Enumerative Induction in Mathematics

Alan Baker

Abstract: In my 2007 paper, "Is There a Problem of Induction for Mathematics?" I rejected the idea that enumerative induction has force for mathematical claims. My core argument was based on the fact that we are restricted to examining relatively small numbers, so our samples are always biased, and hence they carry no inductive weight. In recent years, I have come to believe that this argument is flawed. In particular, while arithmetical samples are indeed biased, my new view is that this bias actually strengthens the inductive support that accrues from them. The reason is that small numbers typically provide a more severe test of general arithmetical claims due to the greater frequency of significant properties and boundary cases among such numbers. In this paper, I describe and defend this new view, which a call Positive Bias Pro-Inductivism.

Keywords: Enumerative induction, Arithmetic, Non-deductive reasoning

1. Introduction

In 1919, Hungarian mathematician George Polya was pursuing questions concerning prime factorization and began noting down, for each of the first few natural numbers, whether a given number had an even number of prime factors or an odd number of prime factors. What Polya was counting was not distinct prime factors, but just the raw number of prime factors. So, for example, $10 = 2.5$ has an even number of prime factors, 11 has an odd number, and $12 = 2.2.3$ has an odd number. Next, Polya recorded, for each given number, n, how many numbers in the set $\{1,2,3,\ldots n\}$ have an odd total number of prime factors and how many have an even total number. Looking at all the numbers up to 100, Polya noticed that at every point at least half of members of the set $\{1,2,3,\ldots n\}$ had an odd total number of prime factors. Intrigued, Polya extended his calculations up to n = 1,500 and found that the pattern continued to hold. At this point, Polya conjectured in print that this holds universally. In other words, for all n, at least half of the numbers less than n have an odd total number of prime factors. This came to be known as *Polya's Conjecture* (PC).

Polya presumably made the conjecture because he believed that it was likely to be true. What grounds did he have for this belief? According to the narrative outlined above, Polya's belief in the truth of PC was based exclusively on enumerative induction. 1,500 instances from a larger domain were examined and found to fit the hypothesis, and on this basis the hypothesis was conjectured to hold universally. As such, this seems like a particularly pure case of enumerative induction, with few extraneous complicating factors. But was Polya *rational* in taking the results of the first 1,500 cases to lend support to the universal hypothesis in this way? More generally, can enumerative induction lend genuine support to mathematical claims? It is these questions that I shall be taking up in what follows.

Alan Baker, Swarthmore College, USA, abaker1@swarthmore.edu, ORCID:0000-0003-2896-7961

Journal for the Philosophy of Mathematics, Vol. 1, 5-21, 2024 | ISSN 3035-1863 | DOI: 10.36253/jpm-2931

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.
Competing Interests: The Author(s) declare(s) no conflict of interest.

Firenze University Press | https://riviste.fupress.net/index.php/jpm

2. Scene Setting

Before proceeding it may be helpful in clarifying the scope of our inquiry to distinguish the questions I am asking from some other, related (and also interesting) questions, and also to make explicit several presuppositions that I will be taking for granted.

Firstly, I am not asking whether enumerative induction in mathematics can yield *knowledge* of general mathematical claims. My focus here is on justified belief, not knowledge, and I will be taking no stand on the potential link between enumerative induction and knowledge. Secondly, I am not asking whether enumerative induction in mathematics can (or should) ever be on a par with *deductive proof*. As it pertains to mathematical methodology, everything I say is compatible with deductive proof continuing to be the gold standard for demonstrative mathematical reasoning.[1] Thirdly, I am not asking whether enumerative induction can provide *compelling* grounds for belief in a general mathematical claim. Rather, I am interested in whether enumerative induction can achieve the more modest goal of providing some positive support.

As for presuppositions, the work being done in this paper will be carried out against the background of a general anti-scepticism about induction. In other words, I will be presuming that enumerative induction in the empirical sciences is generally (other things being equal, *ceteris paribus*) a rationally acceptable tool for acquiring justified belief. Also, although I will continue to talk sometimes about "enumerative induction in mathematics" my focus will almost exclusively be on enumerative induction in number theory. The examples I will be discussing will all involve claims about the totality of the natural numbers. Putting together these two points, the question I asked at the end of Section 1 can be precisified as follows: is there something distinctively problematic about using enumerative induction to bolster beliefs in arithmetical claims?

3. An Argument Against Enumerative Induction in Mathematics

In my 2007 paper, "Is There a Problem of Induction for Mathematics?," I rejected the idea that enumerative induction has force for mathematical claims. More specifically, I argued for the following two theses, one normative and one descriptive:[2]

[EIM Norm] Enumerative induction *ought not* to increase confidence in a universal mathematical generalization (over an infinite domain).

[EIM Desc] Enumerative induction *does not* (in general) increase confidence in a universal mathematical generalization (over an infinite domain).

I am not an inductive sceptic in a more general sense, so my argument for EIM Norm rest on features that I take to be peculiar to the mathematical case. These distinctive features can also be framed as *disanalogies* between pure mathematics and the empirical sciences.

In increasing order of importance, the three main disanalogies that I pointed to are the following:

(i) The domain of number theory is infinite.

The challenge here for enumerative induction is clear enough. When dealing with an infinite domain, such as the natural numbers, however big our finite sample of instances of a given universal claim, it will still only represent an infinitesimally small portion of the whole domain. What is less clear is how much difference this makes for enumerative induction in mathematics as compared to the empirical sciences. Firstly, it may be that the universe is in fact infinitely large, in which case fully general scientific claims would also have an infinite domain. Secondly, even if the universe is not infinitely large, most general empirical claims seem to presuppose some sort of indefinitely open-ended domain, and this may be not much different in its implications for enumerative induction than an absolutely infinite domain.[3] In summary, while the presence of an infinite domain certainly poses a problem, it is not enough on its own to justify abandoning enumerative induction in mathematics while retaining it in the empirical sciences.

[1] Relatedly, I am not advocating for any relaxation of the notion of theoremhood, whereby a mathematical conjecture could be elevated to the status of "theorem" given sufficient enumerative inductive support.

[2] Baker (2007, p.73).

[3] Note that even scientific claims that have an implicit spatial limitation – for example, many biological claims that are about how genetic code, metabolism, etc. work among organisms found on Earth – are likely to be temporally open-ended.

(ii) The domain of number theory is non-uniform.

Here I drew explicitly from Frege's discussion of induction in the *Grundlagen*. The following three quotes from that work give the flavor of Frege's position:

> [In mathematics] the ground [is] unfavorable for induction; for here there is none of that uniformity which in other fields can give the method a high degree of reliability.
>
> An even number can be divided into two equal parts; an odd number cannot; three and six are triangular numbers, four and nine are squares, eight is a cube, and so on.
>
> In ordinary inductions we often make good use of the proposition that every position in space and every moment in time is as good in itself as every other. ... Position in the number series is not a matter of indifference like position in space.[4]

Following Frege, I argued that the domain of the natural numbers is *non-uniform*. In fact, the above quotes suggest two different scales on which this non-uniformity might manifest. At the scale of individual numbers, if we arbitrarily pick two separate numbers, m and n, just because m has property P gives us no *prima facie* reason to think that n also has property P. (For example, just because m is a triangular number gives us no reason to think that n will also be a triangular number.) But there is also, potentially, non-uniformity at the scale of different segments of the number line. For example, consider the sequence of numbers between 200 and 300 (call this S_1), and the sequence of numbers between 10^{10} and $10^{10}+100$ (call this S_2). Next consider some property, P^*, that holds for every member of S_1, or which holds collectively for the set. Each of the two sets S_1 and S_2 consists of 100 consecutive numbers. So should we expect P^* also to hold for S_2? Not necessarily, and in fact there are many properties that we know to vary according to where in the number line we are sampling. For example, every number in S_1 has eight or fewer prime factors, but this is not true of every number in S_2. And at the level of collective properties, the density of primes, of square numbers, of Fibonacci numbers, and so on are all significantly lower in S_2 than in S_1.

