doesn't follow the set of natural numbers is simply infinite.

The reasoning can be repaired. If the axiom of (countable) choice is assumed, then any two (countable) cardinals are comparable (Potter 2004, pp.161,164). It follows that a cardinal, if it exists, is either finite or infinite. I quibbled earlier about identifying cardinals with sets, but let me grant that now. Take the set formed by the earlier procedure. Since it was not finite, it follows from the axiom of choice that it is an infinite set.

With this repair, the reasoning is better but less attractive, because adopting countable choice would be a poor strategy in justifying the axiom of infinity. For one, it seems to get the explanatory order wrong: the finitist could contrapose and deny that the axiom of countable choice is justified, precisely because it licenses the surreptitious introduction of infinite sets when their very existence was in question. Second, and more importantly, the axiom of choice has consequences more doubtful than infinity, therefore cries out for justification itself.[15]

---

15 For instance, Skolem (1922, p.300) objects "sets are *not* generated *univocally* by applications of" "the principle of choice." Many choice sets exist as long as not all sets chosen from are singletons. No concept or property is needed to uniquely define the choice set, which runs against the "demand that every set be definable." Along with Skolem, Fregeans should be worried about this. Another worry, more widely shared, is that the Banach-Tarski paradox follows from the axiom of choice (e.g., Maddy 2011, p.33).

This reasoning might be better than before, but it still isn't valid! Let's repeat the procedure. We take natural numbers and form their singletons. Then we do the union of the singletons and a set results. Stop here and ask: what guarantees that we can do this union? Assume, as before, the axiom of choice, so that a set is either finite or infinite. Then the set of sets we apply the union to is either finite or infinite. If that set (of singletons of natural numbers) is finite, doing a union over it results in a finite set. But we know it cannot be finite, because there is no highest natural number. Under the axiom of choice, that set (of singletons of natural numbers) has to be infinite. Now we know where the reasoning fails. The procedure for constructing an infinite set by doing the union over singletons of natural numbers presupposes that the union operation can be applied on a set that is infinite, and it presupposes that such a set is available (the set of singletons of natural numbers). That is, it presupposes what it supposedly proves: that there is an infinite set. So the argument is circular.

It seems number theory does not single-handedly justify the existence of an infinite set after all. True, if we added the assumption that there are as many sets as numbers and functions, we would get the result that there are infinite sets. In the present context, however, doing that would be question-begging, as explained when contrasting Russell and Zermelo's outlooks in section 5.

8. Three traditions

We can now drive the point home. Of the three main epistemological traditions – foundationalism, coherentism, and pragmatism – none helps justify the axiom of infinity.

Take foundationalism first. The foundation of our knowledge of infinity is not inferential. I have just surveyed the most common argument in section 7 and found it lacking. Dedekind deserves credit for clearly seeing the axiom of infinity needs to be justified, but his argument has also been questioned in section 6. And, naturally, the axiom of infinity cannot be proven from other set-theoretic axioms.

Indispensability arguments (sourced in other branches of mathematics or mathematical physics) for the existence of infinite sets were briefly criticized in section 5.

Nor is the foundation of our knowledge of infinity *non*-inferential. In section 2, I argued that intuitions about infinite sets conflict. As Boghossian (2003) notes, their source is either mysterious or presupposes what they are supposed to establish. Our intuitive thinking-processes about infinite sets are unreliable, since intuitions about things as crucial as whether any concept has a set as an extension may turn out to be false. (For discussion, see Pantsar 2015 and footnote 15 above.) So we have neither an inferential nor a non-inferential foundational justification for thinking infinite sets exist.

What about coherentism? It might be thought that the axiom of infinity is so central to our mathematical knowledge that a coherentist justification is available (Quine 1951). The problem is with identifying the relevant sense of coherence. If coherence is a probabilistic relation (Douven and Meijs 2007), it cannot help justify the axiom of infinity because the axiom is either necessarily true or necessarily false. If coherence is an explanatory relation, then it is far from clear why more entrenched theories like the calculus need set theory to explain anything, or why representing them in set theory would constitute explanatory progress. Both lines of thought are discussed in section 4. If coherence is neither a probabilistic nor an explanatory relation, it is hard to see what kind of relation it is. So we haven't yet been provided with a coherentist justification for the axiom of infinity.

