Mathematical instrumentalism, Gödel's theorem, and inductive evidence

A.C. Paseau

1. Introduction

Mathematical instrumentalism construes some parts of mathematics, typically the abstract ones, as an instrument for establishing statements in other parts of mathematics, typically the elementary ones. The best-known modern instrumentalism is Hilbert's. Hilbert insisted that a proof of the consistency of instrumental mathematics or, as he called it, 'ideal' mathematics, be given within elementary or 'real' mathematics. However, Gödel's second incompleteness theorem (G2) apparently shows that real mathematics cannot even prove its own consistency, let alone that of the stronger ideal mathematics. G2 is therefore generally thought to defeat Hilbert's instrumentalism. But does G2 defeat mathematical instrumentalism more generally, or merely Hilbert's particular version? Against Hilbert, I shall argue that the requirement of proof is not at the heart of mathematical instrumentalism. Mathematical instrumentalism consequently survives G2.

2. Mathematical instrumentalism

The mathematical instrumentalist divides mathematics into two parts, real and ideal. (We stick to the vivid labels 'real' and 'ideal' without commitment to the particulars of Hilbert's instrumentalism. Other forms of instrumentalism, not discussed here, might distinguish more than two parts of mathematics.) According to the instrumentalist, real mathematics is in some manner superior to ideal mathematics. Because of the different metaphysics of the two areas, or because of our differential epistemic access to them, or both, or for some other reason, the role of the ideal part of mathematics as the instrumentalist sees it is to deliver statements in the real portion. For a given instrumentalism, we may label the precise form this constraint must take—the type of real statements the ideal part must deliver if it is to discharge its instrumental role successfully—the instrumentalist constraint. And we may call the type of argument required to show that the instrumentalist constraint is satisfied the instrumentalist argument.

Hilbert's instrumentalist ideas in ''On the Infinite'' (1925) and ''The Foundations of Mathematics'' (1927) illustrate this general characterisation.[1] In these papers Hilbert characterised the real/ideal division[2] in three ways, intended to be extensionally equivalent: (i) the part of mathematics that is intuitable versus the unintuitable part; (ii) the part of mathematics requiring only a grasp of the finite versus the part requiring a grasp of the infinite; (iii) elementary statements, truth-functional compounds and bounded quantifications thereof as well as finitary generalisations, versus the rest of mathematics.[3] It is now generally held that for Hilbert real mathematics is Primitive Recursive Arithmetic (PRA),[4] and the rest is ideal mathematics. More generally (for any instrumentalism whatsoever, not just Hilbert's), let '$\vdash_{I+R}$' denote 'provable in ideal and real

---

[1] We set aside Hilbert's other views on the philosophy of mathematics.

[2] The distinction was intended to mirror the observation/theory distinction in natural science (1927, p. 475).

[3] Finitary generalisations are distinguished from elementary real statements by containing free variables, e.g. '$x + y = y + x$' or '$2 \times y = y \times 2$'.

[4] Tait (1981) is the *locus classicus* justification for identifying finitism with PRA.

mathematics' (i.e. mathematically provable), '$\vdash_R$' denote 'provable in real mathematics only', and '$\bot$' denote a real contradiction (e.g. '$0 = 1 \wedge \neg\, 0 = 1$'). Here are some possible instrumentalist constraints:[5]

Consistency: $\nvdash_{I+R} \bot$

Reliability: For all real p, if $\vdash_{I+R} p$ then p is true

Conservativeness: For all real p, if $\vdash_{I+R} p$ then $\vdash_R p$

Avoidance: For all real p, if $\vdash_{I+R} p$ then $\nvdash_R \neg p$

Reliability is sometimes called 'Soundness' and may be thought of as a constraint of semantic conservativeness, as opposed to *Conservativeness*, which is a constraint of deductive conservativeness. Assuming that real proof (the relation denoted by '$\vdash_R$') is sound with respect to real truth, *Conservativeness* entails *Reliability* and its special case *Consistency* (a special case because '$\bot$' is a false real statement). If real proof is sound and complete with respect to real truth, *Reliability*, *Conservativeness* and *Avoidance* are equivalent, as is *Consistency* if the underlying logic is also classical (as '$0 = 1 \wedge \neg\, 0 = 1$'is classically derivable from any contradiction).

Hilbert himself insisted on a proof of *Consistency* and of *Conservativeness*.[6] The conjunction of the two (or simply the stronger of the two, *Conservativeness*, on the assumption that real proof cannot establish contradictions) therefore constituted his instrumentalist constraint. The instrumentalist argument he sought was a real proof—a proof within real mathematics—that this constraint holds.[7] Hilbert insisted on this because he thought real mathematics grounds the epistemology of mathematics. Yet G2 implies that consistent standard deductive systems cannot prove their own consistency. According to the G2-based argument, for any system R with a standard proof predicate:

$$\nvdash_R (\nvdash_R \bot)$$

[5] For Hilbert these statements should be construed as finitary generalisations. The list is not intended to be exhaustive.

[6] ''For there is a condition, a single but absolutely necessary one, to which the use of the method of ideal elements is subject, and that is the proof of consistency; for, extension by the addition of ideals is legitimate only if no contradiction is thereby brought about in the old, narrower domain, that is, if the relations that result for the old objects whenever the ideal objects are eliminated are valid in the old domain.'' (1925, p. 383; see also 1927, p. 471). In this passage Hilbert seems to be conflating Consistency with Reliability. But on the assumption of real-completeness and classical logic Consistency and Conservativeness are equivalent; and Hilbert had a proof-based conception of real truth, which collapses Conservativeness into Reliability. A version of the consistency problem was of course the second of Hilbert's twenty-three problems from the 1900 address: ''But above all I wish to designate the following as the most important among the numerous questions which can be asked with regard to the axioms: To prove that they are not contradictory, that is, that a finite number of logical steps based upon them can never lead to contradictory results.'' (1900, p. 9).

[7] Claims about proofs can be made in real mathematics using coding.

Since $I+R$ extends R, it follows that $\nvdash_R (\nvdash_{I+R} \perp)$. Hence neither *Consistency* nor *Conservativeness* can be proved in R. This seems to damage Hilbert's instrumentalism beyond repair. In our more general terms, G2 apparently implies that there is no prospect of a Hilbertian instrumentalist argument for the satisfaction of the Hilbertian instrumentalist constraint. One could object that real mathematics may not be formalisable within a single system R. But that does not help since the consistency of $I+R$ isn't even provable in $I+R$ and $I+R$ outstrips (the union of) any open-ended series of formalisations of real mathematics. Hence the orthodox conclusion: G2 sinks Hilbert's instrumentalism.[8]

A few philosophers, notably Michael Detlefsen, dissent from orthodoxy and maintain that G2 does not show that the consistency of the ideal part of mathematics cannot be proved by its real part. Unlike Detlefsen, we shall not dispute the claim that G2 shows that ideal mathematics cannot be proved a good instrument using real mathematics.[9] Instead, our defence is that the instrumentalist argument need not take the form of a proof. The best-known modern instrumentalist may have insisted on proof; but Hilbert's insistence notwithstanding, proof is not a necessary ingredient of a successful mathematical instrumentalism.

3. The inductive defence

Since our main interest is in defending the idea that the instrumentalist argument need not take the form of deductive proof, it does not matter terribly which instrumentalist constraint

---

[8] With this in mind Bill Tait writes: ''The attempt to found mathematics on finitism fails definitively, as is well known.'' (1981, p. 21); Giaquinto comments: ''The argument stemming from Gödel's Second Underivability Theorem is convincing [given acceptable assumptions] . . . So I conclude that there is no serious possibility of achieving the goal of Hilbert's Programme for arithmetic.'' (2002, p. 196); Kreisel: ''[in arithmetic and metamathematics] the broad claim behind the program is refuted, and the program and its relevance are dubious'' (1976, p. 93); and Smoryński, ''the passage of time has erased the optimism of the turn of the century and has supplied a mathematical refutation of both Hilbert's philosophy and faith'' (1979, p. 117). Gödel himself in his original paper (1931, p. 195) was cautious about G2's implications for Hilbert's philosophy of mathematics.