The fact that the properties of one number are not a reliable guide to the properties of other numbers, and similarly for groups of numbers, is undeniably a challenge for the effectiveness of enumerative induction in supporting general claims about all numbers. What is less clear is whether it is enough on its own to justify Frege's wholesale rejection of enumerative induction in an arithmetical context. After all, there is plenty of variability in many of the empirical domains in which enumerative induction is employed, yet this is not considered a fatal flaw. If the issue is simply non-uniformity, then it can be counteracted by selecting a diverse array of instances to examine. If the conjectured hypothesis holds across this diverse array, which will presumably exemplify the same kind of non-uniformity as the domain as a whole, then this enumerative inductive evidence may still provide genuine support for the hypothesis.

(iii) We can only sample from among the very small numbers.

The infinite size and the non-uniformity of the domain of the natural numbers are not enough on their own to undermine the effectiveness of enumerative induction in mathematics. However, they do become enough when combined with a third feature, which concerns how and where we are sampling the domain.

The basic point is straightforward enough. Whether we are talking about the instances initially collected by Polya in connection with his conjecture, PC (i.e. all the numbers up to 1,500), or the current range of instances checked in connection with the Goldbach Conjecture (i.e. currently, all the even numbers up to about 4×10^{18}), all the instances that we can ever feasibly collect for any general arithmetical conjecture are in an important sense *small*. In my (2007), I introduced the term "minute" to make this latter notion more precise, defining a natural number to be *minute* if it can be written down (by us) in ordinary decimal notation, including exponentiation (but not iterated exponentiation). We are practically limited to searching among the minute numbers for instances of general arithmetical claims.

4 Frege (1884, Section 10).

So what? Even if we are in practice restricted in this way, what relevance does this have to the effectiveness of enumerative induction? This is a fair question, and – as with the first two distinctive features of mathematics discussed earlier – it could justifiably be claimed that on its own it is not a serious problem. However it becomes a problem, or so I argued, when combined with the two earlier features. The core argument of my 2007 paper can be summarized as follows:

(1) All instances of general arithmetical conjectures that we can feasibly check are drawn from among the minute numbers.

(2) Hence, there are an infinite number of non-minute numbers, none of which we can ever sample.

(3) Size is known to make a difference to the properties of individual numbers and of sets of numbers.

(4) Hence, any sample used as a basis for enumerative induction in arithmetic is unavoidably biased.

(5) Hence, enumerative induction in arithmetic carries (and ought to carry) no weight.

Returning to the example that we began with, the above argument suggests that Polya's examination of the first 1,500 natural numbers provides no substantive support for his conjecture, PC, because there is no reason to think that this initial sample is representative of the number line as a whole.

My 2007 view on induction in mathematics can be seen as a targeted form of inductive scepticism. Paseau (2021) uses the term "size scepticism" to refer to this position, and helpfully distinguishes three different types of size scepticism. According to Paseau's classification scheme, Baker (2007) is a defense of *u-scepticism* (the "u" here stands for "unrepresentative").[5] Sampling only from the minute numbers amounts to not varying the sampling space along an axis (in this case size) that is *known* to make a difference. So, in this sense, any basis for enumerative induction in mathematics is (unavoidably) *unrepresentative*.

It is worth noting that the conclusion of my 2007 paper is quite radical (though not historically unprecedented, as seen from the remarks of Frege quoted above), for my anti-inductive thesis is not merely that enumerative induction fails to provide compelling evidence for general arithmetical claims, but that it provides no supporting evidence whatsoever. This 'no-support' thesis can best be made sense of as arising not just from the unrepresentativeness of our sample set but also from the infinite number of objects in the domain (as alluded to in step (2) of my core argument). In drawing exclusively from the minute numbers, we end up with a sample set that is unrepresentative of the domain as a whole and is also an infinitesimally small portion of that whole. If either of these two obstacles was removed then there would be the possibility of enumerative inductive evidence yielding substantive support for general conjectures. A representative sample would give grounds for projecting the partial result across the whole domain. And a non-infinitesimal sample would raise our subjective probability for the general conjecture by verifying a substantive proportion of cases.

4. Bolstering the Biased-Sample Claim

In recent years, I have come to believe that the core argument from my 2007 paper is flawed, and that inductive scepticism with respect to mathematics is unwarranted. Most of the criticism of my core argument has focused on undermining the intermediate conclusion, (4), that our sampling from the domain of the natural numbers is unavoidably *biased*.[6] Implicit in such attacks is the presumption that the final step in the argument, from

[5] This is to be distinguished from *s-scepticism* (size scepticism), according to which small numbers are distinctively problematic. i.e. sampling 1,000 non-small numbers would provide better evidence than sampling 1,000 small numbers. And it is also to be distinguished from *c-scepticism* (comparative scepticism), according to which a set of larger numbers provides better evidence than a set of smaller numbers.

[6] See e.g. Walsh (2014), Waxman (2017). One common line of objection is that in gathering confirming instances for a general scientific conjecture we are also stuck with potentially biased samples, since all instances that we observe must be "close" to us (in space and in time). But induction in empirical contexts is (by presumption) acceptable. Hence it cannot be this "bias" that leads to problems in the mathematical context. However, this objection ignores the fact that the "close to us" bias is very different from the "close to the beginning of the number line" bias. Firstly, being "small" is an objective property of numbers, while being "close to us" is (manifestly) observer-relative. Secondly, we have *antecedent* reason to believe that smallness can make a systematic difference to the behavior of numbers, whereas we have no antecedent reason to believe that being close to us in space and/or time makes a systematic difference to the behavior of physical objects.

sample bias (step (4)) to lack of inductive support (step (5), is unproblematic. My own approach will be rather different. I continue to believe that my 2007 argument for sample bias is convincing. However, I want to resist the inference that sample bias automatically undermines the force of enumerative induction. To this end, my strategy will be as follows. Firstly, I will sharpen and defend the biased-sample thesis by focusing on an initial subset of the natural numbers that is radically smaller than the set of minute numbers. Secondly, I will argue that the particular kind of bias that afflicts this initial subset actually serves to *strengthen* inductive inferences that are based on it. The counterintuitive-sounding takeaway is that bias in our sampling of the natural numbers provides a rational boost to enumerative induction applied to general arithmetical claims in a way that helps to overcome the challenges of generalizing over a demonstrably infinite domain.

The first task is to formulate and defend a more circumscribed version of the biased-sample claim. I shall begin by zeroing in (no pun intended) on a much smaller initial subset of the natural numbers that features in my core argument. In my 2007 paper, I focused on the "minute numbers," where a natural number is defined as *minute* if it can be written down by us in ordinary decimal notation, including exponentiation (but not iterated exponentiation). It is worth noting that, despite the label, the minute numbers include a lot of numbers that are very far from being small in any ordinary sense. For example, current estimates put the number of atoms in the universe at something on the order of 10^{80}. Included among the minute numbers are numbers vastly larger than this, such as 10^{800}, 10^{8000}, and so on. My reason for placing such a relaxed bound on the minute numbers was to make absolutely sure that all putative inductive evidence for unproven general conjectures consists exclusively of instances that count as minute. However, for current purposes this is overkill, and in fact the liberal characterization of minute number has a negative impact on the core argument. I consider these two points separately below.

