Hilary Putnam on the Philosophy of Mathematics

Daniel Waxman
National University of Singapore

Penultimate draft – please cite published version

1 Introduction

Hilary Putnam was one of the great philosophers of the twentieth century. This chapter aims to give an opinionated introduction to his contributions in the philosophy of mathematics, one of the central preoccupations of his career.[1] This is no easy task. One difficulty is that Putnam was a highly prolific author: one bibliography of his writings in the philosophy of mathematics and logic alone comes in at six pages (Cook and Hellman 2018, Chapter 2). His papers are also extremely dense: not in the sense of being unreadable – actually the opposite is the case (Putnam is a clear and lucid writer with a remarkable ability to make even the most venerable topics feel fresh) – but in the sheer ratio of arguments to pages. Another challenge is that Putnam was notoriously philosophically open-minded, so much so that some authors adopt the measure of adding subscripted dates to his name to make clear which time-slice of Putnam is under discussion.[2]

At a high level of description, Putnam's philosophical writings can be thought of as falling into three periods: a 'scientific realist' phase, dating from his earliest publications to around 1976, focused to a large extent on developing a post-logical-empiricist philosophy of science, mathematics, mind, and language; an 'internal realist' phase, from 1976 to around 1990, criticizing and exploring alternatives to the 'metaphysical realist' desire to attain a God's-eye perspective on a mind-independent world; and

---

[1] He also made substantial contributions to mathematics and mathematical logic proper, most notably as the P in the so-called MRDP theorem that resolved Hilbert's tenth problem. I will not discuss this work here: it will be the subject of the chapter by John Burgess.

[2] It is not even clear that this measure is sufficient: in Putnam 1975h (first published in 1967) he expresses scepticism that the continuum hypothesis has a determinate truth value; but in Putnam 1975c, *published in the same year*, he ridicules a closely related view. (In his defense it is likely that the former was written considerably earlier for a probably-delayed volume on Russell.) Putnam's views on the continuum hypothesis underwent further shifts: Putnam 1983c returns to something like an indeterminacy view; but he again endorses the determinacy of set-theoretic truth in his Reply to Field in Auxier, Hahn, and Anderson 2016.

a final, 'pragmatist realist' phase.[3] And even within these periods, Putnam's views changed considerably. As a result, there is very far from a single, neat and tidy, 'Putnamian' view in the philosophy of mathematics. One option in writing a handbook article on this topic would have been to try and trace out the detailed historical development of his views. But the number of twists and turns we would have needed to explore would run the risk of losing track of the overall significance of Putnam's work. So, instead, my approach will be to focus on the views of some influential Putnam time-slices on three main clusters of topics. First I discuss Putnam's 'quasi-empiricism', his views on the continuity of mathematics with empirical science, and his rejection of conventionalism (§§2–3). Then I discuss Putnam on 'mathematics as modal logic' and his doctrine of equivalent descriptions (§4). Finally, I discuss Putnam on objectivity and realism, including the 'Quine-Putnam' indispensability argument, Putnam's relation to contemporary nominalism, and the model-theoretic challenge and rejection of 'metaphysical realism' found in Models and Reality (§§5–6).[4] I hope that a sense of the diachronic sweep of his work, if not exactly its unity, will nevertheless emerge.

2 The continuity between mathematics and empirical sciences

It is difficult to understand Putnam's early work in the philosophy of mathematics apart from his scientific realist convictions. As he describes his view of science, 'one of our important purposes in doing physics is to try to state "true or very nearly true' laws . . . not merely to build bridges or predict experiences'.[5] For Putnam, mature scientific theories, especially physical theories, are to be regarded as true or approximately true descriptions of reality; no other picture of science could account for its otherwise miraculous success. (This, in short, is his famous 'no-miracles' argument for scientific realism). But how should *mathematics* be viewed by scientific realists, given the intimate connection between the two subjects? Many of Putnam's contributions before his 'internal realist' turn in around 1976 came from confrontations with this question; his answer was the development of a 'quasi-empirical' view of mathematics.

---

[3] A good overview is his remarkable intellectual autobiography in Auxier, Hahn, and Anderson 2016.

[4] Key papers outlining Putnam's quasi-empiricism include Putnam 1975i, Putnam 1975h, Putnam 1975c, Putnam 1975j, and Putnam 1971. The main source for 'mathematics as modal logic' is Putnam 1975c; on equivalent descriptions, see also Putnam 1983b; Putnam's later discussions of conceptual relativity, as in e.g. Putnam 1987, are also relevant. On indispensability the main source is Putnam 1971 (though for an interesting earlier paper see Putnam 1956). On internal realism see especially Putnam 1983c and the essays in Putnam 1981. The substantial (co-authored) introduction to the influential collection of readings Benacerraf and Putnam 1983 is also illuminating (although how much it reflects *Putnam's* views is not clear). For collections of essays on Putnam's views, including in philosophy of mathematics, see Auxier, Hahn, and Anderson 2016, Baghramian 2012, and Chakraborty and Conant 2022; the first two include replies from Putnam.

[5] Putnam 1971, p. 338.