[9] Detlefsen (1986)'s negative argument has several prongs, an important one being that the derivability conditions have not been shown to be acceptable constraints on any notion of provability. That the conditions might not be acceptable does not alter the fact that no proof of *Reliability* has been given, so that an explanation of why no proof is needed is still welcome. One of Detlefsen's more positive considerations in his (1986) is to appeal to so-called Rosser systems, a certain kind of 'consistency-minded' system. A Rosser system S^C (the 'C' is for 'consistency-minded') is defined in terms of a base system S (e.g. Peano Arithmetic, Zermelo-Fraenkel-Choice set theory). Given an enumeration of S's theorems, we define theoremhood in S^C by: 'θ is a theorem of S^C iff it is the nth *S*-theorem and ¬θ is not the mth *S* -theorem for any $m < n$'. Such systems are consistent only in the narrow sense that no theorem is the negation of any other. What remains to be shown is that they are consistent in the broader and more relevant sense that their theorems do not jointly lead to contradiction. And although the narrow consistency of Rosser-systems is easy to prove real-mathematically, there is no reason to think *Reliability* is (cp. Detlefsen, 1986, n. 48, pp. 140–141). In the end, the principal conclusion of Detlefsen (1986) is the avowedly negative one, that the G2-based argument does not show that *Reliability* is not real-provable, rather than the positive one that there is reason to think that *Reliability* is real-provable

we pick. We adopt *Reliability* as the main constraint for the sake of concreteness and for two other reasons. First, in the absence of reasons for *Reliability*, we lack reason for thinking that ideal mathematics cannot lead us astray by implying real falsehoods. In contrast, absent a proof-based conception of real mathematical truth,[10] the requirement that every ideally-provable real truth is real-provable is motivated only by the thought that each use of the instrument is in principle eliminable. This seems an optional addition to the instrumentalist tenet that the role of the instrument is the derivation of truths in the real kernel. So much the better, the instrumentalist might think, if the instrument proves more truths in the real kernel than are non-instrumentally provable. Instrumentalism in the philosophy of science opts for the analogue of *Reliability*, and there is no reason to see a disanalogy with the philosophy of mathematics on this point. Second, an instrumentalism that insists on *Conservativeness* already falls prey to Gödel's first incompleteness theorem (G1). On almost any construal of the real/ideal division, ideal mathematics proves some real statements not provable by real methods.[11] Since our aim is to examine the soundness of the G2 argument against instrumentalism, we may assume that instrumentalism has survived the G1 argument in order for the G2 argument even to arise. We will therefore assume that Conservativeness is not the required instrumentalist constraint. Having said that, our discussion will sometimes focus on the weaker constraint of Consistency when the implications for *Reliability* depend on where the real/ideal boundary lies.[12,13]

The inductive defence, then, claims that though it may sink Hilbert's instrumentalism, G2 does not sink mathematical instrumentalism more generally. The reason is that instrumentalists can settle for evidence for *Reliability* that does not take the form of proof. We need a word for this kind of evidence, so let us settle on 'inductive'. In a nutshell, this paper's point is that inductive evidence can be sufficient to warrant rational high credence in the appropriate instrumentalist constraint, thereby meeting one of instrumentalism's key aims.

By inductive evidence for p we mean any kind of evidence for *p* other than a proof of *p*.[14] The most familiar type of inductive evidence is enumerative. For example, $F(a_1)$, $F(a_2)$, …, and $F(a_n)$ are inductive evidence for the claim that all *a*s are *F*.[15] Inductive reasoning

[10] Which was in fact present in Hilbert's thinking.
[11] As Hilbert could not have known in the mid-1920s.
[12] Detlefsen (1990), which defends Hilbert's instrumentalism from G1 in more detail, may be interpreted as taking (what I call) Reliability to be the more important instrumentalist constraint. Giaquinto (2002, pp. 148–149, 162–163, 179–180) emphatically takes this line. Several commentators have argued that Hilbert should, *qua* instrumentalist, have plumped for Reliability (e.g. Kreisel 1976; Prawitz 1981, p. 253). Note that *Reliability* is not a Hilbertian constraint, because what Hilbert put stock in was finitary evidence, not the semantical property of truth. Such a notion of truth was not acceptable for Hilbert because propositions concerning it are not refutable by finitary evidence.
[13] What if the instrument is known to lead us astray in certain areas? Instead of considering this a failure of Reliability, the instrument could be restricted to the areas in which it is held to be reliable.
[14] Anyone who construes arguments ordinarily held to be inductive as deductive should understand inductive evidence as evidence that is ordinarily taken to be inductive.
[15] In mathematical circles this kind of reasoning is often casually called 'empirical'. This usage is unfortunate, as the claims '$F(a_1)$, $F(a_2)$, …, $F(a_n)$' may be a priori.

also includes other forms of non-deductive reasoning such as inference to the best explanation.[16] It also encompasses any evidence for *p* that is deductive in nature but is not a deduction of *p*. For example, a proof of *q* where *q* is a weakening of *p* may in this sense count as inductive evidence for *p* (examples to follow). Our construal of 'inductive evidence for *p*' only excludes evidence that takes the form of a deduction of *p* itself. It does not preclude the use of deduction for convincing oneself of the truth of *p*, so long as this use does not take the form of a proof of *p*. Our notion of inductive evidence is broad and subsumes many varieties of evidence. It may or may not be of epistemological interest in other contexts. But that doesn't matter for present purposes since our aim is only to defeat the G2-based argument against mathematical instrumentalism. This argument insists that the evidence for *p* must come in the form of a deductively valid proof whose conclusion is *p*. If there is sufficient evidence of a different form for *p*, of whatever kind, its existence is enough to block an argument that sees the absence of a proof of *p* as fatal for instrumentalism. Our broad construal of 'inductive evidence for *p*' is precisely what is required for this dialectical purpose.

Mathematical instrumentalism's success condition is that the upper tier of mathematics—the instrument—be deemed successful from the perspective of the lower tier. Our defence of mathematical instrumentalism is that, whatever the upper tier's success comes to (consistency, reliability, etc.), establishing it need not consist in producing a lower-tier proof to that effect. Strong inductive reasons can achieve the same end. Inductive evidence in this broad sense may rationally justify a high level of confidence in ideal mathematics' reliability (and in particular in its consistency).

So far as I am aware this argument has not been properly articulated and defended before, though no doubt it has crossed the minds of many who have thought about mathematical instrumentalism and to some extent underpins parts of the mathematical work in the tradition of Hilbert's Programme.[17] The best way to develop our defence following this bald statement is to parry various objections to it. This will take up the rest of the paper. Before that, I mention two caveats about the discussion's relation to Hilbert's philosophy of mathematics.

First, we will not be concerned with post-Gödelian work in the tradition of the Hilbert Programme (or with other aspects of Hilbert's philosophy of mathematics). In particular, we won't be concerned with results concerning the proof-theoretic reduction of one mathematical theory T_1 to another T_2 (e.g. where T_1 is some subsystem of second-order arithmetic and T_2 is PRA). Some of these results are pregnant with philosophical significance, illuminating as they do the epistemology of mathematics by revealing what can be known on the basis of what. But the partial revival of Hilbert's philosophy of mathematics that reconstrues it as a series of relativised reductionisms is no part of our project.[18] We are concerned only to defend the thesis that ideal mathematics can be justified wholesale from within a single real-mathematical perspective.

[16] Which perhaps subsumes enumeratively inductive reasoning—we won't take a stand on this.
[17] Kreisel obliquely hints at something along the inductive response's lines; see e.g. his (1968, pp. 358–362) or his (1976). But as usual with Kreisel's philosophical writings, it is hard to tell what exactly he had in mind. Detlefsen (1986), a book-length defence of Hilbert's philosophy against the G2 objection, does not consider the inductive response.
[18] For discussion, see Sieg (1988), Simpson (1988) and Feferman (1988).