How is this overkill? The fact is that in the vast majority of actual cases where mathematicians gather instances of a general arithmetical conjecture, these instances are not especially large. Historically, when instances were calculated by hand, these rarely exceeded a few thousand (in number, and in magnitude), and this persisted until the advent of the digital computer.[7] Once computers could be used to speed up this checking process, the sampling size massively increased. But even now, the limitations of computing speed and memory place fairly tight bounds on how many cases can feasibly be calculated. Consider the oft-cited example of the Goldbach Conjecture, which has been checked by computer for all even numbers up to around 4×10^{18}. This is among the larger sample sets for any currently open arithmetical conjecture, yet it is nowhere close to pushing the bounds of the minute numbers (which, recall, include anything we can feasibly write down using decimal notation and non-iterated exponentiation, so 10^{180} , 10^{1800}, and 10^{18000} are all minute numbers). The upshot is that we could focus on a much, much smaller set of numbers than the minute numbers and still capture nearly all actual cases of putative enumerative induction in mathematics.

Even if the above point is true, it might seem that there is still no harm in using the minute numbers as a basis for the core argument. However, it turns out that there is a clear downside, namely that referencing an excessively large set of numbers risks undermining the biased-sample claim itself. There are a couple of ways in which this happens.

Recall that the basis of my claim that our sampling of the natural numbers is (unavoidably) biased is that the sample space is unrepresentative of the numbers as a whole. So if the focus is on the minute numbers, then this amounts to the claim that the minute numbers behave differently than the non-minute numbers. The (Frege-inspired) intuition that position along the number line matters to what properties a number is likely to have was motivated, back in Section 3, by comparing the sequence of one hundred numbers between 200 and 300 (call this S_1) with the sequence of one hundred numbers between 10^{10} and $10^{10} + 100$ (call this S_2), and noting the sharp difference in the relative occurrence of properties such as being prime, square, Fibonacci, and so on. Notice, however, that both of these sequences fall squarely within the set of minute numbers, so this observation does nothing to support the claim that the minute numbers are themselves systematically different from the non-minute numbers. If instead we take a set of one hundred numbers from further along the sequence of minute numbers and compare it with a set of one hundred non-minute numbers, it is much less clear that the

7 The Polya Conjecture example discussed in Section 1, in which cases were checked up to 1,500, is very typical in this respect.

'systematic difference' claim holds up.[8] The basic problem is that the set of minute numbers includes numbers of such magnitude that maintaining that all minute numbers are unrepresentative becomes much less plausible.

A second problem with the excessive magnitude of the minute numbers is that it becomes harder to find examples where a general conjecture has held for all minute numbers but ended up failing for some non-minute number. In my (2007) paper, I came up with just one putative example where this may happen, and even this is conditional on the unlikely eventuality that the Riemann Hypothesis is false. Under this assumption, the upper bound of the point at which the logarithmic density function is eventually exceeded is given by the second Skewes number, which is a non-minute number. However, as D'Alessandro points out, this lone example no longer holds up.[9] A result of Saouter and Demichel, proved in 2010, shows that the upper bound is at most a little over 10^{316}, which is well within the realm of the minute numbers. The difficulty of finding actual examples of conjectures for which the minute numbers behave differently from the non-minute numbers puts further pressure on my thesis that the minute numbers are "unrepresentative" of the numbers as a whole.

I propose to cut through the difficulties associated with the minute numbers by restricting attention to a drastically smaller initial segment of the natural numbers, and focusing on the biased-sample thesis as applied to this smaller subset. Consider the natural numbers from 1 to 1,000. At risk of multiplying terminology beyond necessity, I shall refer to the numbers in this set as *tiny numbers*. And I hope to make plausible the following two theses associated with this set:

(A) A sample of positive instances selected only from among the tiny numbers is biased, and is unrepresentative of the domain of natural numbers as a whole.

(B) A sample of positive instances of a general conjecture, C, that includes all of the tiny numbers provides substantive support for C, other things being equal, *in virtue* of the fact that the sample is biased.

The broad arguments for the more general thesis (A) have already effectively been made back in Section 3. Now that we are restricting attention to just the first 1,000 numbers, it is very plausible to maintain that sequences of numbers from this set look very different in their individual properties, and distribution of properties, than sequences of much larger numbers. In the next section, I will turn to defending thesis (B), firstly by highlighting three specific ways in which the tiny numbers are unrepresentative, and then by showing how each of these biases actually *increases* the level of inductive support that tiny numbers provide.

5. From Bias to Inductive Support (I): Unrepresentative Representativeness

Bias in a sample is not usually considered to be a positive feature for inductive inference. However, certain specific types of bias can boost inductive support, or so I shall argue. Counterintuitive though this sounds, it is easy enough to think up examples where this happens even in a non-mathematical context. Consider some claim about the upper physiological limits on human strength or endurance, and imagine that we test this claim by measuring the capacities of 1,000 professional athletes and finding that they all fall within the conjectured limit. Clearly this is a biased sample, since professional athletes are not representative of human beings as a whole. However, it is equally clear that in this case the bias ought to strengthen our belief in the physiological limit conjecture more than if we had picked 1,000 people at random from the general population. This is because it is *prima facie* more likely for a professional athlete to exceed a postulated performance limit. So here the bias of the sample serves to boost the inductive support that it provides.

Returning to the mathematical case, I will highlight three features that manifest differently among the tiny numbers in comparison to arbitrary larger numbers. For each feature, I will begin with a characterization of what the feature is, then show how the feature plays out differently for the tiny numbers, and finally argue that this difference boosts the inductive force of samples involving the tiny numbers.

8 For example, consider the (minute) number $M = 10^{1,000,000}$, and compare an arbitrary set of minute numbers, $\{M+1, M+2, \ldots, M+100\}$, with an arbitrary set of non-minute numbers, $\{10^M+1, 10^M+2, \ldots, 10^M+100\}$.

9 D'Alessandro (2021, p.33).

The first feature involves *significant properties*. Every natural number has an infinite number of mathematical properties, but some of these properties are more significant than others.[10] What do I mean here by "significant"? I am not going to try to give a precise definition of this term, partly because I do not know how to do so and partly because the larger argument I shall be giving does not require it. For present purposes, I shall proceed with two complementary approaches to characterizing the notion of significance: a loose definition of significance, and then some contrasting cases of *in*significant properties.

Starting with the loose definition, we shall count a mathematical property as *significant* to the extent that it features in the statements and proofs of important mathematical results, and to the extent that it is systematically linked to other significant properties.[11] The boundary between significant and insignificant properties is vague, and significance is itself plausibly a matter of degree.[12] Nonetheless, it seems clear that some mathematical properties are significant *simpliciter* and others are not. Any list of canonical significant mathematical properties would presumably include properties such as being prime, being even, being square, and being a Fibonacci number.[13]

As a second aid to drawing the above distinction, it may be helpful to enumerate a few ways in which number properties can fail to be significant. The following list of kinds of insignificant property is obviously not meant to be exhaustive, but it provides a few examples:

(i) properties that relate to non-mathematical facts
e.g. the property of being the number of weeks in a year

(ii) properties that are representation-dependent
e.g. the property of having the decimal digits sum to a prime number

(iii) properties that are highly gerrymandered (or disjunctive)
e.g. the property of being either a square number or a Fibonacci

(iv) properties that are arbitrarily specific
e.g. the property of having exactly 23 distinct prime factors

Most of the canonical significant properties decrease in frequency as numbers get larger. Sometimes this decrease is fairly gradual (as with prime numbers) and other times it is more drastic (as with perfect numbers). This general pattern suggests that the tiny numbers are biased in manifesting a greater frequency of significant properties, but we need to be careful in how this claim is formulated. It is not the case, for instance, that there is a greater density of numbers with significant properties among the tiny numbers. The reason why not is that *every* number has (some) significant properties. How so? Because in addition to the kinds of properties mentioned above, there are also significant properties that do not decrease in frequency among larger numbers. Typically, these are significant properties whose complements are also significant.[14] An example of a significant property whose frequency does not decrease with number size is the property of being an odd number. This property is significant because many other properties are linked to it, and its complementary property, being an even number, is also a significant property, and also does not decrease in frequency as number size increases. Another example is the property of being composite. This is a significant property whose frequency actually increases with number size (since its complementary – and also significant – property, primality, decreases in frequency). Every number is either even or odd, and every number (apart from 1) is either prime or composite, so every number has significant properties.