Putnam's quasi-empiricism has a few components. What lies behind them all is the conviction that mathematics is continuous with science. This amounts to a kind of 'anti-exceptionalism' about mathematics (and logic) – a denial that there is anything methodologically, semantically, or epistemically special about mathematics that distinguishes it as a matter of kind from empirical theorising or that would warrant treating it as immune from revision or as governed by fundamentally different epistemic norms. The influence of Quine here is palpable; Putnam shares (or at least comes very close to) Quinean scepticism about the analytic and the apriori, and cites Quine's critique of conventionalism approvingly.[6]

In saying that mathematics is quasi-empirical, Putnam is not of course saying that mathematicians go around taking readings on some special kind of measuring device; he recognizes that the central activity involves giving proofs from axioms. But he firmly rejects the usual 'foundationalist' accounts of what makes those axioms justified or warranted. Against (some traditional forms of) platonism, the axioms are not self-evident, at least in some self-certifying sense, nor do we have some special faculty of intuition that gives us *a priori* access to them. Actually Putnam accepts that there is some role for intuition in mathematics, but this is also an aspect that he sees as continuous with empirical science, for he thinks that it plays an important (though fallible and uncertain) role there too, in for instance helping us to come up with theories to test in the first place.[7] And against conventionalism, the axioms are not 'rules of language' or analytic; axioms can (and, crucially for Putnam, have been!) revised in a way that he thinks rules of language cannot.

An extremely influential case study in Putnam's thinking is the rise of non-Euclidean geometry, which at one point he describes as 'the most important event in the history of science for the epistemologist'.[8] He draws many lessons from it, but in the philosophy of mathematics his main conclusion is that it shows the bankruptcy of the views of axioms mentioned above. Before Riemann and Lobachevski, Euclid's axioms might have seemed self-evident, or 'virtually analytic';[9] the thought that two parallel lines could meet would have seemed totally incoherent. But the development of alternative geometric theories in which the Parallel Postulate fail, and the development of General Relativity, which raised the possibility that such a theory may in fact describe the structure of the physical world, showed this to be untenable. Putnam did not deny that the Euclidean axioms may have seemed self-evidently true; his point is that, as he put it, 'our notions of what is 'self-evident' have to be subject to revision, not just in the

---

[6] See for instance Putnam 1975g and Putnam 1983a.

[7] Putnam 1975j, p. 67.

[8] Putnam 1975e, p. x.

[9] Putnam 1975f, p. 88

light of new observations, but in the light of new theories.'[10] No body of theory, even if it is apparently self-evident or centrally entrenched in our theorising, is in principle immune from revision. This applies not just to theories, like geometry, that historically arose with close ties to physics, but also to arithmetic and logic: 'I believe that under certain circumstances revisions in the axioms of arithmetic, or even of propositional calculus (e.g. the adoption of a modular logic as a way out of the difficulties in quantum mechanics), is fully conceivable.'[11] This was no idle threat: one of Putnam's time-slices famously advocated overthrowing classical logic in favour of a quantum logic in which conjunction fails to distribute over disjunction (later to be taken back); though, to the best of my knowledge, he never endorsed revising arithmetic.[12]

Another component of Putnam's quasi-empiricism involves the methodology of mathematics. He was an early proponent of the use of non-deductive methods, for instance enumerative induction, analogical reasoning, and inference to the best explanation, in mathematics. This can come in a weak form – according to which mathematicians can use non-deductive methods in a roughly heuristic way, for instance in generating hypotheses and influencing the direction of investigation – and a strong form – according to which these methods can positively confirm or justify mathematical claims. Putnam accepted both: at one point he suggests that claims like the Riemann hypothesis might be confirmed in the event that 'extensive searches with electronic computers [fail] to find a counter-example'.[13] Here again there is an analogy with empirical science: the results of particular calculations or proofs play the role of empirical observation reports in allowing the confirmation of generalizations of which they are instances.

Abductive methods also lie at the heart of Putnam's positive account of axiom adoption in terms of their necessity to science, which we will discuss at length in the context of his indispensability argument in §5. It is worth noting that Putnam was an early advocate of the possibility of what Maddy later called extrinsic justification of axioms: the view that necessity to *mathematics* itself, not only empirical science, could form the basis for a good abductive argument for an axiom.[14]

---

[10] Putnam 1975c, p. 50. See Putnam 1975b for an analogous discussion of necessity.

[11] Putnam 1975c, p. 50.

[12] For the defense of quantum logic see Putnam 1968; for the retraction see Putnam's reply in Baghramian 2012, pp. 265–80.

[13] Putnam 1975j, p. 61.

[14] See Maddy 1988. For instance in Putnam 1975i, p. 66, he approvingly discusses Zermelo's case for the axiom of choice on the basis of its use *within set theory*.

3 The rejection of conventionalism

At the very beginning of his career Putnam dallied with a conventionalist view of mathematics, likely under the influence of his logical empiricist teachers Reichenbach and Carnap. As he put it in Putnam 1956, p. 83, the axioms of mathematics are 'meaning conventions, or meaning postulates' and are 'empty of factual content'. But not long afterwards he came to firmly reject this view, ending up as one of its severest critics. Putnam's attack on conventionalism is sometimes assimilated to Quine's – Putnam was broadly sympathetic to Quine's scepticism of the analytic/synthetic distinction, and endorsed something along the lines of Quine's famous critique of Carnap in Truth by Convention – but he also had novel arguments that are worth considering on their own terms.[15]

I will discuss two of Putnam's arguments against conventionalism here. The first involves the revisability of our mathematical beliefs, a component of his quasi-empirical view that we have already encountered. According to conventionalism, to accept a mathematical claim is to adopt it as a 'rule of language' that governs the meaning of the distinctive vocabulary it contains, in some suitable sense.[16] But consider, Putnam urges, our attitude towards possible *revisions* of our mathematical claims, how we might view certain eventualities as *counting against* them. Consider the claim that there are no integer solutions to $x^3 + y^3 = z^3$ (a special case of Fermat's Last Theorem first proven by Euler). Suppose we cube 1769 and add the cube of 269 and discover that it is the cube of 1872. How would we react? Putnam's answer is that:

> I would first go back and look for a mistake in the proof of $x^3 + y^3 = z^3$. If I decided that the latter proof was a correct proof in First Order Arithmetic, then I would have to modify First Order Arithmetic – which would be shattering. But it is clear that in such revision purely singular statements ($5 + 7 = 12, 189 + 12 = 201, 34 \times 2 = 68$, etc.) would take priority over generalized statements ($x^3 + y^3 = z^3$). A generalized statement can be refuted by a singular statement which is a counter-example. Putnam 1975i, p. 9.