Second, the inductive defence would be anathema to Hilbert. As the latter wrote, ‘‘there is a condition, a single but absolutely necessary one, to which the use of the method of ideal elements is subject, and that is the *proof of consistency*’’ (1925, p. 383, cp. 1927, p. 471). We have accordingly distinguished Hilbert’s instrumentalism from the broader category it belongs to, namely mathematical instrumentalism. As explained, two ideas make up the latter’s ideological core. First, that mathematics has two philosophically quite different portions: real and ideal mathematics. (To repeat, we are using Hilbert’s terminology—itself drawn from mathematical analysis of course—without commitment to his particular form of instrumentalism.) Second, that one can give an argument from outside ideal mathematics—the ‘instrumentalist argument’—that ideal mathematics discharges its instrumentalist function satisfactorily—that it satisfies the ‘instrumentalist constraint’. I have identified this constraint as *Reliability*, though most of the discussion applies to other constraints the instrumentalist might choose. And I have identified the instrumentalist argument as seeking to establish rational high credence in the satisfaction of this instrumentalist constraint. The inductive defence respects these two ideas, seeking to give an instrumentalist argument that Hilbert would not have countenanced, by providing an inductive rather than a deductive argument. The conclusion is that though Hilbert’s instrumentalism may be undermined by G2, not all mathematical instrumentalisms are.[19]

4. Objections to the inductive defence

Objection 1. A proof would offer certainty in *Reliability*, though no inductive argument could. Mathematical instrumentalism’s success condition is to establish certainty in *Reliability*. Hence the inductive defence fails.

Reply. Both premisses are false. A proof would not offer rational certainty in *Reliability* nor *Consistency* nor any other constraint, and nor does instrumentalism require it. First of all, that we are in possession of a proof of *p* does not imply that we should be certain of *p*. We may not be certain of the proof’s premisses; and it would be irrational to be more certain of an argument’s conclusion than its premisses if that is our sole reason for believing it. Equally, the proof may be long and hard to follow, so that any flesh-and-blood mathematician should assign non-zero probability to its being invalid. The longer and more complex the proof, the less secure its conclusion, as Hilbert’s great collaborator Paul Bernays regretfully noted.[20]

---

[19] Why did Hilbert not appreciate this? Well, Hilbert was a mathematician par excellence and his professional instincts compelled him always to look for deductive proof of mathematical claims. Besides, at the time he was propounding his philosophy, in the 1920s, flush with the success to which he himself had contributed of finding axiomatic bases for any mathematical theorem one cared to mention, it was reasonable for him to believe that mathematics is complete, that is, that every mathematical claim is provable from some fixed axiom system. Hence he naturally assumed that *Consistency* was provable. Having quite justifiably zeroed in on provability, he was subsequently blind to the available retrenchment following Gödel’s theorems. Anyhow, why Hilbert did not see it that way is a matter for intellectual history, and our interest here is in the normative point that an instrumentalist can legitimately see it otherwise.

[20] ‘‘Gradually the expectations [of finding a proof of *Consistency*] were disappointed, and in the process it became palpable that the danger of an oversight is particularly great in the area of metamathematical reasoning.’’ (in Hilbert (1935/1970, p. 210), translation by Stephan Leuenberger). Kreisel (1976, p. 116) makes a similar point.

Arguably, no one should rationally believe any statement with certainty, or perhaps no statements other than Cartesian propositions such as ‘I exist now’ or ‘I am thinking’. Less contentiously, it is rationally permissible to give degree of belief less than 1 to virtually all claims. Since certainty entails degree of belief 1,[21] and given the normative (some would say constitutive) link between degrees of belief and betting behaviour, certainty in a statement implies that you should would be willing to stake your life, risk eternal damnation in hell, etc., for no return on its truth. Yet it is rationally permissible not to go to the stake over the soundness of most if not all arguments; it is not irrational of fallible inquirers, reliably prone to error, to maintain a degree of doubt in virtually any claim. Thus even if, *per impossibile*, a real proof of *I*’s reliability existed, we—fallible inquirers, reliably prone to error—would almost certainly be rationally entitled to give its conclusion less than full credence.

Second, mathematical instrumentalism should not require that *Reliability* be established with certainty. As explained, two theses make up mathematical instrumentalism, the second of which is that the instrumental part of mathematics performs its instrumental function satisfactorily. Is a philosophy of mathematics unacceptable if one of its components is not believed with certainty? No. It is acceptable if there are strong reasons for believing it, reasons that warrant a high degree of belief. Compare scientific instrumentalism: it would be unreasonable to ask for a *proof* of the empirical adequacy of the instrumental part of science over the observational part. Observe also that the first tenet of instrumentalism, that there is an epistemologically significant distinction between real and ideal mathematics, can hardly be established with certainty. Even if we could establish the second tenet with certainty, then, our overall credence in instrumentalism should still be lower than 1. There is no reason to insist that the second component of a philosophical position be established with certainty when the first cannot be.

Objection 2. A proof of *Reliability* would rationally entitle the instrumentalist to a higher credence in this claim than any inductive argument could deliver.

Reply. Perhaps. I am not denying that a proof of *Reliability* is desirable as far as rational credence-raising is concerned. Depending on its length and complexity, it might offer us a more secure warrant for belief in *Reliability*. It might rationally raise our credence in *Reliability* more than any piece of inductive reasoning could. To pick some numbers fairly randomly, it might raise our credence in *Reliability* to 0.99, whereas inductive reasons might never take our credence beyond 0.97. The credence-raising power of proof is not in question, nor the fact that, typically—but only typically—a proof of *p* in relatively elementary mathematics can raise our credence in *p* more than the available inductive evidence can. However, this does not affect the moral that an inductive argument can offer good reasons for belief in *Reliability*. If inductive grounds can rationally raise our credence in *Reliability* to a high degree, instrumentalism’s second tenet is vindicated.[22] Instrumentalism may not be in as

[21] No commitment to the converse is made, that degree of belief 1 entails certainty.
[22] I’m not denying that Hilbert thought that certainty was attainable in these matters, nor am I claiming that for him highly credible grounds for belief in *Reliability* were good enough. Hilbert was explicit about the need for ‘definitive’ reliability: ‘‘That, then, is the purpose of my theory. Its aim is to endow mathematical method with the definitive reliability that the critical era of the infinitesimal calculus did not achieve’’ (1925, p. 370). But as before, Hilbert has overegged the pudding. Instrumentalism can take as its

good a place as it would have been had a proof of *Reliability* been available; but it is in a good place nonetheless. For details of why on one natural way of drawing the real/ideal division the inductive evidence for *Reliability* is strong, see the reply to objection 8.

We may put the point more generally. A key end of mathematical instrumentalism is to establish a rationally high degree of belief that the instrumentalist constraint is satisfied. The means by which Hilbert sought to achieve that end was a proof that the constraint obtains. Yet the unavailability of this particular means does not show that the end itself is unattainable. So long as there are other ways of rationally achieving a high degree of belief in the instrumentalist constraint's satisfaction, instrumentalism is not defeated by G2. To take the unprovability of consistency as a knock-down objection to instrumentalism is to mistake a means of the project for the end itself.

Objection 3. Recalling Hilbert's desire for a permanent solution to the consistency problem, a deductively valid argument is monotonic in the sense that new premisses will not convert it into an invalid one, in contrast to inductively valid arguments. The latter are systematically subject to one type of threat of non-permanence that deductive arguments are not.

Reply. It is characteristic of non-deductively valid arguments that addition of premisses can invalidate them. This, however, is a *logical* sense of permanence. The question remains what the *epistemic* significance of this logical permanence (monotonicity) might be. Given that deductive argumentation almost never (if at all) brings with it certainty, proof almost never (if at all) achieves epistemological permanence in the sense of rational unrevisability of its conclusion. As mentioned, we may come to reject some of the premisses or inferential steps of any given proof that we currently accept, or we may for some other reason come to suspect its validity (e.g. because of its length). In response, one might say that deductive arguments in mathematics are typically more secure than inductive ones. However, this would be to concede that the difference is of degree rather than of kind. It takes us back to the second objection, answered in broad terms above and in more detail in reply to objection 8. To make the objection stick, then, one must find an important epistemological sense of permanence that inductive arguments lack, or possess to a considerably lesser extent than deductive arguments. We have already ruled out certainty, unrevisability, credence-raising and rational credence-raising. We shall rule out some other candidates in what follows.

Objection 4. The instrumentalist argument (whatever form it may take, inductive or deductive) should aim to do more than establish rational high credence in the instrumentalist constraint (whichever it may be, e.g. *Reliability* or *Consistency*).

Reply. Achieving rational high credence in the satisfaction of the instrumentalist constraint should at least be a key aim of instrumentalism. This aim is at the core of the various versions of instrumentalism in the philosophy of science, the paradigm of philosophical instrumentalism. Suppose you have empiricist grounds for thinking that the unobservable claims made by the theory of relativity should not be believed. You can nevertheless trust its observable predictions (made on its own or in conjunction with some other theories) if you are able to show, by recourse to the history of the theory's applications or otherwise, that it is rational to place a high degree of belief in the soundness of these (joint or individual) predictions. In passing, we note that the inclusion of the word 'rational' respects the idea that

---

purpose to endow mathematical method with a high degree of reliability. There is no call for 'definitive' reliability, whatever that might be.

the instrumentalist argument will have succeeded if it is rational, on the basis of the evidence, to have a high degree of belief that the instrumentalist constraint obtains. Rational credence is constitutively tied to what the evidence supports; it is not purely subjective.