[10] For example, each natural number, n, has the (infinitely large) set of properties of being $< n+1, < n+2, < n+3$, etc.

[11] In his *Mathematician's Apology*, G.H. Hardy spends some time exploring the notion of "seriousness" as a property of certain mathematical theorems and results. Hardy writes: "A 'serious' theorem is a theorem which contains 'significant' ideas and I suppose that I ought to try to analyse a little more the qualities that make a mathematical idea significant." (Hardy, 1940, p.21) Hardy ends up linking significance (of mathematical ideas) both to generality and to depth. For an interesting recent discussion of Hardy's notion of seriousness, and how initial judgements of non-seriousness may be overturned by subsequent developments in mathematics, see Weisgerber (2024).

[12] I will return to discuss the issue of *degree* of significance in Section 6 below.

[13] For discussion of the – potentially related – notion of "mathematical natural kind," see Corfield (2004) and Lange (2015).

[14] This is not the case for the majority of significant properties. For example properties such as '_ is not a square number' and '_ is not a Fibonacci number' are not significant.

The upshot of the above discussion is that the tiny numbers are not different from sequences of larger numbers in having fewer elements with significant properties. However, it is nonetheless true that most significant properties decrease in frequency with number size. As a consequence, if we compare the tiny numbers with a sequence of 1,000 consecutive much larger numbers, we should expect a greater *range* of significant properties to be instaConsider, for example, the properties of being a square number and of being a prime number. The difference between consecutive squares is $(n+1)^2 - n^2 = 2n+1$, so once we get above 500^2, there will be sequences of 1,000 consecutive numbers none of which has the property of being square. For prime numbers, the density of primes less than n is well approximated by $1/\log n$, and among numbers as 'small' as 10^{15}, sequences of 1,000 consecutive composite numbers are known to occur.[15] For both of these properties, being square and being prime, we don't have to go very much further along the number line beyond the tiny numbers before the chances of either being instantiated in a given 1,000-number sequence is extremely low. The vast majority of numbers will have the bare minimum of significant properties, being either odd and composite or even and composite, and only very occasionally will significant properties such as being prime, being perfect, being square, or being Fibonacci be instantiated.

This, then, is the sense in which the tiny numbers are biased with respect to significant properties. The range of significant properties that is encompassed by the tiny numbers radically exceeds that encompassed by similar-sized samples from among much larger numbers. The tiny numbers are *diverse* in a way that collections of larger numbers are not. What impact does this 'diversity bias' have on the role of tiny numbers in enumerative induction? The tiny numbers are atypical in this respect and so in one sense they are not "representative" of the numbers at large. However, in another sense, the tiny numbers *are* representative: the variability and density of properties instantiated by tiny numbers allows them to effectively stand in (or "represent") a much broader range of numbers.[16]

I shall use the term *unrepresentatively representative* to refer to this – initially paradoxical seeming – aspect of the tiny numbers. Normally, unrepresentativeness correlates with lack of variation. This is what underpinned my (2007) u-scepticism, and what decreases the value of unrepresentative samples, other things being equal. However, in this context the unrepresentativeness in question pertains to a precisely contrary feature, namely *increased* variation. My claim is that this makes the tiny numbers *more* effective as a basis for enumerative induction, not less. Why, exactly? If some large number has a significant property, p^*, then it is likely that there is some tiny number that is p^*. So if the large number is a counterexample to a given conjecture, C, in virtue of the number having property p^*, a corresponding counterexample to C will occur among the tiny numbers.

By way of illustration, consider the following non-mathematical analogy, which is based loosely on an actual historical example. In the 1930's, the U.S. medical establishment drew up infant growth charts in order to give guidance to doctors about when a baby's rate of growth might be a cause for concern. These growth charts gave upper and lower bounds of what counted as 'normal' weight, height, and weight to height ratio, for each age. In devising the charts, the medical authorities measured thousands of healthy infants of different ages. The surveys were carried out in the rural Midwest, where most of the infants were bottle-fed and of northern European and Scandinavian origin. How representative was this sample? In one sense it was very representative: the 'typical' or 'average' American baby in 1930 may well have been similar to the sampled population. But in another sense, the sample was not representative: there were many other sub-populations present in the U.S. at the time that were not included in the sample. Imagine a second sample of infants, of similar size to the first sample, taken from New York City. As a sub-region of the U.S. in 1930, New York City was far from typical; most other regions would have looked very different. So in our first sense above, the NYC sample is not representative. However, an important way in which New York City was different was in the diversity of its population. This diversity holds across multiple axes: ethnicity; country of origin; socio-economic class; and so on. If we focus for the moment on the geographic origin of the parents of NYC infants in 1930, it may well be the case that there was significant representation from across all 48 states of the U.S. So in this second sense, the NYC sample is representative. This "unrepresentative representativeness" of the NYC sample would have made it a better basis for drawing up widely applicable infant growth charts than the sample from the Midwest.

[15] Caldwell (2021).

[16] Are there any good examples of a significant property that is not instantiated by a tiny number, but only by some large number? If we restrict attention here to individual properties, rather than combinations of two or more properties, it is surprisingly difficult to come up with a good example of a significant property whose first instance is greater than 1,000.

6. From Bias to Inductive Support (II): Severe Tests

The second feature that makes the tiny numbers different from larger numbers is the occurrence of *boundary cases*. For any instantiated mathematical property, there is a smallest natural number with that property.[17] Without any restriction on what properties we are considering, a given number being a boundary case is of no particular import. Indeed, it is trivially the case that every number, n, is a boundary case for the corresponding property of being greater than or equal to n. But if we restrict attention to *significant* properties (in the sense characterized in the previous section), then boundary cases carry more weight.

As a starting-point, let us consider where the boundary cases for various significant mathematical properties lie. The boundary case for being prime is 2, for being even is 2, for being square is 1, for being perfect is 6, and for being a Fibonacci number is 1. These examples are all tiny numbers, indeed they are toward the very beginning of the tiny numbers. We can also consider boundary cases for *combinations* of significant properties. For example, the boundary case for being both odd and prime is 3, for being both even and square is 4, and for being both a square and the sum of two squares is 25. Not all boundary cases for significant properties (and for combinations of significant properties) are tiny numbers, but the vast majority are. In this respect, the tiny numbers are very unlike subsets of larger numbers.

What are the implications of this fact, that there are disproportionately more boundary cases among the tiny numbers, for the value for enumerative induction based on samples involving the tiny numbers? The key point here is that general claims, if they fail, often fail to hold for extreme cases. For general mathematical claims, this may not be readily apparent, because failure to hold for the smallest case often means that the claim is never seriously entertained. However, the fact that we sometimes have to modify general conjectures in order to avoid the pitfalls of boundary cases shows that this phenomenon is fairly widespread. For example, the Goldbach Conjecture (GC) is formulated as the claim that every even number *greater than 2* is expressible as the sum of two primes. 2 is a boundary case, since it is the smallest even number. And GC fails for 2, so it needs to be explicitly excluded in order for the conjecture to (potentially) be true. Or consider the theorem that every number has a unique prime factorization, which is true as long as we exclude the boundary cases of 0 and 1.[18]

Alan Hajek has suggested that it is not just in mathematics where boundary cases are more likely to yield counterexamples to general hypotheses. In a paper on "philosophical heuristics," one of the techniques that Hajek recommends wielding against philosophical claims is to check extreme cases:

> [L]ook for trouble among the extreme cases – the first, or the last, or the biggest, or the smallest, ... It is a snappy way to reduce the search space. Even if there are no counterexamples lurking at the extreme cases, still they may be informative or suggestive.[19]

The upshot of the above observations is that the prevalence of boundary cases among the tiny numbers tends to make any sample containing these numbers an unusually *severe test* of a general mathematical conjecture. It is a familiar point that samples that severely test a conjecture but then end up being positive instances of that conjecture provide stronger inductive support for the conjecture than positive instances that never really put the hypothesis at much risk.[20]

There is also a third feature of the tiny numbers that contributes to the severe testing of general hypotheses, and that arises from the feature discussed in the previous section concerning the wider range of significant

[17] There may also be a largest number with that property, if there are only a finite number of occurrences of the property, but for many mathematical properties the only boundary is at the lower end.