A similar example, discussed also in Putnam 1983a and Putnam 1994b, Chapter 28,

---

[15] Putnam's views on the analytic/synthetic distinction are slightly more subtle, but in ways that we will not pursue here; see e.g. Putnam 1975g. On conventionalism see e.g. Putnam 1975i, which explicitly targets Carnapian conventionalism (as well as the ontological characterization of mathematics as describing a realm of distinctively mathematical *objects*, which will be discussed shortly) and Putnam 1983a, p. 116. Putnam was also a spirited opponent of conventionalism in the philosophy of science: see for Putnam 1975a, although I won't be able to discuss this aspect of his work here.

[16] Perhaps only the axioms are really adopted directly as rules, in which case the conventionalist should say that we can accept mathematical claims in virtue of their being *logical consequences* of rules. Traditional (and contemporary) mathematical conventionalists have typically also been conventionalists about logical consequence too.

involves our finding an inconsistency in Peano Arithmetic. Our dispositions in these imagined cases contain a great deal of coherent internal structure – we have sophisticated and detailed preferences about how to revise our mathematical beliefs in the face of recalcitrant 'observations' (calculations). Of course, Putnam agrees, if a claim like that there are no solutions to $x^3 + y^3 = z^3$ is true, then it is mathematically and even metaphysically necessarily true. But that is all compatible with the epistemic possibility of finding a counterexample. (Indeed, our dispositions to make the relevant revisions might be triggered by our mistakenly but rationally believing that we have in fact found a counterexample).

Putnam's thought is that these facts bring the practice of mathematics into conflict with conventionalism. Here he is appealing to a further premise: that accepting something as a rule of language involves adopting the attitude that nothing counts against it (and is incompatible with sophisticated dispositions concerning ways to revise it). This premise, I take it, is where all the action is. Many conventionalists will want to deny it; on a Carnapian view, for instance, we can indeed compare 'frameworks' (= systems of conventions), and come to prefer one over another. Of course for Carnap we cannot do so on the basis that one framework is correct from an 'external' perspective, but we can do it on pragmatic, external grounds; and it is not clear why our preferences in doing this could not involve considerable fine-grained structure.[17] The question seems fundamentally to boil down to this: is meaning preserved through hypothetical revisions of this kind? For conventionalists the answer is no, since the rules of language are meaning-constituting. Putnam has two reactions to this conventionalist response. First, as he puts it, coming to change our mind in this way (for instance by finding what seems like a counterexample to a generalization) 'certainly seems to be Discovery in every sense of the word' Putnam 1975i, p. 11. One way to take the point is that the conventionalist line is not true to the phenomenology of the case: giving up a mathematical claim on such grounds feels more like changing one's mind on a factual question than simply changing the meaning of the constituent words. The second response is that this move threatens to trivialize conventionalism: 'any statement can be regarded as a 'rule of language' if we are willing to say 'well, they changed their rule' if speakers ever give it up.'[18] This is not entirely decisive: whether it succeeds will turn on the question of whether conventionalists can give a principled distinction between rules of language and other claims – in effect, whether they can make out an analytic/synthetic distinction. That, of course, is a big question – one on which Putnam had much to say, but which would take us too far into his views in philosophy of language to discuss further here.

---

[17] See e.g. Carnap 1937, p. 318; for a modern conventionalist perspective see Warren 2020, Chapter 28.II.

[18] Putnam 1975i, p. 11.

Putnam has another, related, argument against conventionalism worth mentioning. The starting point is the observation that some mathematical claims – for example, those that express that a certain calculation has a certain result, or that a certain logical system is syntactically consistent (i.e. no contradiction can be proven within it) have a "brute fact' character (somewhat analogous to the character of an observation statement in empirical science)…'.[19] A proof is a kind of syntactic object of the sort that can be written down or instantiated by a concrete inscription; so a consistency claim has the 'synthetic' content that no such concrete proof can be exhibited. But, as Gödel showed, claims about syntax can be represented by purely mathematical claims – by the arithmetization of syntax there is a sentence in the language of arithmetic, Con(PA), that is true if and only if Peano arithmetic is consistent. Putnam uses this as the basis for an objection:

> Once we grant that there is at least one mathematical fact which is not simply our stipulation …– that, for example, the consistency of our stipulations/practices is not itself just another stipulation or practice – then logical positivist/Wittgensteinian accounts of logical and mathematical truth are seen to be bankrupt. Putnam 1994b, p. 501.

In other words, if there is an objective, non-conventional fact about whether PA is consistent – as Putnam argues there is – and if Con(PA) expresses that fact, then the truth of Con(PA) cannot simply be a matter of convention.

Many philosophers have taken this objection to be conclusive.[20] However, the argument is the subject of some ongoing discussion. One response, due to Warren, argues that Putnam makes a subtle mistake. Although the arithmetical sentence Con(PA) expresses the consistency of arithmetic, Putnam seems to be assuming that it would continue to do so *even in a setting where different rules of language are in play*.[21] The alternative conventionalist picture is something like this: rules of language fix the meaning of some part of pure mathematics; but to apply that mathematics to a non-mathematical question (e.g. the outcome of some empirical calculation or the possibility of writing down a proof of a certain claim) requires a bridge principle roughly to the effect that the non-mathematical domain is isomorphic to or otherwise accurately modeled by the

---

[19] Putnam 1994b, p. 501.