What epistemic stance should an instrumentalist aim for other than rational high credence? We have already encountered some unattainable epistemic goods: certainty and permanence understood as unrevisability. Another aim might be justification. However, as will become clear, the same arguments go through for justification as for rational high credence. Given reasonable construals of these notions, we are highly justified in thinking that ideal mathematics is consistent and indeed reliable with respect to real mathematics. (On several ways of understanding the notion of rational high credence, it covers justification anyway.) The same goes for warrant and other evidential notions in the area.

Objection 5. Since mathematics insists on deductive proof, the grounds on which we judge the acceptability of mathematics should also be deductive. Otherwise instrumental mathematics and the rules by which we judge whether it is an acceptable instrument are dissonant.

Reply. The reply is twofold. First, as we shall explain in response to the next objection, inductive reasons can be rationally suasive within mathematics. Second, there is nothing wrong with rejecting a certain kind of evidence within real mathematics while accepting the very same kind of evidence for judging mathematics' acceptability. The following are distinct: (i) reasons within mathematics; and (ii) the instrumentalist's reasons for accepting mathematics.

To illustrate this distinction, take the formalist construal of ideal mathematics as a game (arguably espoused by Hilbert). Consider the game of whist. The aim of whist is to take as many tricks as possible. Contrast the reasons for playing whist which are, I suppose, to disport oneself in a structured, intellectually stimulating and sociable way. It would be absurd—a kind of category mistake—to suggest that the way to evaluate whist's success in achieving sociability is to count tricks. Likewise, the rules and methods of the mathematical game are one thing, the rules and methods for determining whether the mathematical game should be played another. There is no dissonance between taking the former to be strictly deductive and the latter to be inductive (or a mixture of inductive and deductive). As is readily seen, this argument does not depend on taking a formalist view of ideal mathematics and can be generalised to any mathematical instrumentalism.

Objection 6. The only evidence for mathematical statements is deductive. Given that *Reliability* (or *Consistency* etc.) is a mathematical statement, inductive evidence for it is no evidence at all.

Reply. A defence of the role of what I called inductive reasons in mathematics has been offered by many writers, for example famously in Pólya (1959) and in Lakatos (1967). What follows must perforce fall short of a full defence, and I can do no more than summarise some of the main points. To illustrate the strength of inductive reasons in mathematics, consider Goldbach's Conjecture (GC). GC states that every even number greater than 2 is the sum of two primes. Number theorists have a very high degree of belief in its truth on inductive grounds. First, GC has been checked for every even number up to about 3×10^{17}—as of 2005, so that number is no doubt greater today. Second, various slightly weaker claims, for example that every sufficiently large odd number (greater than approximately $4 \times 10^{43,000}$) is the sum of three primes, or that every sufficiently large number (in another sense) is the sum of a prime and a product of at most two primes, have been proved. (This illustrates our broad use of 'inductive evidence for *p*' to include deductive evidence for statements other than *p*.) Third, let $G(n)$, the Goldbach number of *n*, be the number of different ways in which *n* can be written as the sum of two primes. GC can then be expressed as the claim for all even *n* greater than 2, $G(n) \geq 1$. Computer evidence shows that the function $G(n)$ broadly increases for even *n* as *n* increases (with oscillation, but with a steady increasing trend), so that for instance for $n \approx 10^5$, $G(n) \geq 500$. In light of this evidence, the thought that $G(n)$ will suddenly drop to 0 is deeply unlikely. This and much other evidence accounts for mathematicians' virtual certainty in GC's truth.

Probabilistic tests are another example of compelling inductive reasons in mathematics. Perhaps the best known is Rabin's (1980) primality test, which performs a simple test on a randomly chosen number *k* smaller than *n*. If the verdict is that *n* is composite, then *n* must be composite. If the verdict is that *n* is prime, then the chances of *n* being prime are over 3/4. The tests are independent for different *k*, so if we run the test for 1000 numbers smaller than *n* and each time the verdict is that *n* is prime then the likelihood that *n* is prime is more than $1 - 4^{-1000}$. Mathematicians take at face value the probabilities the test delivers and, allowing for updating, adopt them as credences.

These facts are compatible with the truism that mathematicians prefer deductive proof and actively look for it even in the presence of overwhelming inductive evidence. The reason for this is that they are mathematicians and as such value deduction. As Frege put it, ''It is in the nature of mathematics always to prefer proof, where proof is possible, to any confirmation by induction'' (1884, §2; notice the caveat). Given that the ends of mathematical activity are not solely to raise rational degree of belief in mathematical truths, it is accordingly not a cogent objection to our argument that mathematicians go on seeking proofs of Goldbach's Conjecture or the primality of a given number in the presence of overwhelming inductive evidence. Compare marathon-running, whose ends are to get to a destination 24 miles away and to run there. If our sole end is to reach the final destination then any means will do; in particular, we may drive the 24 miles by car. It is not a cogent objection to driving that if one had marathon-runners' ends one would shun the car in favour of running. Likewise with proof: the instrumentalist's ends are not the mathematician's. Mathematical practice enshrines proof as an end in itself, but there is no reason for an instrumentalist to adopt this as an end as well, given that her project is philosophical. This is not the place to broach the question of why mathematics elevates proof to an end, though we will touch on this question below. Fallis (1997, 2002) interestingly defends the radical claim that proof serves *no* epistemic objectives that inductive evidence cannot also serve.

It should be stressed that as a matter of fact most mathematical knowledge is based on inductive evidence (and is a posteriori). For one thing, most mathematical beliefs are acquired by testimony: picked up from books, journals, conversations with other mathematicians, etc. And testimonial beliefs are in our broad sense inductive (and a posteriori). For example, I believe Whitney's Embedding theorem (about the embedding of manifolds) because the Springer book says so and past evidence suggests that Springer textbooks are highly reliable. That is an a posteriori, non-deductive reason for believing the theorem.[23] Or I believe Fermat's Last Theorem on the good authority of Andrew Wiles and

[23] Informants' prima facie trustworthiness may or may not be an a posteriori assumption, but either way one's evidence for a belief p acquired by testimony always includes the empirical

the handful of mathematicians who have the ability to check it, as well as on the reliability of the media and individuals via which the result has been relayed to me. Even Andrew Wiles' belief in Fermat's Last Theorem has an a posteriori, testimonial component. Wiles, I take it, did not personally check all the premisses on which his 1995 paper rests: he too relied on others' reliability for various claims, theorems, lemmas, techniques, etc., to which he helped himself in that sprawling paper. The same moral applies to theorems that rely on computer proof, such as the Four Colour Theorem, which rely on inductive evidence that the computer is functioning properly; to so-called folk theorems that are taken for granted and airily quoted by the community but for which no one seems to have a ready proof and whose credibility thus rests on the community's dependability; and to theorems whose various parts have been verified by groups of researchers but which no single person is able to take in, now or ever, for example the classification theorem for finite simple groups.

The research frontier serves up two further examples. A good research mathematician typically tries to prove statements she has strong grounds for believing are both true and susceptible to her proof techniques. Having reliable credences about the truth of unproved statements is thus an essential part of mathematical research as we know it. Second, reasons for adopting new (independent) axioms are not deductive. By definition, reasons for adopting new axioms, for example adding the Axiom of Inaccessibles or the Axiom of Projective Determinacy to ZFC or the Axiom of Choice to ZF, are inductive in our broad sense.

Though the research frontier illustrates the point excellently, the illustrations need not be recherché. The majority of lay mathematicians who believe Pythagoras' Theorem, I suspect, have never personally worked their way through a proof of the theorem. There are also examples aplenty of simple mathematical statements whose non-deductive justification mathematicians generally trust more than their deductive justification. Take the claim that the universal function for general recursive functions is general recursive. Given Church's Thesis, the claim has a one-line proof. It can also be proved in tedious detail without resort to the thesis; a typical textbook presentation (e.g. Cutland, 1980, pp. 95–99) takes up five pages. Most people believe Church's Thesis on inductive grounds (involving a large measure of testimonial evidence); indeed it is generally thought that the thesis cannot be deductively proved. Which argument for the universal function statement do mathematicians who have seen both have more confidence in? Invariably, it is the one based on Church's Thesis rather than the longer deductive proof.[24] Another example is the statement that $2^{2^{24}}+1$ is composite, whose deductive proofs at the time of writing are, on account of their complexity, less trusted than the computation to the same effect.