[18] A less well-known example, that fails at both its lower and upper boundaries, is the following: For a given number, n, assume that we have n boxes and n objects. Let Q be the property that holds of n if and only if it is possible to assign all the objects to boxes in such a way that fewer than n boxes are used and fewer than n objects are in each box. If we pick an arbitrary number, say 17, then it is easy to see that property Q holds for it. In this case, for example, we could put 13 objects in box 1, 4 objects in box 2, and leave the other 15 boxes empty. So we have used only 2 boxes ($2 < 17$), and each box contains less than 17 objects. Does Q hold for every number? No. It is straightforward to verify that Q fails to hold for the two lower boundary cases of 0 and 1. And it also fails to hold for the upper boundary case of ∞.

[19] Hajek (2014, p.292). As one example, Hajek considers testing the claim "every event has a cause" by looking at the first event as a boundary case.

[20] For example, following Hempel, both a white shoe and a black raven can be seen as positive instances of the general hypothesis that all ravens are black. However, observing a raven and determining that it is black puts the hypothesis at risk, while observing a shoe and determining that it is white does not.

properties among the tiny numbers. This wider range increases the chances of two or more significant properties being co-instantiated by a single number, and for more different combinations of significant properties to be co-instantiated. Because most significant properties become less common as the size of numbers increases, the frequency of coinstantiation (especially for properties that are not directly related) dramatically decreases beyond the tiny numbers. To give just one example, consider the question of which numbers are both square and Fibonacci. It turns out that there are only three numbers that co-instantiate these two (significant) properties: 0, 1, 144, and these are all tiny numbers.

Call a number *interesting* if it co-instantiates two significant properties such that the proportion of numbers that co-instantiate in this way is much smaller than the proportion of numbers with either property alone. So while significance is a feature of individual properties of numbers, interestingness is a feature of pairs of properties.[21]

Interestingness (like significance) is not a precisely defined notion, but it is clear nonetheless that 0, 1, and 144 are all interesting numbers, in the above sense, in virtue of exhibiting the – very rare – co-instantiation of the properties of being square and being Fibonacci. Or consider the following pair of cases:

> 2 is interesting in virtue of co-instantiating the significant properties of being even and being prime, because this co-instantiation is unique and is therefore much rarer than either being prime or being even.
>
> 3 is not interesting in virtue of co-instantiating the significant properties of being odd and being prime, because this co-instantiation is almost as common as the property of being prime *simpliciter*.[22]

As was the situation with boundary cases, interesting numbers are a potentially rich source of counterexamples to general conjectures about numbers. Interesting numbers combine significant properties in unusual ways, and as such they are more likely to be anomalous and thus to provide severe tests for universal claims.[23] The excess of interesting numbers among the tiny numbers is therefore another way in which samples drawn from tiny numbers provide disproportionately severe tests of general mathematical conjectures.

Not only is the frequency of interesting numbers much greater among the tiny numbers, in comparison to the non-tiny numbers, so is the typical *degree* of interestingness, or so I shall argue. What makes one interesting number more interesting than another? A couple of features come to mind. Consider a number, n, that has significant properties P and Q. Intuitively, the degree of interestingness of n is determined partly by the significance of P and Q, and partly by how rare the combination of P and Q is relative to these properties considered individually.[24]

In the previous section, we discussed how the range of significant properties is greater among the tiny numbers than among comparable-sized sequences of larger numbers, and how this makes the tiny numbers "unrepresentatively representative." Thinking now of significance as a matter of degree, perhaps a better way to put the above point is that the range of properties of *higher significance* is greater among the tiny numbers.[25] This latter claim seems very plausible, and although I do not have any knockdown argument to establish it, some circumstantial evidence in its favor may be gathered.

The Penguin Dictionary of Curious and Interesting Numbers is, as the title suggests, a (non-exhaustive!) catalog of numbers whose properties are deemed to be interesting for one reason or another.[26] In our terminology, the numbers listed include both those with very significant properties and those with very interesting combinations of properties.[27] We can get some sense of the relative degree of interestingness of numbers of

21 Note that in a situation in which there are two significant properties and one is a special case of the other, this will not result in the number in question being interesting. Why not? Because if all numbers with property P also have property Q then the proportion of numbers that have the twin properties (P, Q) is the same as the proportion of numbers that have property P.

22 Which is not to say that 3 is not interesting, it is just that it is not interesting in virtue of *this* combination of significant properties.

23 Consider the number 2 in relation to the Goldbach Conjecture, as was previously discussed. Arguably the reason that 2 is a counterexample (and thus needs to be excluded from the domain of application of GC) is because it (uniquely) instantiates both the property of being even and of being prime.

24 In some sense, then, degree of interestingness involves a trade-off. A pair of highly significant properties whose combination is not rare relative to the occurrence of the properties individually may result in the same level of interestingness as a pair of less significant properties that very rarely occur together.

25 One way to formulate this more precisely is as follows: for any cut-off that is made between "higher significance" and "lower significance" for mathematical properties, the range of properties above this cut-off is greater among the tiny numbers than among almost all sequences of 1,000 consecutive non-tiny numbers.

26 Wells (1986).

different sizes by looking at snapshots from different places along the number line and seeing what is listed for each sequence of numbers. For reasons of space, I will restrict this survey to looking at three different sequences of three consecutive numbers, picking the most intuitively interesting pairs of properties in each case:[28]

2 The only even prime number

3 The smallest odd prime number

4 The smallest composite number

102 The smallest 7^{th} power to be the sum of eight 7^{th} powers

103 One of a pair of twin primes (with 101)

104 A semi-perfect number (i.e. equal to the sum of all or some of its proper divisors)

For the third sequence, $\{1002, 1003, 1004\}$, we enter the realm of the non-tiny numbers. It is perhaps no coincidence that, even by this relatively early point on the number line, the proportion of numbers that get a mention in *The Penguin Dictionary* is very low, and that none of the numbers in this third sequence appear. Looking elsewhere for information related to interestingness, the following properties may be noted:[29]

1002 A sphenic number (i.e. the product of 3 distinct primes)

1003 The product of a prime, n, and the n^{th} prime

1004 A heptanacci number[30]

The difference, even just intuitively, between the degree of significance of the properties involved in these three sequences is striking. In addition, the numbers in the first sequence either exhibit combinations of core significant properties that are unique, or are boundary cases of a core significant property, and thus are also unique. This gives all three numbers in this sequence close to a maximal degree of interestingness. In the second sequence, the first number, 102, is a boundary case, but the property involved (being a 7^{th} power that is the sum of eight 7^{th} powers) involves the combination of two properties that are of low significance. The remaining two numbers in the sequence, 103 and 104, exhibit individual properties that are of moderate significance. However, in the absence of a second identified property, the only option is to combine with one of the 'default' significant properties (even, odd, composite) and this will not result in the number being interesting. By the time we get to the third sequence, which is also the first to feature non-tiny numbers, the properties involved have become even more obscure and even less significant, while also not combining with a second, non-trivial significant property to boost any prospects of interestingness. I conclude that this survey, brief and partial though it is, lends further support to the claim that the tiny numbers typically exhibit a higher degree of interestingness and thus function especially well as severe tests of general mathematical conjectures.