[20] See for instance Koellner 2009, p. 84, who applies the point more generally against pluralism of the form that claims that both PA + Con(PA) and PA + $\neg$Con(PA) are equally good arithmetical theories – if the argument succeeds, then either PA + Con(PA) or PA + $\neg$Con(PA) is objectively unsound. See also Clarke-Doane 2020, Chapter 6.2.

[21] See also Azzouni 2023 for a similar argument. Interestingly Floyd and Putnam 2000, p. 628 later endorsed something like this point, attributed to Wittgenstein: 'the "translation" of [the Gödel sentence G of a theory T] as [G] is unprovable in [T] is not cast in stone, but is something that we have to give up in certain contexts'.

mathematical one. On this approach, pure mathematics remains pure; it is the bridge principles that are responsible for any 'synthetic' character it may have. If that account is right then the objection is undermined, for it shows how the mathematical sentence Con(PA) might be both true by convention and express a synthetic, non-conventional fact.

This response to Putnam turns on a particular account of applications: what Pincock 2011, p. 282 calls a 'two-stage' account, on which, first, pure mathematics is characterized independently of any applications, and only then, second, applications are introduced by means of further relations between the mathematics and the domain of application. By contrast, one-stage accounts of applications 'build them in' to the content of mathematics.[22] It is not entirely clear whether Putnam would have tried to appeal to a one-stage account here; there is evidence that he once rejected something like a two-stage picture of the applications of geometry.[23] But in other work (e.g. Putnam 1971) he says things that strongly suggest a two-stage account, and his discussions of Con(PA) and other 'synthetic' sentences of pure mathematics show no evidence of appealing to the distinction.

Where does this leave Putnam's argument? Warren's response shows that it is not decisive: the conventionalist can maintain that pure arithmetic is true by convention while locating any synthetic content in deniable bridge principles. However, Picollo and Waxman 2025 argues that the bridge principles are not so easily denied. It is plausibly essential to the conceptual role of arithmetic that induction be extended to include instances containing extra-mathematical vocabulary. But, combined with plausible theories of syntax, this entails that arithmetic and syntax are provably isomorphic. It follows that the arithmetical sentence Con(PA) is *provably equivalent to* a purely syntactic sentence that directly expresses consistency – a bridge principle of just the kind that Putnam's argument requires. If instances of extended induction, involving nonmathematical vocabulary, are essential to the content of arithmetic, it is by no means clear that it can be viewed as a 'pure' mathematical theory in the way the two-stage account requires.[24]

---

[22] Think of the neo-Fregean picture of arithmetic that gives pride of place to Hume's Principle (i.e. the application of numbers to counting) as in e.g. Wright and Hale 2001 and Wright 2000.

[23] 'It will not do to wave about the distinction between Euclidean space as an abstract mathematical object and physical space, i.e. the space in which bodies move. For what seems to be a necessary truth . . . is precisely that physical space is a Euclidean space' Putnam 1975e, pp. ix–x.

[24] For the full argument, see Picollo and Waxman 2025. We endorse Putnam's conclusion that pluralists about arithmetic cannot view syntax as objective, though unlike Putnam we do not regard this as a straightforward reductio of pluralism.

4 Mathematics as modal logic and equivalent descriptions

A recurring and central idea in Putnam involves the possibility of formulating mathematics in *modal* terms. In Putnam 1975c, Putnam offers two 'pictures' of mathematics: mathematics as set theory and mathematics as modal logic, which he viewed as 'equivalent descriptions'. Mathematics as set theory is the familiar conception of mathematics as involving quantification over a universe of mathematical objects, which Putnam takes to be reducible (in some sense or another) to sets.[25] The alternative picture, mathematics as modal logic, instead regiments mathematical claims in terms of a modal operator expressing a notion of mathematical necessity. I will first outline the mathematics as modal logic picture before discussing Putnam's doctrine that the two pictures constitute equivalent descriptions.

Take the claim that there is a counterexample to Fermat's last theorem, i.e. that there are positive integers a, b and c such that $a^n + b^n = c^n$ for some integer $n > 2$. As stated this is an explicitly existential claim, asserting the existence of numbers possessing various properties; write it as $\exists x\phi(x)$. But, Putnam says, we can give an alternative regimentation as follows. Let us suppose that AX is the conjunction of finitely many arithmetical axioms sufficient to disprove Fermat's last theorem (if it is false).[26] These axioms will contain terms for addition and multiplication relations; suppose they are S and T respectively. And now suppose that $\Box$ expresses a suitable notion of mathematical necessity (on which more in a moment). Then, Putnam says, the same content as $\exists x\phi(x)$ can be expressed by:

$$\Box(AX(S,T) \to (\exists x\phi(x))^*)$$

where $*$ is a function that rewrites a sentence in the language of arithmetic in terms of S and T. Roughly the claim is that as a matter of mathematical necessity, for *any* objects (Putnam thinks of these as possibly concrete) and relations on them that behave in the way that the usual arithmetical axioms tell us that addition and multiplication behave, there will be a counterexample to 'Fermat's last theorem' reformulated in terms of the relevant relations.[27]

Now, this approach raises many technical and philosophical questions. On the technical side: Putnam's example is special in a few respects, and it is not entirely obvious

---

[25] The reduction to sets is famously problematized in Benacerraf 1965. Putnam views this as yet another example of equivalent descriptions; he briefly but explicitly discusses a related notion of 'equivalent constructions' at the end of Putnam 1971, where he notes that 'numbers can be constructed from sets in more than one way' and that mathematics might equally well proceed by taking functions, not sets, as the basic objects (see Button 2025 for a recent development of this idea).

[26] That there are finite such axioms relies on the fact that the statement is Π_1.