Arguably, a similar point to the one about testimony can be made about mathematicians' reliance on memory. As, it is controversial, consider instead mathematical beliefs that apparently rest on purely deductive grounds. Imagine that I have carried out a proof of Fermat's Last Theorem without assistance from memory or from others. In this apparently paradigm case of solitary deductive reasoning is the reasoning really exclusively deductive? Is it a priori? Well no, typically it isn't either. Unless my proof has been carried out with extraordinary rigour it will contain several large gaps. My deductive argument almost certainly contains shortcuts, steps skipped, abbreviated sub-arguments. So what entitles a

---

fact that so-and-so told you that *p*. Tyler Burge (1993) has recently questioned this received wisdom about testimony; see Malmgren (2006) for a reply.

[24] This is not to deny that they have a high degree of confidence in the latter.

mathematician to think that a sketchy argument, missing significant argumentative steps, can be expanded into a proper proof? The answer: inductive evidence. I know that in the past when I have judged a proof sketch with this-sized gaps in it to be valid I or others have managed to fill in the gaps to anyone's satisfaction (when challenged, for fun, etc.). Like any half-decent mathematician, I appreciate when an argument amounts to a proof, not through insight into its fully rigorous archetype, but in virtue of my past experience that relatively ungappy arguments of this kind can be filled in successfully if necessary. This filling need not result in a fully rigorous proof; it may simply make the proof less gappy, satisfying the higher standards of a stricter context. The key point is not that rigour must always take the form of fully articulated deduction, but that some arguments are sketchy even with respect to a quotidian conception of rigour. As all this shows, much actual mathematical knowledge turns out upon close inspection to be inductive.[25]

Objection 7. Agreed, inductive evidence for mathematical statements is often compelling. But enumerative induction cannot support a generalisation over infinitely many cases, for two reasons. First, if we know of only finitely many instances, our evidence is infinitesimally small compared to the generalisation's infinite scope. Second, our sample is biased because it consists only of the small cases. This scuppers any hope for supporting *Reliability* (or *Consistency*) inductively.

Reply. We take the two versions of the objection in turn. The inductive evidence for *Reliability* (or *Consistency* etc.) is not limited to verification of finitely many instances of each, as we shall see below. But the objection is misconceived anyway. Mathematics is rife with enumeratively inductive reasoning over infinitely many cases, as we saw with GC. On what authority does the objector accuse mathematicians of irrationality? To anyone uncomfortable with attributing mass irrationality to mathematicians, it is more reasonable to treat the objection as a *reductio* of modelling confidence levels naively as the ratio of the number of verified cases to the number of total cases. A Bayesian treatment helps explain how one may have a high degree of confidence in an infinite generalisation on the basis of finitely many instances. Let '$\forall I$' be an infinite generalisation and 'I_1', 'I_2', .., 'I_n' its known instances (i.e. statements asserting that a given instance conforms to the generalisation). From Bayes' Theorem, it follows that

$$Pr(\forall I / I_1 \wedge I_2 \wedge .. \wedge I_n) = Pr(I_1 \wedge I_2 \wedge .. \wedge I_n / \forall I) \times Pr(\forall I) \div [Pr(I_1 \wedge I_2 \wedge .. \wedge I_n / \forall I) \times Pr(\forall I) + Pr(I_1 \wedge I_2 \wedge .. \wedge I_n / \neg\forall I) \times Pr(\neg\forall I)]$$

(These probabilities are credences.) Since '$\forall I$' entails '$I_1 \wedge I_2 \wedge .. \wedge I_n$' (as we are perfectly aware), and '$Pr(I_1 \wedge I_2 \wedge .. \wedge I_n / \neg\forall I) \times Pr(\neg\forall I)$' equals '$Pr(I_1 \wedge I_2 \wedge .. \wedge I_n \wedge \neg\forall I)$', this is equivalent to

$$Pr(\forall I / I_1 \wedge I_2 \wedge .. \wedge I_n) = Pr(\forall I) \div [Pr(\forall I) + Pr(I_1 \wedge I_2 \wedge .. \wedge I_n \wedge \neg\forall I)]$$

Now the smaller $Pr(I_1 \wedge I_2 \wedge .. \wedge I_n \wedge \neg\forall I)$ becomes relative to $Pr(\forall I)$ for increasing n, the closer $Pr(\forall I / I_1 \wedge I_2 \wedge .. \wedge I_n)$ approaches 1. It is an empirical fact that sometimes the

[25] Fallis (2003) discusses this mathematically familiar point, whose philosophical consequences are not sufficiently appreciated.

mathematical community believes that $Pr(I_1 \wedge I_2 \wedge .. \wedge I_n \wedge \neg\forall I)$ is very small compared to $Pr(\forall I)$.

Of course, as is well known, the orthodox Bayesian framework cannot be applied wholesale to mathematics (e.g. it forces logical truths to have probability 1).[26] But the above application is not affected by that problem. Moreover, even if the Bayesian analysis is thought problematic, the basic point that the modelling is intended to capture stands: mathematicians give some claims, for example the Goldbach Conjecture or the Riemann Hypothesis, extremely high credence even in the absence of proof.

A certain sort of philosopher might complain that precisely because we don't have a good comprehensive mathematical model for these credences, they should not be trusted. But this would be an extreme 'theory first' perspective. On this view, however systematic, reliable, and informally understood certain phenomena might be, in the absence of an undergirding theory they are not to be trusted. I submit that this is to put the cart before the horse. Compare the situation before the emergence of probability theory: would it have been unreasonable to think that some claims are more credible than others, that some evidence can confirm a hypothesis, etc., in the absence of a satisfactory mathematical model? Is it unreasonable now for anyone who has not taken a course in probability or formal epistemology?

One could insist that although mathematicians do in the relevant cases take $Pr(I_1 \wedge I_2 \wedge .. \wedge I_n \wedge \neg\forall I)$ to be low, they are wrong to do so. This criticism could be intended as either internal or external to mathematics. If the former, it might succeed in some cases. No doubt there are some generalisations of which mathematicians are more (or less) confident than their evidence warrants judged by the mathematical standards of the day. Past examples *perhaps* include Fermat's confidence that $2^{2^n} + 1$ is prime for all n, or until 1914 when Littlewood proved otherwise, that the logarithmic integral $Li(x)$ (defined as $\int_2^x dt/ ln(t)$) is larger than $\pi(x)$, the number of primes up to x (it turns out that the first crossover between the two occurs at about 10^{316}). Mathematicians have in the past got such estimates wrong, and their confidence levels have at times failed to match the available evidence. However, once we move into the computer age, misplaced confidence of this kind becomes less frequent, and expert mathematicians have become more cautious about extrapolating small-sample trends to larger samples. Experts' credences are seasoned by awareness of misplaced past confidence. When such an extrapolation is upheld by the experts in the relevant field, it is a considered judgment, typically underpinned by other inductive evidence, as in the case of GC. Almost invariably, it conforms to the highest mathematical standards of the day. To challenge mathematicians' confidence in the truth of GC from an internal mathematical perspective therefore requires a mathematical argument, in the form of new evidence or a reason that the available evidence has been improperly weighed. We shall consider the available evidence for the case of interest to us in reply to objection 8.

Alternatively, if the objection is intended as external to mathematical practice, it loses its force for all but the most anti-naturalist philosopher. If we assume that scientific practice in the broad sense is by and large reasonable, grounds for overturning expert opinion in a mathematical field concerning a mathematical question must be grounds recognisable by the

[26] For a recent discussion of some of the issues, see Part II of Corfield (2003), which also contains several more examples of inductive reasons in mathematics.

experts in that field. This is not to say that mathematicians' views are sacrosanct. The point, rather, is that arguments that mathematicians are collectively unreasonable on this scale lead to a much more general scepticism.