7. Induction and Mathematical Practice

To summarize where we have got to thus far, the main thesis that I have been defending – in contrast to Baker (2007)'s inductive scepticism about mathematics – is that substantive enumerative inductive support can be accrued for a general mathematical conjecture from successfully testing the conjecture on the tiny numbers (i.e.

27 The catalog also includes many examples of properties that are not significant in our sense, either because they are not internally mathematical (e.g. being related to some physical world phenomenon) or because they are notation dependent (e.g. concerning some feature of the decimal digits of the number in question).

28 Why confine attention just to combinations of *two* properties in the diagnosis of interestingness? This is a fair question, and I do not mean to rule out analyses of interestingness that go beyond pairwise concatenations of significant properties. However, my worry is that if we allow interestingness to be determined by combinations of greater numbers of significant properties, it will become too easy to find very rare such combinations and as a result the bar for interestingness will become too low.

29 *Wikipedia* has entries for individual numbers and their notable properties, which is one useful source of information for non-tiny numbers.

30 The heptanacci sequence is formed analogously to the Fibonacci sequence, except starting with seven consecutive 1's and then forming the next entry in the sequence by summing the previous seven entries.

the natural numbers up to 1,000). In this section, I will argue that this modest pro-inductive stance fits better with mathematical practice than the sceptical view.

Some aspects of mathematical practice connected to enumerative induction are equivocal between the pro- and anti-inductive positions. I have in mind relatively well-known conjectures such as the Riemann Hypothesis, the Goldbach Conjecture, and $P \neq NP$, for which numerous instances have been surveyed, and about which mathematicians tend to have a very high confidence in their truth.[31] The complicating factor is that there are many considerations other than pure enumerative induction that feed into this confidence. Baker highlights one such consideration in the case of the Goldbach Conjecture, namely that the number of ways that an even number, n, can be expressed as the sum of two primes increases steadily with the size of n. In the current context, we are trying to adjudicate between the thesis that induction over the tiny numbers provides substantive inductive support and the contrary thesis that it provides no support. Given the prevalence of factors other than enumerative induction in examples such as the Goldbach Conjecture (GC), each of the above theses is seemingly consistent with mathematical practice.[32] According to the pro-inductivist, mathematicians' high level of confidence in the truth of GC is rooted in partial support from enumerative induction which is then boosted by other (also non-deductive) considerations. According to the inductive sceptic, mathematicians' high level of confidence comes entirely from these other considerations.

Are there other patterns of mathematical practice that give clearer guidance one way or the other? I think that there are, and I will present and discuss one important kind of case. A good illustration of what I have in mind is the Polya Conjecture (PC) example that I introduced right at the start of the paper. Recall that Polya initially verified the claim – about the total number of prime factors of a number being even versus being odd – for the first 1,500 natural numbers, before postulating that it might hold for all numbers. Polya (and, subsequently, other mathematicians) then set about trying to prove PC. A standard way to frame this kind of story is in terms of the distinction between "context of discovery" and "context of justification." The checking of the first 1,500 cases belongs to the discovery phase, while the search for a proof belongs to the justification phase. And this allows Polya's actions to be explained without attributing any justificatory force to enumerative induction over the 1,500 positive instances.

But is this explanation actually adequate? How rational is it to devote time and energy to trying to prove a conjecture if the enumerative induction carried out during the discovery phase provides no substantive evidentiary support for it? Imagine that a mathematician is faced with two *prima facie* equally plausible conjectures, C_1 and C_2. She has verified C_1 for the first 1,000 natural numbers, but she has only verified C2 for the first handful of numbers. If C_1 and C_2 are potentially on a par in all other respects (significance, relevance to her other work, etc.), surely it makes rational sense for the mathematician to devote her energy to trying to prove C_1. James Franklin makes essentially this point in the context of defending non-deductive methods in mathematics more generally:

> Anyone can generate conjectures, but which ones are worth investigating? . . . Which might be capable of proof by a method in the mathematician's repertoire? . . . [T]o say that some answers are better than others is to admit that some are . . . rationally better supported by the evidence.[33]

Interestingly, Polya himself has written about the role of induction in mathematics, and while he does not mention the example of his own conjecture, PC, he does defend the view that enumerative induction can provide substantive evidence for a general mathematical claim:

> [H]aving verified the theorem in several particular cases, we gathered strong inductive evidence for it. The inductive phase overcame our initial suspicion and gave us strong confidence in the theorem. Without such confidence we would have scarcely found the courage to undertake the proof which did not look at all like a routine job. "When you have satisfied yourself that the theorem is true, you start proving it" – the traditional mathematics professor is quite right.[34]

31 These three examples are mentioned by D'Alessandro in his (2021).
32 Baker (2007, pp.69-70).
33 Franklin (1987, p.1).
34 Polya (1954, pp.83-4). See also Zeng (2022) for an interesting discussion of the contrast between Polya's positive views about enumerative induction in mathematics and the anti-inductivist views of Lakatos.

In summary, a view that attributes positive (though not compelling) evidentiary value to enumerative induction over tiny numbers seems to fit much more naturally with the above aspect of mathematical practice than does a blanket inductive scepticism.

The philosophical debate here, however, is not just between inductivism and anti-inductivism. Part of the thesis that I am defending is that the claim of positive bias for tiny numbers is a key factor in what allows enumerative induction in mathematics to carry evidentiary weight. In other words, the pro-inductivist position that I have been arguing for above is to be distinguished from more mainstream versions of inductivism that are based on downplaying or dismissing size bias as a potential factor.

There is no standard terminology here, but *prima facie* there are six basic positions that one could adopt concerning size bias and enumerative induction in arithmetic. One could take size bias to weaken enumerative induction, to be neutral, or to strengthen enumerative induction. And one could combine this viewpoint with an overall stance that either sees enumerative induction in arithmetic to be rationally justified or does not. Using the terms "positive bias," "neutral," and "negative bias" for the first component, and "pro-inductivist" and "anti-inductivist" for the second component, we can generate labels for all six positions. I don't think that every position has actually been defended in the philosophical literature, but three of them certainly have.

Negative Bias Pro-Inductivism is the closest to the mainstream position. This is the view that size bias negatively affects the force of enumerative induction, but not to a degree that undermines induction carrying significant evidentiary weight. Among defenders of this position, I include Walsh (2014), Waxman (2017), D'Alessandro (2021). Negative Bias Anti-Inductivism is the view articulated and defended in Baker (2007), according to which the inevitability of size bias undermines the legitimacy of enumerative induction for arithmetic. And Positive Bias Pro-Inductivism is the position that I am defending in the current paper, according to which size bias can positively affect the force of enumerative induction such that samples including the tiny numbers can carry significant evidentiary weight.

Is there anything in mathematical practice that might help differentiate between the plausibility of the two alternative versions of Pro-Inductivism mentioned above? One very widespread feature of the way that mathematicians gather information about instances of a general conjecture is that they tend to try the smallest numbers first. In itself this has no particular implications for the issue of size bias, because there are obvious non-bias-related reasons for this pattern of practice. Typically, smaller numbers are easier to do calculations on, so it makes sense to try these numbers first. Also, for many conjectures, the calculation of the $n+1$ case requires that the n case be calculated first.[35]

Nonetheless, if Negative Bias Pro-Inductivism is true then any sample of smaller numbers is biased in a way that negatively impacts the evidentiary force of that sample, because in sampling from just one very small corner of the domain of natural numbers it is less representative than it could be. So if there were a relatively easy way to diversify the size range of the sample, we would expect to see mathematicians doing this. For example, it would be much better to randomly select one thousand numbers between 0 and 1,000,000 to sample, in support of a general conjecture, rather than simply sampling the first one thousand numbers (i.e. the tiny numbers). Yet we almost never see this happening, even when the relevant properties of the larger numbers are easy enough to calculate.