[27] Putnam 1975c, p. 49 suggests that the modal translation of a claim will also add that it is *possible* for there to be suitable objects and relations on them.

how it can be generalized. How are arbitrary arithmetical statements to be handled? (We will not in general have a guarantee that some particular finite theory will be sufficient to decide them, in which case we may not be able to even formulate the required conjunction that figures in the antecedent of the conditional.) How can sentences in parts of mathematics other than arithmetic be regimented? This question is particularly pressing for the language of set theory, given that the modal picture is supposed to be able to capture any fact that mathematics as set theory can. And how can mathematics as modal logic handle the applications of mathematics?

Putnam does not say a great deal to resolve these questions. While he did sketch an account of how set theory might be given a modal translation in terms of 'standard models' (Putnam 1975c, p. 54), I think it is fair to say that the technical details were not entirely worked out. As an example, consider the claim that for every ordinal there is a larger ordinal, $\forall\alpha\exists\beta(\alpha < \beta)$. In the mathematics as modal logic picture, this would come out as roughly:

> Necessarily, for any standard model M, and for any object α that behaves in an ordinal-like way in M, it is possible that there is a standard model M' that expands M, with an object β that behaves in an ordinal-like way in M', such that $\alpha < \beta$ according to M'.

A much fuller development can be found in Hellman 1989 who develops a modal-structuralist theory of mathematics along Putnamian lines. Hellman there appeals to second-order theories: since second-order Peano arithmetic and Zermelo-Fraenkel set theory are finitely axiomatized theories, this allows their axioms to be expressed in a single conjunction; schemata (as with e.g. S and T above) can be replaced by explicit quantification into the relevant positions. The use of second-order notions is also arguably necessary to express Putnam's notion of a 'standard' model. A considerable amount of work has subsequently been done on modal interpretations of mathematics; for some of the highlights, see Burgess and Rosen 1997, Parsons 2007, Linnebo 2013, Studd 2013, Warren 2017, Roberts 2019, Scambler 2021, Button 2021, Berry 2022, Hamkins and Linnebo 2022, and Barton 2024.

On the philosophical side, one of the major questions concerns the interpretation of the relevant modality. Putnam does not say a great deal here, either. He does say that mathematical necessity can be compared to the set-theoretic notion of being true in all models; but this can presumably be at best a heuristic if mathematics as modal logic is supposed to be intelligible without presupposing concepts from mathematics as set theory. I think it is fair to say that this is still very much an open question in the Putnam-inspired literature on modal approaches to mathematics. This is not a minor lacuna: if we cannot give a substantive account of the relevant modality, the natural worry is that ’mathematics as modal logic’ is less an alternative picture than a notational variant whose content parasitically depends on the set-theoretic picture. Hellman appeals to a notion of ‘logico-mathematical’ possibility and is explicit that it should not be assimilated to physical possibility, but does not say much more about its interpretation; Field 1989 gives a broadly inferentialist development (which he puts to work in giving a nominalist account of logical consequence); and other writers, most explicitly Linnebo 2018b (who traces it back to Parsons) consider an *interpretational* account which considers different possible interpretations of the mathematical vocabulary.[28]

Putnam’s central philosophical claim about mathematics as modal logic is that it constitutes an ‘equivalent description’ to mathematics as set theory. This doctrine of equivalent descriptions was a recurring theme of Putnam’s work: Linnebo recounts that it was frequently discussed in his seminars at Harvard during the period 1996-2002.[29] The basic idea is that ‘what is in some sense the same fact can be expressed in two strikingly different ways.’[30] Putnam applied this idea at various times to a wide range of alleged theoretical disputes:

- Wave vs particle formulations of quantum mechanics;
- Mathematics as modal logic vs mathematics as set theory;
- Nihilism vs universalism about mereological composition;
- Taking sets as opposed to functions as a primitive mathematical concept;
- Different ‘reductions’ of e.g. natural numbers to particular sets.

The first of these examples, taken from his teacher Hans Reichenbach, is the source of one of his earliest expositions of the idea:

> The description of the world as a system of particles . . . may be associated with a different picture than the description of the world as a system of waves . . . ; but the two theories are thoroughly intertranslatable, and should be viewed as having the same physical content. The same fact can be expressed either by saying that the electron is a wave with a definite wavelength A or by saying that the electron is a particle with a sharp momentum p and an indeterminate position. Putnam 1975c, p. 45.

Putnam is clear that the notion of two facts being ‘the same’ is *not* supposed to be understood in terms of the synonymy of the sentences expressing them; that would

---

[28] See also Laan 2025 for criticism.

[29] Linnebo 2018a.

[30] Putnam 1975c, p. 44.

be implausible in the quantum mechanics example, since it is hard to see how any sentence about wavelengths could be synonymous with a sentence about momentum. The question, then, is how to positively understand the notion of theoretical equivalence. The idea has certainly not been found persuasive or even intelligible by everyone. For instance John Burgess complains that:

> As the Council of Nicæa declared that the Father and the Son are somehow the same and yet somehow different, so Putnam declares the 'mathematics as set theory' and 'mathematics as modal logic' pictures . . . are somehow the same and somehow different. I find the Nicene Creed easier to understand than Putnam's notion of equivalent descriptions. Burgess 2018, p. 136.

In Putnam's defense, he says considerably more about the relation of equivalence even in his early presentations. For example, in Putnam 1975c, p. 46 he gives more substantial criteria:

- (i) 'the primitive terms of each admit of definition by means of the primitive terms of the other theory, and then each theory is a deductive consequence of the other'
- (ii) 'there is no particular advantage to taking one of the two theories as fundamental and regarding the other one as derived';
- (iii) 'The two theories are, so to speak, on the same explanatory level. Any fact that can be explained by means of one can equally well be explained by means of the other.'