The second version of the objection is that initial segments of the natural numbers are a biased sample and therefore that our enumerative inductive evidence for infinite numerical quantifications is by its nature misleading. This objection is distinct from the one just considered because it is based on the alleged bias of the sample (a number's place in the number line is known to affect its properties) rather than the sample's infinitesimal size compared to the whole set. It is, however, just as implausible as the first because it sets its face against ingrained mathematical practice, as we saw in the case of GC. The responses just given to the infinitesimal sample objection apply to the biased sample objection as well. But the latter objection is also weak in two further respects. First, we may have reason to believe that the sample is biased in a favourable way. As we saw in the case of GC, the partition function evidence suggests that if a counterexample to GC exists, it is likely to occur early on in the number sequence. Naturally, the partition function evidence is yet more inductive evidence, so we haven't produced a non-circular argument in favour of inductive reasoning. But since our aim is not to defeat the Humean sceptic, we may use our available evidence to demonstrate that initial samples may be more reliable indicators of the generalisation's truth than later samples.

Second, scepticism about enumerative inductive evidence for a numerical generalisation does not easily apply to our defence of instrumentalism, because proofs of real statements do not come in a natural numerical order. On some numerical codings, proofs with more steps are assigned a higher number than shorter proofs, while on others it is the number of logical connectives that matters most. Codings can start arbitrarily at 1 or 10 or 10^{10} or $10^{10^{10}}$, and so on. What unifies the class of proofs available to us is therefore not that they correspond to some initial subset of the number line comprising only small numbers. The class's unity lies rather in their accessibility to us. Scepticism about finite inductive evidence regarding the infinite totality of proofs therefore cannot be based on the thought that initial segments of the natural numbers are biased samples.

This is not to deny that there is a fairly natural complexity ordering on proofs and that the accessible proofs are the simpler ones. But there is much less of a tendency, within mathematics, to think of simpler proofs as a biased sample of the totality of proofs than to think of initial segments of the natural numbers as biased. (A tendency which, we have argued, should be resisted given sufficient evidence.) A pronounced scepticism here can only be based on the thought that humanly accessible proofs are a biased sample of proofs. However, this scepticism is of the general, Humean kind, since all evidence we may acquire, mathematical, scientific or other, is humanly accessible. In the scientific case, too, we happily extrapolate from accessible instances to generalisations covering all instances. There does not appear to be a special problem of induction for *Reliability*.

Objection 8. Granted, inductive evidence can rationally increase our confidence in a claim. But there is little inductive evidence for *Reliability* or *Consistency*.

Reply. The response must be on a case by case basis for different instrumentalisms. To illustrate the type of defence that the instrumentalist will wish to offer, we take R to be PRA (the received view of what Hilbert took real mathematics to be) and $I + R$ to be ZFC set theory.[27] An analogous case could be made for similar instrumentalisms. We mostly focus on *Consistency* in what follows and return to *Reliability* towards the end.

Since it would not have been unreasonable to doubt the emergent ZFC's consistency in the mid-1920s, to accuse Hilbert of not appreciating the inductive response to the G2 argument is, among other things, anachronistic. But our present vantage point is not Hilbert's. Almost a century later, things look different: it is no longer reasonable to have serious doubts about the consistency of ZFC.[28]

One strong form of current evidence for the consistency of ZFC (or its reliability over PRA) consists in the verification of instances. All existing ideal proofs of real statements have issued in apparently true statements. Even when the real conclusions of ideal proofs are not real-provable (e.g. Goodstein's theorem or the Gödel sentence for PRA), many of their instances have been verified. Note that the search for a counterexample to *Consistency* (or *Reliability*) may proceed backwards from a contradiction or other falsehood as well as forwards. Yet no route to a falsehood from the axioms of ideal mathematics (say ZFC) or from a falsehood back to the axioms has ever been found.

As the locus of many of the original paradoxes, the consistency of ZFC has been consistently probed in the past hundred years. Indeed, Zermelo's avowed aim in his (1908) axiomatisation was precisely to encompass as much set-theoretic reasoning as possible within a consistent system that blocks the paradoxes. Given its history, people have actively tried to look for contradictions in ZFC, starting with the handful of paradoxes available at the start of the twentieth century: Cantor's, Russell's, Burali-Forti's, etc. A passive, half-hearted search that does not turn up any evidence of the quarry is little evidence of its absence; but an active search that comes back empty-handed a century later is considerable evidence that there is no quarry to be found. Thus Bourbaki, writing several decades ago:

> Since the various mathematical theories are now logically attached to the Theory of Sets, it follows that any contradiction encountered in one or another of these theories must give rise to a contradiction in the Theory of Sets itself. Of course, this is not an argument from which we can infer the consistency of the Theory of Sets. Nevertheless, during the half-century since the axioms of this theory were first precisely formulated, these axioms may have been applied to draw conclusions in the most diverse branches of mathematics without leading to a contradiction, so that we have grounds for hope that no contradiction will ever arise. (1968, p. 13).

The case of ZFC's consistency is in fact in some ways similar to that of GC. As a result of ongoing work in the system as a well as the large searches carried out by automated proof systems, we have lots of enumerative inductive evidence on behalf of its consistency. Perhaps

---

[27] More precisely, *R* and *I* consist of the mathematics formalisable in these systems.

[28] Hilbert himself never doubted the consistency of analysis, though he took it to be a part of ideal mathematics. (For instance in his Paris address he presupposes its consistency: ''Indeed, when the proof for the compatibility of the axioms shall be fully accomplished, the doubts which have been expressed occasionally as to the existence of the complete system of real numbers will become totally groundless.'' (1900, p. 10).) Where did this confidence stem from? Whatever its source, it was not proof. In practice, then, even the arch-deductivist's confidence in consistency rested on inductive grounds.

not as many instances as we have in the case of GC, but plenty nevertheless. Another similarity with the case for GC is the existence of partial results, most of them familiar.[29] ZFC without the Axiom of Infinity is equi-interpretable with Peano Arithmetic (over PRA as a base system), about whose consistency there is no genuine doubt. ZFC without the Axiom of Replacement has $\langle V_{\omega+\omega}, \in\rangle$ as a model, a model so low down in the set-theoretic hierarchy and graspable with such minimal infinitary resources that there is no real doubt about its consistency either. (Whether this kind of evidence is available to the instrumentalist will be discussed below.) ZFC is equiconsistent with ZF, which tells us that any inconsistency in the former would not be due to the Axiom of Choice. Since Choice is the only ZFC-axiom that has been seriously questioned in the system's hundred-year history, that is good grounds for confidence in ZFC's consistency. One could of course maintain that mathematicians' confidence levels are much higher than they ought to be: not just slightly optimistic (which would be compatible with our instrumentalist defence) but wildly optimistic. But as with the analogous scepticism regarding the inductive grounds for GC, this extreme antinaturalism has no cogency.

Whether 'favourable bias' is also a point of similarity is less clear. Set theories known to be inconsistent, such as naïve set theory and variations thereof, can typically be shown inconsistent in a small number of steps. Indeed, Cantor's, Russell's and Burali-Forti's Paradox are all one- or two-liners. This suggests, as in the case of GC, that any bias in our evidence is in fact favourable, i.e. that if ZFC were inconsistent, its inconsistency could be established in a small number of lines. However, there is a prima facie credible extension of set theory, ZFC plus the claim that there is a non-trivial elementary embedding of the set-theoretical universe into itself, that is inconsistent (as shown by Kunen) but not obviously inconsistent. And if New Foundations should turn out to be inconsistent after decades of investigation it too would be an example of a set theory whose inconsistency is hard to spot. Whether there really is a 'favourable bias' case for ZFC's consistency is therefore hard to say.