Pragmatism seems to have the best chances, with its replacement of absolute justifications with employment in problem-solving. As James (1907, Lecture II) explains, “theories thus become instruments”. After all, there are numerous problems involving infinity, and dealing with them might be better approached in a unified framework like set theory. The worry here is that assuming that infinite sets exist fares no better, in point of problem-solving, than *if-thenism* (e.g., Carnap 1939, p.45). For any problem solved with the help of the axiom of infinity, the finitist can solve it by reasoning hypothetically: *if* infinity sets exist, *then* so-and-so follows. Crucially, the hypothetical reasoning doesn't presuppose that the axiom of infinity is justified. Its potency in solving problems doesn't justify the axiom of infinity either.

9. Conclusion

Current set-theoretical practice takes an *axiom* of infinity as part of the axiomatic approach to set theory. Before an axiom of infinity was felt to be necessary, Dedekind had attempted to prove the existence of an infinite set. But his 1888 argument is unsuccessful. The natural lesson was drawn early enough by Russell (1908) and Zermelo (1908): one cannot *prove* the existence of an infinite set, so an axiom to that effect is needed. (1)-(5) are formulations of the axiom of infinity. The axiom cannot be proven from other set-theoretic axioms. So much is old news.

This essay mobilized old and well-trodden mathematical resources to make a basic philosophical point. The essay contributes an answer to the following question: Is there any good *argument* one can advance to justify the axiom of infinity? Of several justifications proposed, I have argued that none is tenable. Even though his argument based on the objects of thought fails, Dedekind's approach was right: the axiom of infinity needs justification. It currently lacks one.

Despite its simplicity, the point has, for some time, stopped being made in its blunt obviousness. Earlier writers on this issue, including the founding fathers, were keenly aware of the problem. The debate moved on and the problem was buried rather than addressed. More recently, with Maddy's (1988 and onwards) program of justifying set-theoretical axioms, we find ourselves in a curious situation: that project is only compelling if we can justify its starting point – the axiom of infinity. Yet that is precisely what is missing.

I end on a cautionary note. I have suggested that mathematical practice by itself doesn't help justify the axiom of infinity; the axiomatic method similarly fails to justify the axiom; so does application of mathematics to empirical science. What, then, could do the trick? The axiom of infinity lacks justification, one may protest, because it is *impossible* to provide one given the strictures imposed on what should count as an adequate justification. I am sympathetic to the epistemological conundrum a friend of the axiom might find herself in. In reply, I issue a challenge: if some of the suggestions I made against attempts to justify the axiom are not appropriate, *which ones and why so?* Answering these questions would, at one blow, help address the epistemological problem I raise concerning the axiom of infinity, and help advance the project of justifying it.

It is natural to think this diagnosis doesn't fit the facts well. To think an axiom foundational of the whole of mathematics fails to garner the support it needs is to undermine foundations in full generality. Unhinged skepticism would seem to ensue. That, however, is far from my intentions. So it might be more prudent to cast the moral of this text as pointing to a radical *mismatch* between set-theoretical practice and the epistemology by the lights of which we evaluate it. Set-theoretical practice *does* often rely on formal systems, trained intuitions, and bottom-up support for set theory from other branches of mathematics. Yet that fits poorly with the epistemic norms we might bring to bear in order to provide a verdict of justification – or lack thereof – of set-theoretical axioms.

A dilemma thus arises: should we change the epistemic norms, or should we reform practice? Or is there, perhaps, a solution in between which does a bit of both? This text doesn't make grand recommendations of the sort. But it does a crucial bit of preliminary work: it fleshes out the conflict implicit in believing – mistakenly – that the axiom of infinity is justified by the lights of current epistemic standards.

References:

Baker, A. (2005) Are There Genuine Mathematical Explanations of Physical Phenomena? *Mind* 114

(454):223-238.

Boghossian, P. (2003) Blind Reasoning. *Proceedings of the Aristotelian Society,* Supplem. Vol.77 (1):225–248.

Boolos, G. (1971) The Iterative Conception of Set. *Journal of Philosophy* 68 (8):215-231.

Brouwer, L.E.J. (1913/1983) Intuitionism and formalism. In: H. Putnam & P. Benacerraf (eds.) *Philosophy of mathematics* (2nd ed., pp.77-89). Cambridge: Cambridge University Press.

Carnap, R. (1939/1983) The Logicist Foundations of Mathematics. In: H. Putnam & P. Benacerraf (eds.) *Philosophy of mathematics* (2nd ed., pp.41-51). Cambridge: Cambridge University Press.