Although Putnam uses the word 'definition' in (i), his explicit denial that equivalence is to be understood as synonymy suggests that this is intended as a *formal* criterion; and indeed this is substantiated in his most comprehensive discussion of the issue in Putnam 1983b, where he defends what logicians have come to call mutual relative interpretability as a necessary, formal condition on equivalence.[31] This requires a little explanation.

A theory T_1 *relatively interprets* T_2 if there is a mapping from the primitive symbols of T_2 into expressions of T_1 such that, when we carry out the translation on sentences induced by that mapping, every theorem of T_2 becomes a theorem of T_1.[32] For example, Zermelo-Fraenkel set theory relatively interprets Peano arithmetic via the map that

---

[31] It is actually unclear whether mutual interpretability holds in all of the examples that Putnam wants to diagnose as equivalent descriptions. For instance, a recurring example is the rejection of certain ontological disputes as insubstantial – for instance between universalist and nihilist views of mereology – on the grounds that they are equivalent descriptions. But the relevant theories are not exactly mutually interpretable; Warren 2015 proves a result that comes close, but requires a slightly non-standard formulation (due to the use of plural resources).

[32] The canonical reference is Tarski, Mostowski, and Robinson 1953.

takes (expressions denoting) the natural numbers to the set ω, 0 to $\emptyset$, $S(n) \mapsto x \cup \{x\}$, and the usual definitions of $+$ and $\times$ on ω. A relative interpretation can be viewed as a way of simulating one theory within another by means of 'suitable definitions' (again, understood in the purely formal sense of a mapping from one language to another, without any connotation that meaning is preserved). Two theories are mutually relatively interpretable if each relatively interprets the other.

Relative interpretability is generally taken to be too weak to capture a notion of theoretical equivalence in a full-blooded sense. For instance, PA is mutually interpretable with PA + ¬Con(PA); but it is clear enough that these two theories cannot reasonably be viewed as expressing the same facts in any plausible sense.[33] Of course, none of this is necessarily problematic for Putnam, who never to my knowledge advocated mutual interpretability as sufficient; it always went hand-in-hand with an explanatory criterion as in (ii) and (iii) above.[34]

Putnam is explicit that the explanatory criterion is informal; but although the notion of explanation is informal, he thinks that it has enough content (and generates enough agreement in particular cases) to give rise to a workable criterion.[35] His most fleshed out version is as follows. Start by selecting some class of phenomena that we desire our theories to explain. Then we assess whether two mutually interpretable theories are equivalent by asking whether the translations – induced by the 'definitions' of the vocabulary of one theory within the other – preserve explanations for all phenomena in the relevant class. There is bound to be some slack in this process, because it relies on a prior decision about which phenomena are in the class to be explained; but elsewhere Putnam suggests that the appropriate methodology here is to defer to the explanatory standards of the subject in question (Auxier, Hahn, and Anderson 2016, p. 85).

Let us briefly return to the claim that mathematics as modal logic and mathematics as set theory constitute equivalent descriptions. Partly because he did not work out the technical aspect of the modal picture in much detail, Putnam never gave a proof of mutual interpretability or any other kind of formal equivalence between modal and non-modal theories. Later developments have clarified matters substantially.[36] Roberts 2019 shows that there are certain natural modal and non-modal set theories and a Putnam-inspired modal translation pt such that the modal theory proves ϕ^{pt} if

---

[33] The result is in Feferman 1960. The point is especially forceful for those who, like the Putnam time-slice discussed above, criticize conventionalism on the grounds that it misdiagnoses the source of truth of ¬Con(PA).

[34] Somewhat oddly this point seems to have been missed in some of the literature criticizing Putnam. For instance Barrett 2020 argues that, despite the explicit embrace of mutual interpretability in Putnam 1983, the relation of definitional equivalence – a stronger criterion of equivalence between theories – would better fit Putnam's purposes. (A similar complaint is made in Linnebo 2018a). But the arguments seem just to amount to the insufficiency of mutual interpretability on its own.

[35] Putnam 1983b, p. 39.

[36] Burgess and Rosen 1997 contains the most comprehensive technical study of modal approaches.

and only if the non-modal theory proves ϕ. And Button 2021 has shown that a levels theory (which axiomatizes the stages in the set-theoretic hierarchy) is what he calls 'nearly synonymous' (a relatively strong form of formal equivalence, close to definitional equivalence) with a natural modal theory of stage formation. Although neither of these results demonstrate mutual interpretability exactly, they nevertheless disclose close formal equivalences between modal and non-modal theories, and so, perhaps, are of use to a defender of the equivalent descriptions thesis.

Turning from the formal criterion to the explanatory criterion, it is perhaps more surprising that Putnam nowhere seems to have given an argument that the two pictures are *explanatorily* equivalent. One difficulty here is how to think of the relevant class of phenomena to be explained. Mathematics as set theory seems capable of providing satisfying explanations of mathematical facts *non-modally construed*, and likewise mathematics as modal logic for mathematical facts *modally construed*; but it is hard to see what could be a 'neutral' class of facts that are equally well explained by both.

In fact, Putnam himself later came to retract the claim (although he maintained the general doctrine of equivalent descriptions) in a series of blog posts near the end of his life, because he doubted that mathematics as modal logic preserved explanations in a way that would be recognized as such by the internal standards of mathematics.[37] As he puts it, 'the modal-logical version . . . is not one mathematicians are even *aware* of.' While this is true, it is arguably not conclusive: it is hard to deny that modal notions play at least a heuristic role in mathematical practice, as in e.g. informal motivations of the iterative hierarchy of sets. So, even defenders of the equivalence thesis who adhere to Putnam's use of internal standards to adjudicate explanatory equivalence may have further resources to appeal to here. For interesting further discussion see Linnebo 2018a and Button 2021.