Another way of putting the above point is that mathematicians do not try to counteract the size bias in their sampling, even when doing so would be relatively straightforward. Indeed mathematicians actually seem to act in ways that *increase* the size-bias effect, by always starting at the very beginning of the number line, and by plodding through successive small numbers rather than randomly sampling larger numbers. As Baker (2007) argues, the fact that a sample has *some* size bias is not something that can ever be overcome: any finite sample will inevitably include only numbers that are "small" relative to the set of natural numbers as a whole. Nonetheless, there are *degrees* of size bias, and if the size bias in question were in fact a weakening factor for enumerative induction, we would expect straightforward measures that could be taken in order to overcome it to be manifested in actual mathematical practice. The fact that we rarely see this is further evidence that Positive Bias Pro-Inductivism is the more plausible of the two versions of pro-inductivist argument that we have canvassed.

[35] This is true of the Polya Conjecture, for example.

8. Objections

In this section I briefly survey, and respond to, several objections that might be raised against the position that I am arguing for. This position, which I have been calling Positive Bias Pro-Inductivism, maintains that substantive inductive support can be provided for a general arithmetical conjecture through sampling the tiny numbers (the natural numbers up to 1,000), and that this support stems crucially from the fact the tiny numbers are a biased sample.

Objection 1: 1,000 positive instances is not very many. How plausible is it that such a small number of instances yields substantive inductive support for a general mathematical conjecture?

It is important to re-emphasize that what I have been arguing for is that induction over the tiny numbers provides some substantive positive evidence for a general conjecture, and not that it ever (on its own) provides *compelling* evidence.[36] How high should the bar be set for evidence to count as "substantive"? I don't have a ready answer to this question, but a couple of approaches come to mind. One could define the evidence to be substantive if it leads mathematicians to the belief that the conjecture is more likely than not to be true. Or one could take a more operational approach, and deem evidence to be substantive if it makes it rational for mathematicians to pursue a proof of the given conjecture.

Objection 2: Cutting off the tiny numbers at 1,000 seems arbitrary. Why think that there is some significant difference between numbers slightly smaller than 1,000 and numbers slightly larger than 1,000?

No privileged status is being claimed for the cut-off of 1,000 between tiny numbers and non-tiny numbers. The thesis being defended is simply that enumerative induction over the tiny numbers is *sufficient* (other things being equal) to provide substantive evidentiary support. Obviously if this basis is sufficient then so is a larger basis that goes beyond 1,000 (for example, the basis of the first 1,500 numbers that Polya established in support of his Conjecture). It also may well be the case that for many conjectures a smaller initial basis is also sufficient. The cut-off point of 1,000 was chosen not because it has any specific, special features. The motivation was to pick a cut-off that is large enough so that the tiny numbers clearly provide a sufficient basis for substantive inductive support, and small enough that the bias claim (for tiny numbers being different from numbers at large) still has intuitive force.[37]

Objection 3: The 'positive bias' thesis is based largely on the claim that the range of significant properties is much greater among the tiny numbers than it is within comparably-sized stretches of non-tiny numbers. But how can we rule out there being lots of more complex significant properties that we have failed to recognize and which are well represented among the non-tiny numbers? In other words, perhaps the larger numbers just *look to us* to be relatively impoverished with respect to significant properties.

This is an important objection, since the basic possibility that it raises seems hard to rule out. However, even if we grant the assumption that we may be failing to recognize an array of significant mathematical properties that are too complex for our limited minds to grasp, I think that it is still possible to defend Positive Bias Pro-Inductivism. The reason is that the general conjectures that we – as human mathematicians – put forward will of practical necessity feature significant mathematical properties that are graspable by us. Mathematicians seek to prove theorems that are important, and the clearest way for a theorem to be important is through relating significant mathematical properties. A property that is too complex for us to grasp will be unlikely to feature in the theorems that we prove, and similarly for those properties that we can grasp but whose significance escapes us. Thus, for the purposes of defending the core thesis of Positive Bias Pro-Inductivism, we can simply

[36] Let alone that it provides *knowledge*, which we have already set aside as not being an issue that we are addressing here.

[37] As mentioned in Section 4, one problem with my (2007) delineation of the "minute numbers" is that many minute numbers are so large that it is difficult to discern systematic differences between minute numbers and non-minute numbers.

relativize to the significant mathematical properties that are graspable by us. The general conjectures that we formulate feature significant *and* graspable mathematical properties, and significant and graspable mathematical properties occur across a much wider range among the tiny numbers than among the non-tiny numbers. In other words, even if we do show a bias towards significant-to-us mathematical properties, this does not prevent induction involving these properties from supporting important-to-us general conjectures.

One might worry that this line of response to Objection 3 introduces a problematic circularity into the defense of enumerative induction. Are we not now making the relatively unsurprising claim that induction over the tiny numbers supports important-to-us conjectures, where "important-to-us" is defined as involving properties that are densest among the tiny numbers? I think that there are a couple of things that the defender of Positive Bias Pro-Inductivism can say here. Firstly, "important-to-us" is not being *defined* in reference to the tiny numbers. Rather an empirical claim is being made that the mathematical conjectures that we make tend to involve significant properties that are densest among the tiny numbers. Secondly, as was mentioned in Section 6, we are thinking of significance for mathematical properties as a matter of degree. While it seems plausible that there may be hitherto unrecognized significant properties that are well represented among the non-tiny numbers, it seems much less plausible that many (or any) of these will be as highly significant as properties such as being prime, being square, and being even, which feature in the conjectures that we have tested up to this point.

Objection 4: There is never any such thing as "pure" enumerative induction in mathematics: when induction is utilized, there are always other, conjecture-specific considerations that affect the degree of belief in the general conjecture. Hence there are no compelling grounds, based on mathematical practice, for attributing substantive evidentiary weight to enumerative induction.

Even if the factual claim is conceded, it is not clear why the ubiquitous presence of non-inductive considerations should undermine the Positive Bias Pro-Inductivism position. On a Kuhn-style theory choice model, enumerative induction is just one element among many that feeds into the overall non-deductive evidence for a given conjecture. But if enumerative induction never provides any significant positive weight, it remains puzzling why mathematicians devote time and energy to checking small instances of general conjectures. In other words, the 'no pure enumerative induction' claim does not weaken the argument from mathematical practice that was presented in Section 7 above.

Objection 5: The motivating example for the value of tiny numbers in lending inductive support for a general conjecture was Polya's Conjecture, and George Polya's survey of the first 1,500 numbers. But Polya's Conjecture turns out to be false! Doesn't this undermine the claim of inductive support in this case?

Before responding to the above objection, I will start with a brief overview of what transpired following Polya's original conjecture in 1919. In addition to the putative inductive evidence, there was also a heuristic argument in favor of Polya's Conjecture (PC). Recall that the claim is that, for every natural number, n, at least as many numbers less than n have an odd number of prime factors as have an even number of prime factors. The heuristic argument in favor of PC is as follows. All prime numbers have an odd number of prime factors, while composite numbers seem *prima facie* equally likely to have either an even number of prime factors or an odd number of prime factors. Hence we should expect numbers with an odd number of prime factors to predominate. On the other hand, there were also theoretical considerations that spoke against the truth of PC. It was discovered fairly early on that PC implies the Riemann Hypothesis. This fact was considered evidentially neutral,[38] however PC also implies linear dependence relations among the positive imaginary parts of the non-trivial zeroes of the Riemann zeta function. This latter claim is considered unlikely to be true, and this cast doubt in turn on the truth of PC.

[38] Since the Riemann Hypothesis is generally considered to be true.

This is how things stood until 1958, when Haselgrove not only proved that PC is false, but also that it fails for some $n < e^{832}$. Two years later the first specific counterexample to PC was found, by Lehman: PC fails for $n = 906,180,359$. In 1980, Tanaka proved that $906,150,257$ is the smallest number for which PC fails.