However that may be, further evidence for ZFC's consistency is provided by the course taken by set theory in the past half-century. I don't have the space or competence to detail the pertinent results comprehensively, so let me summarise two key points. Confidence in ZFC's consistency is greatly increased by results in inner model theory. In particular, the universe of constructible sets $\boldsymbol{L}$ is known to settle pretty much any set-theoretic question of interest. Such is the constructible universe's resolution power and transparency that set theorists are convinced that if there were a contradiction in ZFC it would make itself manifest in $\boldsymbol{L}$. Yet no contradiction has made itself manifest in L. Second, the network of results concerning ZFC's proposed extensions strengthens the case for consistency. For example, the potential axiom $\mathrm{AD}^{\boldsymbol{L}(\mathbb{R})}$ seems to be entailed by plausible axiom candidates for extending ZFC (some of which are inconsistent with one another).[30] Given this and $\mathrm{AD}^{\boldsymbol{L}(\mathbb{R})}$'s fruitful and plausible consequences—in particular, it has great simplifying power in analysis without contradicting the Axiom of Choice, in contrast to unrelativised AD—set theorists tend to believe that any correct extension of ZFC must be an extension of ZFC + $\mathrm{AD}^{\boldsymbol{L}(\mathbb{R})}$. This commonality between proposed extensions is taken to be strong evidence for the truth of ZFC + $\mathrm{AD}^{\boldsymbol{L}(\mathbb{R})}$, never mind the mere consistency of ZFC.[31]

---

[29] Proofs of these facts may be found in many set theory textbooks.
[30] $\mathrm{AD}^{\boldsymbol{L}(\mathbb{R})}$ is the Axiom of Determinacy relativised to $\boldsymbol{L}(\mathbb{R})$, the class of constructible sets with $\mathbb{R}$ as given.
[31] Thanks to Peter Koellner for discussion of these points.

Contemporary set theorists' confidence in ZFC is perhaps best illustrated anecdotally. At the Gödel centenary conference in 2006, Hugh Woodin promised to resign his chair and demand that it be filled by the first person to find an inconsistency in ''ZFC + There are ω-many Woodin cardinals'', a theory much, much stronger than ZFC. Today's set theorists would in short consider you a crank or a humorist if you were to question the consistency of a set theory as relatively down to earth as ZFC. They are not necessarily right. But their expert view should weigh heavily with us regarding a theory whose ins and outs they know better than anyone, at least regarding the mathematical question of its consistency as opposed to, say, the philosophical question of its truth.

Perhaps the strongest form of inductive evidence for ZFC's consistency is that there is an informal model for it, the iterative conception of set (developed after Hilbert propounded his philosophy of mathematics).[32] In a nutshell: start with the empty set at the first stage, then take power sets of already-formed sets through all the available stages, including the first infinite one and beyond. This picture of the universe of sets now guides research and teaching in the subject. Realists think the picture corresponds to an independent reality, anti-realists demur. Whatever one's metaphysics, the point is that whenever we have as clear a mental picture as this, its associated theory is very likely consistent. We know of this link not by some deductive insight but by empirical induction over past cases.[33]

You might object that this evidence is not available to the instrumentalist because according to her our intuitive picture of the set universe is illusory. The instrumentalist does take the iterative conception to be illusory; she thinks it does not correspond to reality. But it is consistent with the conception's failure to correspond to mathematical reality that it is perfectly coherent, just as a fictional world can be coherent even if it does not exist. Perhaps the instrumentalist goes further and thinks that our intuitive picture of sets is illusory in a stronger sense. This sense cannot be incoherence of course, otherwise the associated theory would be inconsistent, contrary to what the instrumentalist in fact believes. The stronger sense in which the intuitive picture is illusory might be that it is meaningless (this would be the formalist construal of instrumental mathematics, arguably espoused by Hilbert). In that case, her evidence cannot include the fact that we intuitively grasp what a transfinite model of set theory would look like if it existed—a grasp we might have even such models do not exist (the weaker sense in which the intuitive picture is illusory). Her evidence is exhausted by the sociological facts that set theorists possess some sort of mental picture which they all seem to agree on and to develop in the same way, which leads them to see some axioms as natural, others as less natural, etc. These abilities have developed over several decades, by thousands of researchers and many more students. This kind of evidence does not require infinitary capacities to appreciate: it is straightforward sociological evidence about what

---

[32] Though it is due in part to Mirimanoff, the iterative conception first properly emerges in Zermelo (1930). The conception was still relatively unknown at the time of Hilbert's death in 1943. Boolos (1971) offers a clear exposition. The conception's existence sets ZFC apart from Quine's NF.

[33] This too is enumeratively inductive evidence. But the inference based on it is of the form 'theory$_1$ with characteristic ϕ is/is most likely consistent, . . . , theory$_n$ with characteristic ϕ is/is most likely consistent, therefore ZFC since it has characteristic ϕ is/is most likely consistent', rather than the already mentioned inductive inference of the form 'proof Π_1 does not lead to $\bot$, . . . , proof Π_n does not to lead to $\bot$, therefore no proof leads to $\bot$'.

many people have said and done. It can be appreciated even by someone who thinks our intuitive picture of sets is meaningless. It is obviously available within the real perspective.

The weight of numbers and history is relevant in assessing the sociological evidence's strength. It will not do to object to our argument by citing some localised instances of apparently coherent conceptions which ultimately issued in an inconsistent theory. The case of ZFC's consistency is very different from the case that springs to every philosopher's mind of Frege's theory of extensions undone by Russell's Paradox. No sooner had a few logicians laid eyes on Frege's theory than it was shown inconsistent. Not so with ZFC, which has been widely used for much longer by a great many mathematicians. It is also important that the historical induction be a positive one, from widespread mathematical visualisability to consistency, rather than the negative one that infers inconsistency from widespread mathematical unvisualisability (e.g. the pre-nineteenth century unvisualisability of the negation of Euclid's fifth postulate). Positive inductions have a stronger track record than negative ones. And perhaps this evidence only kicks in conjunction with other evidence of the kind available in the case of the consistency of ZFC, for example that no derivations within the theory suggested by the intuitive picture have resulted in contradictions or other falsehoods.

The evidence we have given for ZFC's consistency is inductive in our broad sense, which is the sense relevant to rebutting the G2-based argument. Further analysis of this inductive evidence would of course be of great interest, but is beyond the scope of this paper. Our aim is not to analyse and systematically classify the inductive evidence behind Consistency but to present a chunk of it. Appreciating the cumulative force of the inductive evidence is one thing, classifying and analysing it another.

Finally, the case for *Reliability* is just as strong as the case for *Consistency* when R = PRA and $I + R$ = ZFC. For suppose that $\vdash_{I+R} p$, where p is a real falsehood. If p is not a generalisation, then $\vdash_{I+R} \neg p$, since $\neg p$ is true. (Any real p that contains no variables is formalisable as a Δ_0-statement and ZFC is Δ_0-complete.)[34] Thus $\vdash_{I+R} p \wedge \neg p$. If p is a false real generalisation, say $\phi(x_1, \ldots, x_n)$, then some instance of it, say $\phi(a_1, \ldots, a_n)$, is false. Thus $\vdash_{I+R} \neg\phi(a_1, \ldots, a_n)$ since $\neg\phi(a_1, \ldots, a_n)$ is a true finitary non-generalisation, but since $\vdash_{I+R} \phi(x_1, \ldots, x_n)$, it follows that $\vdash_{I+R} \phi(a_1, \ldots, a_n) \wedge \neg\phi(a_1, \ldots, a_n)$. Either way, *Consistency* is violated. Thus *Reliability* inherits all of *Consistency*'s inductive support.

Objection 9. Granted, it is rational to give *Reliability* or *Consistency* high credence, at least for natural choices of I and R. But in the absence of a real-mathematical proof the instrumentalist has no explanation for why the instrumentalist constraint obtains.

Reply: To begin with, a deductive proof might not offer an explanation either. Many deductive proofs aren't explanatory. For example, proofs of generalisations that proceed disjunctively, with different instances having entirely different sub-proofs, are not explanatory. Deductive reasons need not be more explanatory than inductive ones. Second, we have at least a sociological explanation of ideal mathematics' reliability: it has been developed precisely with the satisfaction of the reliability constraint in mind, as everyone—

[34] Even Robinson Arithmetic is Δ_0-complete.

instrumentalist and non-instrumentalist alike—can appreciate. If, say, techniques from analytic number theory had led to number-theoretic falsehoods, one of the two would have been revised. Third, some inductive evidence offers an intrinsic explanation of ZFC's consistency, independently of sociology. In particular, that ZFC has an intended model explains its consistency. ZFC is consistent because it expresses facts about this model, whose coherence we grasp. Compare the parallel explanation in the case of arithmetic: PA is consistent because it expresses facts about an intended model, whose coherence we grasp. This explanation of PA's consistency is stronger than Gentzen's proof and similar deductive proofs of PA's consistency.[35]

A final observation. You might think that mathematical instrumentalism is explanatorily disadvantaged vis-à-vis mathematical realism if the former's explanation of the instrument's reliability is worsted by the latter's explanation, namely that the instrument is reliable because it is true. But don't forget that the instrumentalist claims to have independent arguments that undercut the realist explanation in the first place. Hilbert for instance thought he had arguments to show we shouldn't believe ideal mathematics. Compare scientific instrumentalism: as long as the instrumentalist has a story about why the successful instrument (say relativity theory or whatever) literally construed does not deserve our credence, that story can be deployed to show that the best explanation of the instrument's success cannot be that it is true. If this first core component of instrumentalism is well-motivated—something we have not touched on in this paper—it follows that the instrumentalist's claims about why *Consistency* and *Reliability* hold are not in explanatory competition with the realist's account.