5 The 'Quine-Putnam' indispensability argument

The typical undergraduate with a course in the philosophy of mathematics is most likely to have heard Putnam's name in connection with the so-called Quine-Putnam indispensability argument that seeks to establish the existence of mathematical objects. In fact, this is probably the most prominently discussed positive argument for a realist view of mathematical ontology. Here is a now-canonical version, due to Colyvan 2001:

(P1) We ought to have ontological commitment to all and only the entities that are indispensable to our best scientific theories.

---

[37] Putnam 2014. He also argues there that, in light of the problem of multiple reductions from Benacerraf 1965, mathematics as modal logic is in fact preferable to mathematics as set theory as a rational reconstruction of mathematical practice.

(P2) Mathematical entities are indispensable to our best scientific theories.

(C) We ought to have ontological commitment to mathematical entities.

In Colyvan's presentation, Premise 1 is motivated by naturalism (our commitments should be limited to *only* the entities indispensable for science) and confirmational holism (we ought to be committed to *all* such entities; empirical confirmation is received by theories as a whole, so there is no principled way to say that only concrete entities get support). Premise 2 is motivated by the fact that mathematics figures in our most successful mature scientific theories. It is clear enough that Quine advanced an argument along something like these lines.[38] It is a paradigmatic application of Quinean ontological methodology: questions of existence are to be resolved by considering the quantificational implications of (suitably regimented formulations of) one's best overall theories of the world. However Putnam's relationship to the argument is more complex. He comes closest to Quine's position in *Philosophy of Logic*, his most extended discussion of indispensability, where he writes that:

> . . . I have been developing an argument for realism along roughly the following lines: quantification over mathematical entities is indispensable for science, both formal and physical; therefore we should accept such quantification; but this commits us to accepting the existence of the mathematical entities in question. Putnam 1971, p. 347.

However, while Putnam shared Quine's scientific realism, his views differed from Quine's in important ways. One source for the differences is a paper (Putnam 2012) based on a talk given in 2006, where Putnam explicitly disavows the Quinean argument as formulated by Colyvan.[39] Of course, we should not necessarily defer to a philosopher's interpretation of their own views almost 40 years on. But in this case, there is substantial evidence that Putnam had a more subtle position in mind all along.[40]

First, some brief background. A main thread running through contemporary philosophy of mathematics is an opposition between realist and anti-realist perspectives. To a rough first approximation, realists view mathematics as a body of objective, determinate facts obtaining independently of the language, thought, conventions, and practices of mathematically engaged subjects; the picture is that mathematics is out there to be discovered, not in any interesting way a human invention. The debate has come to involve two slightly more precise theses. One is an *ontological* question: do

---

[38] There is no single canonical formulation in Quine, to my knowledge, but there are scattered remarks from the late 40s onwards. Putnam 1971 is one of the first explicit formulations.

[39] See also Auxier, Hahn, and Anderson 2016, pp. 60–4.

[40] For helpful discussion of Putnam's views see Liggins 2008 and Lebrun manuscript.

mathematical objects – numbers, sets, functions, and so on – exist, abstractly and independently of us? The other concerns the *truth-values* of mathematical statements: do mathematical statements have objective truth-values independently of our capacity to prove or refute them?[41]

The main difference between Putnam's version of the indispensability argument – at least in his later re-telling, and, I think, with some plausibility, all along – versus Quine's (and the standard contemporary formulations which are typically closer to Quine) is that Putnam's conclusion is not primarily an ontological one. As he put it, 'my "indispensability" argument was an argument for the objectivity of mathematics in a realist sense'.[42] Here is one way to represent the argument:

(P1) We ought to accept truth-value realism about theories that are indispensable to our best scientific theories;

(P2) Mathematics is indispensable to our best scientific theories;

So: (C) We ought to accept truth-value realism about mathematics.

The most salient difference is the focus on truth-value realism as opposed to ontology. This is motivated pretty immediately by the doctrine of equivalent descriptions. Mathematics as set theory and mathematics as modal logic are supposed to be equivalent formulations; but the former is committed (in Quine's sense) to an expansive set-theoretic ontology while the latter is not. It follows that ontology is not invariant across different equivalent descriptions of the same facts; it is more like an artifact of one particular presentation than an intrinsic feature of the domain under description. (For similar reasons, it would be a misreading to view Putnam as trying to eliminate mathematical ontology in the vein of modern nominalists; after all mathematics as set theory is a perfectly adequate description! A better reading is that he is attempting to deflate ontological questions altogether.) So Putnam's thought is that taking mathematics seriously, in the way that we should take seriously science and theories that are indispensable to it, is best expressed by truth-value realism; and a sentence's possessing an objective truth-value plausibly *is* a feature that is preserved by translations between equivalent descriptions.

Putnam's justification of the indispensability premise is worth remarking upon.[43] He considers Newton's law of gravitation as an example and argues that there is no way it could be fully expressed nominalistically. One of his arguments appeals to the thought (which he apparently takes to be a commitment of nominalism) that there are

---

[41] See e.g. Shapiro 2000.

[42] Putnam 2012, p. 183.