Carnap, R. (1950) Empiricism, Semantics, and Ontology. *Revue Internationale de Philosophie* 4 (11): 20-40.

Cohen, P. (1966) Examples of formal systems. Primitive recursive functions. In: *Set theory and the continuum hypothesis* (pp. 20-32). New York & Amsterdam: W.A. Benjamin.

Dedekind, R. (1888/2007) The nature and meaning of numbers. In: *Essays on the theory of numbers* (pp. 14-58). 1st English ed. (1901) Chicago: Open Court; 2nd ed. ebook digitized by Project Gutenberg, available at: http://www.gutenberg.org/ebooks/21016

Douven, I., & Meijs, W. (2007) Measuring Coherence. *Synthese* 156 (3):405 - 425.

Hauser, K., & Woodin, W.H. (2014). Strong Axioms of Infinity and the Debate About Realism. *Journal of Philosophy* 111 (8):397-419.

Heil, J. (2003) Introduction. In: *From an Ontological Point of View* (pp. 1-14). OUP.

James, W. (1907/1981) *Pragmatism.* Hackett.

Kelly, Th. (2010) Peer disagreement and higher-order evidence. In: R.Feldman & T.Warfield (eds.), Disagreement (pp.111-174) OUP.

Lavers, G. (2016) Carnap's surprising views on the axiom of infinity. Metascience 25, pp. 37-41.

Maddy, P. (1988) Believing the Axioms - I. *Journal of Symbolic Logic* 53 (2), pp. 481-511.

Maddy, P. (1992) Indispensability and Practice. *Journal of Philosophy* 89 (6):275-289.

Maddy, P. (2011) The Problem. In: *In Defense of the Axioms*. OUP.

Pantsar, M. (2014) An Empirically Feasible Approach to the Epistemology of Arithmetic. *Synthese* 191 (17):4201-4229.

Potter, M. (2004) *Set theory and its philosophy*. Clarendon: Oxford University Press.

Potter, M. & Smiley, T. (2002) Recarving content. *Proceedings of the Aristotelian Society* 102 pp. 351-354.

Quine, W.V.O. (1948) On What There Is. In: In: *From a Logical Point of View* (pp.1-19). Cambridge, Mass.: Harvard University Press.

Quine, W.V.O. (1951/1999) Two Dogmas of Empiricism. In: *From a Logical Point of View* (pp.20-46). Cambridge, Mass.: Harvard University Press.

Quine, W.V.O. (1960) Ontic decision. In: *Word and Object* (pp.233-276). Cambridge, Mass.: MIT.

Resnik, M.D. (1981) Mathematics as a Science of Patterns: Ontology and Reference. *Noûs* 15 (4):529-550.

Robinson, A. (1969) From a formalist's point of view. *Dialectica* 23, pp. 45-49.

Russell, B. (1902/1967) Letter to Frege. In: J. van Heijnoort (ed.), *From Frege to Gödel: a source book in mathematical logic, 1879-1931* (pp.124-126). Cambridge, Mass.: Harvard University Press.

Russell, B. (1908) Mathematical Logic as Based on the Theory of Types. *American Journal of Mathematics* 30 (3), pp. 222-262.

Russell, B. (1919/2010) The Axiom of Infinity and Logical Types. *Introduction to Mathematical Philosophy* (pp.106-115). London: George Allen & Unwin.

Skolem, Th. (1922/1967) Some remarks on axiomatized set theory. In: J. van Heijnoort (ed.), *From Frege to Gödel: a source book in mathematical logic, 1879-1931* (pp.290-301). Cambridge, Mass.: Harvard University Press.

Sober, E. (1993) Mathematics and Indispensability. *Philosophical Review* 102 (1):35-57.

Vineberg, S. (1996) Confirmation and the Indispensability of Mathematics to Science. *Philosophy of Science* 63 (3):263.

von Neumann, J. (1925/1967) An axiomatization of set theory. In: J. van Heijnoort (ed.), *From Frege to Gödel: a source book in mathematical logic, 1879-1931* (pp.393-413). Cambridge, Mass.: Harvard University Press.

Zermelo, E. (1908/1967) Investigations in the foundations of set theory - I. In: J. van Heijnoort (ed.), *From Frege to Gödel: a source book in mathematical logic, 1879-1931* (pp.199-215). Cambridge, Mass.: Harvard University Press.