According to my main thesis, Polya was justified in looking for a proof of PC because he had surveyed the first 1,500 numbers, thus including all of the tiny numbers as part of his sample. In this case, as we have seen, PC was eventually proven to be false, but this in itself does nothing to undermine my claim. For it is perfectly possible to accrue substantive inductive support for a general conjecture that turns out to be false.

9. Conclusions

I concede that the main thesis I am defending may seem counterintuitive. It certainly sounds odd to cite the bias of our sampling methods as a crucial source of strength for enumerative induction in mathematics. However, in other respects the thesis is fairly conservative. Firstly, only a comparatively modest level of support is claimed for induction over the tiny numbers: they provide evidence that is significant rather than conclusive or compelling. Secondly, even the modest level of support holds only *ceteris paribus*. When might other things *not* be equal? Here are a couple of ways:

(i) Very few testable instances of the given general conjecture occur among the tiny numbers.

e.g. Consider perfect numbers, and the conjecture that all perfect numbers are even. There are only three instances of perfect numbers among the tiny numbers (6, 28, and 496). So testing this conjecture just on the tiny numbers does not accrue substantive inductive support.

(ii) There is some clear size-relative component to the given general conjecture that makes it more likely for larger numbers to provide counterexamples.

e.g. Consider the claim that no number is expressible as the sum of two cubes in three different ways. Plausibly, the larger the number the more different ways it is likely to be possible to express it (and the more cubic numbers there are that are smaller than it). And, indeed, the first example of such a number is 87,539,319, which is well beyond the range of the tiny numbers.[39]

In addition to cashing out the *ceteris paribus* condition, another way that the main thesis can be sharpened is by getting clearer on what it is saying about necessary and sufficient conditions for enumerative induction in the mathematical context to be effective. Thus far, all that I have claimed is that an inductive basis that includes all of the tiny numbers is *sufficient* (absent any conjecture-specific contra-indications) to provide substantive inductive support for a general conjecture. Is the inclusion of all of the tiny numbers also *necessary* for this support to accrue?

In considering this question, it may be helpful to distinguish between two different features of the set of tiny numbers. Firstly, the set includes an *initial segment* of the natural numbers. Secondly, the set includes *one thousand* different potential instances. Once we separate out these two features, we can consider alternative evidential bases that lack one but not the other. Compare, for example, the following two samples:

(A) A sample consisting of the natural numbers from 0 to 10.

(B) A sample consisting of the natural numbers from 11 to 1,000.

How much evidence for a generic general conjecture is provided by each of these samples? Intuitively, neither the A-sample nor the B-sample provides sufficient evidence for reasonable confidence in the truth of a general conjecture about the natural numbers. For the A-sample, there are simply too few positive instances to make a reasonable inductive inference. All sorts of trivially false conjectures happen to hold for the first handful of

[39] In a way, this case illustrates the converse of what I argued in (2007) for the Goldbach Conjecture case. For GC, according to my 2007 argument, there is zero support from 'pure' enumerative induction, but there are other considerations that strongly suggest that counterexamples are more likely to occur among smaller cases rather than larger cases.

numbers.[40] What about the B-sample? This sample has almost as many instances as the entirety of the set of tiny numbers, so it does fine with respect to the second feature mentioned above. However, the B-sample excludes the first eleven numbers, and so it does not comprise an initial segment of the natural numbers. This is a problem because of one of the key aspects of the tiny numbers, mentioned back in Section 6, namely the high frequency of *boundary cases*.[41] Excising the first eleven numbers from the tiny numbers yields a sample that has drastically fewer boundary cases, because the first handful of numbers include boundary cases for a multiplicity of significant properties.[42]

What does consideration of the A-sample and the B-sample tell us about necessary conditions for substantive inductive support? There are unlikely to be sharp boundaries around what counts as a necessary, minimal evidential basis, but what we can say is that any such basis needs to include a reasonable sequence of cases from the very beginning of the natural number line (say at least the first ten numbers) and a reasonable proportion of the tiny numbers (say at least a few hundred numbers).[43] At the end of the day, I think that establishing the sufficiency claim is more important, dialectically, than honing the necessity claim. The key question concerning enumerative induction in mathematics is whether it ever provides substantive evidentiary support. The thesis that the tiny numbers are sufficient to provide such support answers this question in the affirmative.

Acknowledgements

Versions of this paper were presented at Princeton University in November 2021, at the Eastern Division Meeting of the APA in Baltimore in January 22, and at the conference on Mathematics and Analogical Reasoning in Munich in June 2022. I am grateful to audiences at these three events for helpful questions and feedback, and also to Alexander Paseau and to William D'Alessandro for comments on earlier drafts of this paper.

References

Baker, A. (2007). Is There a Problem of Induction for Mathematics?. In M. Leng, A. Paseau, & M. Potter (Eds.), *Mathematical Knowledge* (pp. 59-72). Oxford: Oxford University Press.
Baker, A. (2008). Experimental Mathematics. *Erkenntnis*, 68, 331-344.
Caldwell, C. (2021). Table of Known Maximal Gaps between Primes. *Prime Pages*. OL at https://primes.utm.edu/notes/GapsTable.html
Corfield, D. (2004). Mathematical Kinds, or Being Kind to Mathematics. *Philosophica*, 74(2), 37-62.
D'Alessandro, W. (2021). Mathematical Philosophy. *The Reasoner*, 15(4), 32-33.
Franklin, J. (1987). Non-Deductive Logic in Mathematics. *British Journal for the Philosophy of Science*, 38, 1-18.
Frege, G. (1884). *Die Grundlagen der Arithmetik*. Breslau: W. Koeber. Translated by J. L. Austin as *The Foundations of Arithmetic*, Oxford: Blackwell, second revised edition 1953.
Hajek, A. (2014). Philosophical Heuristics and Philosophical Creativity. In E. Paul & S. Kaufman (Eds.), *The Philosophy of Creativity: New Essays* (pp. 288-318). Oxford: Oxford University Press.
Hardy, G. (1940). *A Mathematician's Apology*. Cambridge: Cambridge University Press.
Lange, M. (2015). Explanation, Existence, and Natural Properties in Mathematics – A Case Study: Desargues' Theorem.
Paseau, A. (2021). Arithmetic, Enumerative Induction, and Size Bias. *Synthese*, 199, 9161-9184.
Polya, G. (1954). *Mathematics and Plausible Reasoning: Volume 1: Induction and Analogy in Mathematics*. Princeton, NJ: Princeton University Press.
Walsh, S. (2014). Empiricism, Probability, and Knowledge of Arithmetic. *Journal of Applied Logic*, 12, 319-348.
Waxman, D. (2017). Deflation, Arithmetic, and the Argument from Conservativeness. *Mind*, 126, 429-464.
Weisgerber, S. (2024). Value Judgments in Mathematics: G. H. Hardy and the (Non-) seriousness of Mathematical Theorem. *Global Philosophy*, 34.
Wells, D. (1986). *The Penguin Dictionary of Curious and Interesting Numbers*. Penguin Books: UK.
Zeng, W. (2022). Lakatos and Quasi-Empiricism. *Kriterion*, 36, 227-246.

[40] To give just one example: "All odd numbers are either prime or square" holds for all numbers in the A-sample.
[41] Also, mathematicians seem to be aware (at least implicitly) of the importance of including an initial segment as part of any enumerative inductive basis, since the initial segment is almost always surveyed when considering any general conjecture.
[42] The numbers from 0 through 10 include the (lower) boundary cases for properties such as being odd, being even, being prime, being square, being perfect, and being a Fibonacci number.
[43] Note that the necessary conditions that I am suggesting here are conditions for providing *substantive* positive evidence. A sample that does not include the first few numbers, or does not include a reasonable proportion of the tiny numbers may still provide *some* evidence in favor of a given conjecture.