5. Conclusion

We have illustrated the moral that the instrumentalist may have compelling inductive reasons for thinking that ideal mathematics does not issue in real-mathematical falsehoods with a case study inspired by (but distinct from) Hilbert's instrumentalism. According to the inductive defence, any epistemic advantage deductive arguments might be said to have over inductive arguments is in fact illusory. It is either unattainable (e.g. certainty); or it is of relatively minor importance; or it is achieved by inductive grounds (e.g. explanatoriness, rational credence-raising), albeit perhaps to a diminished extent. Although we have not examined all alleged epistemic advantages, we have discussed most of the main ones and have seen that they all slot into this trichotomy.[36] The apparent moral of the discussion is that mathematical instrumentalism is not sunk by G2. Instrumentalism's success turns instead on whether the division of mathematics into the real and ideal portions is justified.[37]

---

[35] The strongest deductive proofs of PA's consistency seem to be those that presuppose this explanation at some level.

[36] Knowledge in fact also fits this pattern, since for natural choices of I the instrumentalist can have knowledge that I is consistent or reliable. That non-deductive evidence in mathematics can amount to knowledge is argued in Paseau (2010).

[37] Detlefsen (2005, pp. 306–310) calls this the Division Problem. Kitcher (1976) is a critical discussion of Hilbert's epistemology of mathematics with which I am largely in agreement.

Acknowledgments
Thanks to my Oxford tutorial students on whom these ideas were first tried out; to Elliott Sober, Greg Fried, Tim Williamson, two SHPS referees and especially Mic Detlefsen for comments; and to audience members at Sydney University and its Centre for the Foundations of Science, the Australian National University, Manchester University, Aberdeen University, Bristol University, the University of California at Irvine, and the École Normale Supérieure (Paris), where I presented versions of this paper, in particular Alan Hájek, Alan Weir, Carrie Jenkins, David Braddon-Mitchell, David Chalmers, Erich Reck, Fiona MacPherson, Jeremy Avigad, Jeremy Heis, Jonathan Schaffer, Mark Colyvan, Mary Leng, Matthew Foreman, Øystein Linnebo, Patrick Greenough, Paul McCallion, Penelope Maddy and Peter Koellner.

References

Boolos, G. (1971). The iterative conception of set. *Journal of Philosophy*, 68, 215–231.
Bourbaki, N. (1968). Elements of mathematics: Theory of sets. Reading, MA: Addison Wesley.
Browder, F. (Ed.). (1976). *Proceedings of symposia in pure mathematics: Mathematical developments arising from Hilbert problems* (Vol. 28, Part 1). Providence, RI: AMS.
Burge, T. (1993). Content preservation. *Philosophical Review*, 102, 457–488.
Corfield, D. (2003). *Towards a philosophy of real mathematics.* Cambridge: Cambridge University Press.
Cutland, N. (1980). *Computability*. Cambridge: Cambridge University Press.
Detlefsen, M. (1986). *Hilbert's Program*. Dordrecht: D. Reidel.
Detlefsen, M. (1990). On an alleged refutation of Hilbert's Program using Gödel's First Incompleteness Theorem. *Journal of Philosophical Logic*, 19, 343–377.
Detlefsen, M. (2005). Formalism. In S. Shapiro (Ed.), *The Oxford handbook of philosophy of mathematics and logic* (pp. 236–317). Oxford: Oxford University Press.
Fallis, D. (1997). The epistemic status of probabilistic proof. *Journal of Philosophy*, 94, 165–186.
Fallis, D. (2002). What do mathematicians want? Probabilistic proofs and the epistemic goals of mathematicians. *Logique et Analyse*, 45, 373–388.
Fallis, D. (2003). Intentional gaps in mathematical proofs. *Synthese*, 134, 45–69.
Feferman, S. (1988). Hilbert's Program relativized: Proof-theoretical and foundational reductions. *Journal of Symbolic Logic*, 53, 364–384.
Frege, G. (1884). *Die Grundlagen der Arithmetik* (Transl. by J.L. Austin). *Foundations of arithmetic*. Oxford: Blackwell.
Giaquinto, M. (2002). *The search for certainty.* Oxford: Oxford University Press.
Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia mathematica und verwandter Systeme I (Transl. by J. van Heijenoort as ''On formally undecidable propositions of *Principia mathematica* and related systems I'' and repr. in his S. Feferman et al. (Eds.). *Collected Works* (Vol. 1, pp. 145–195). Oxford: Oxford University Press).
Hilbert, D. (1900). Mathematische Probleme. In *Göttinger Nachrichten* (pp. 253–297) (Transl. by M. Winston Newson and repr. in Browder (1976), pp. 1–34).
Hilbert, D. (1925). Über das Unendliche. *Mathematische Annalen* 95 (Trans. by S. Bauer-Mengelberg as ''On the infinite'' and repr. in van Heijenoort (1967), pp. 367–392).
Hilbert, D. (1927). Die Grundlagen der Mathematik. *Abhandlugen aus dem mathematischen Seminar der Hamburgishcen Universität* 6 (Trans. by S. Bauer-

Mengelberg and D. Føllesdal as ‘‘The foundations of mathematics’’, and repr. in van Heijenoort (1967), pp. 464–479).
Hilbert, D. (1935/1970). *Gesammelte Abhandlungen* (Vol. 3). Berlin: Springer-Verlag.
Kitcher, P. (1976). Hilbert’s epistemology. *Philosophy of Science*, 43, 99–115.
Koellner, P. (2006). On the question of absolute undecidability. *Philosophia Mathematica*, 14, 153–188.
Kreisel, G. (1976). What have we learnt from Hilbert’s second problem? (in Browder (1976), pp. 93–130).
Lakatos, I. (1967). A renaissance of empiricism in the recent philosophy of mathematics? *Mathematics, science and epistemology: Philosophical papers* (Vol. 2, pp. 24–42). Cambridge University Press.
Malmgren, A.-S. (2006). Is there a priori knowledge by testimony? *Philosophical Review*, 115, 199–241.
Paseau, A. (2010). Non-deductive knowledge of consistency. MS.
Pólya, G. (1959). Heuristic reasoning in the theory of numbers. *American Mathematical Monthly*, 66, 375–384.
Prawitz, D. (1981). Philosophical aspects of proof theory. In G. Fløistad (Ed.). *Contemporary philosophy. Philosophy of language and philosophy of logic* (Vol. 1, pp. 235–277).
Rabin, M. (1980). Probabilistic algorithm for testing primality. *Journal of Number Theory*, 22, 128–138.
Sieg, W. (1988). Hilbert’s Program sixty years later. *Journal of Symbolic Logic*, 53, 338–348.
Simpson, S. (1988). Partial realizations of Hilbert’s Program. *Journal of Symbolic Logic*, 53, 349–363.
Smoryński, C. (1979). Review of Browder (1976). *Journal of Symbolic Logic*, 44, 116–119.
Tait, W. (1981). Finitism. *Journal of Philosophy*, 78 (repr. in his *The Provenance of pure reason* (pp. 21–42). New York: Oxford University Press).
van Heijenoort, J. (Ed.). (1967). From Frege to Gödel. Cambridge, MA: Harvard University Press.
Zermelo, E. (1908). Untersuchungen über die Grundlagen der Mengenlehre I. *Mathematische Annalen*, 65, 261–281 (Trans. by S. Bauer-Mengelberg as ‘‘Investigations in the Foundations of Set Theory I’’ and repr. in van Heijenoort (1967), pp. 199–215).
Zermelo, E. (1930). Über Grenzzahlen und Mengenbereiche: Neue Untersuchungen über die Grundlagen die Mengenlehre. *Fundamenta Mathematicae*, 16, 29–47 (Trans. by M. Hallett and repr. in W. Ewald (Ed.) (1996). *From Kant to Hilbert: Readings in the foundations of mathematics* (pp. 1208–1233). Oxford: Clarendon Press).