[43] His fullest discussion is Putnam 1971, though see also Putnam 2006.

only finitely many concrete entities; however, this has not been adopted by subsequent nominalists (most notably Field 2016). There is a more interesting argument, however: Putnam sketches a rough version of what is, in effect, a measurement-theoretic account of how a numerical distance function can be defined over space-time points on the basis of a physical relation of congruence that holds between intervals of points. But, he thinks, making sense of this (what he calls the 'numericalization' of distance) requires substantive mathematics: specifically quantification over functions. From a contemporary perspective, steeped in *Science Without Numbers* and the literature in its wake, it is striking how close Putnam comes to endorsing much of the picture of applied mathematics that underlies Field's program, even if he demurs from the nominalism. Actually in Putnam's earliest paper in the philosophy of mathematics, Putnam 1956, he endorsed (to my knowledge the earliest formulation of) the claim that mathematics is conservative over nominalistic science.[44]

There is a large and valuable literature on indispensability arguments. Despite the differences I have emphasized, much of this work is relevant not only to standard, ontological versions but also to the version I have been attributing to Putnam. Work whose relevance arguably carries over includes: discussions of what it is for a theory to be dispensable;[45] discussions of precisely how much mathematics is dispensable;[46] and nominalist objections to the premise that mathematics is indispensable.[47] At the same time, Putnam's version seems to have novel resources to respond to at least some of the objections to standard presentations of the argument, for instance: objections to confirmational holism;[48] and objections that the indispensability-based epistemology does not account for the obviousness of mathematics.[49] In contrast to those who attempt to use indispensability considerations to provide a positive empiricist epistemology of mathematics, Putnam views his argument as merely *a constraint on* the epistemology of mathematics from a scientific realist perspective.

There is one point of tension worth briefly mentioning that arises for a distinctively Putnamian combination of views. A major traditional threat to indispensability arguments is the 'hard road' nominalist programme that aims to provide adequate refor-

---

[44] I found this striking: if you had asked me before I wrote this chapter, I would have said the claim was first formulated and discussed in Field; many informants specializing in the philosophy of mathematics seem to have had the same impression. As a referee helpfully reminded me, conservativeness claims should strictly be formulated in terms of the conservativeness of one *theory* over another. Putnam specifies (or at least gestures at) the mathematical theory with which he is concerned – a theory of classes – but is unspecific about the nominalistic theories in question.

[45] E.g. Colyvan 2001 and Burgess and Rosen 1997.

[46] E.g. Feferman 1992.

[47] E.g. Field 2016, Melia 2000, and Leng 2010.

[48] For instance Sober 1993. It is worth noting that Putnam explicitly disclaims any commitment to holism in his argument: see Putnam 2012.

[49] Parsons 2007.

mulations of scientific theories in a way that does not use mathematical vocabulary. As proponents of the standard Quinean indispensability argument often remark, typical scientific theories are formulated in *platonistic* terms, appealing promiscuously to mathematical ontology; in Putnam's terms, they seem to presuppose something like the 'mathematics as set theory' picture. If Putnam is right that 'mathematics as modal logic' is an equivalent description, there must, presumably, be some account given of how applied mathematics can be carried out in this setting. (As I mentioned earlier, he does not go into any real detail about how this might be done). The most straightforward way to put modal mathematics to work in applications is to separate a nominalistically statable 'pure' physical theory from a distinct mathematical component. For example, one might try to formulate a purely physical theory of space according to which space is Euclidean, and then combine it with modal results of the form: in any possible Euclidean structure, theorem T holds.[50] This would allow the use of mathematics, construed modally, to obtain physical consequences. But now there are two problems. One is that the procedure of separating theories into pure and mathematical components would seem to require the success of the technical part of the hard-road nominalist project – a controversial and non-trivial commitment. The other is that, supposing the procedure *can* be carried out, the conservativeness of mathematics over 'pure' physics would appear to show that the mathematical component of the theory is dispensable after all (on at least one plausible reading of that claim). Of course, this line of thought is not conclusive; but I think it is fair to say that anyone who wants to endorse both Putnam's indispensability argument *and* his thesis that there are equivalent descriptions of mathematics has their work cut out.

6 Models and Reality

Putnam's most significant contribution to the philosophy of mathematics from his 'internal realist' phase, and perhaps altogether, is Putnam 1983c, published in 1980 but delivered as a Presidential Address to the Association for Symbolic Logic in 1977. Putnam's ambitions are no less than the refutation of 'metaphysical realism', a view that he characterizes around the same time as follows:

> the world consists of some fixed totality of mind-independent objects. There is exactly one true and complete description of 'the way the world is'. Truth involves some sort of correspondence relation between words or thought-signs and external things and sets of things . . . its favorite point of view is a God's Eye point of view. Putnam 1981, p. 49

---

[50] For a related discussion, see Field 1989, p. 253.

Putnam's main target in the paper is what he calls 'moderate' versions of metaphysical realism, which attempt to explain truth and reference without postulating 'non-natural mental powers'.[51] His argumentative strategy is roughly as follows. According to metaphysical realists, the best account of truth and reference for a language is to be given in terms of an *intended model*. Putnam appeals to certain mathematical theorems to argue that, if there is an intended model, there will also be various *unintended* models too. Then moderate realists, in particular, face the constraint of explaining in broadly naturalistic terms *why* these models are unintended; unlike what Putnam calls their 'extreme Platonist' rivals, they cannot appeal to a mysterious, primitive, notion of directly 'grasping' or 'understanding' the intended model. But – in the most controversial step – Putnam argues that moderate realists cannot possibly give the kind of explanation that is required. So, he concludes, to save scientific realism, metaphysical realism must be given up in favour of his preferred internal realist picture.

To generate 'unintended' models, Putnam appeals to three kinds of results:

- Permutation theorems. If a theory has a model, then it has distinct isomorphic models (roughly, the result of permuting the referents of constants and correlatively permuting the extensions of predicates).[52]
- The Löwenheim-Skolem theorem. If a first-order theory has an infinite model, then it has models of every infinite cardinality.
- Constructivization results. Putnam appeals to a result which he states as follows: ZF + V = L has an ω-model which contains any given countable set of real numbers.[53]

Roughly speaking, permutation models disagree with the base model about what terms refer to; Löwenheim-Skolem models disagree about the cardinality of the universe; and constructivization models disagree about which real numbers are constructible. The dialectic is something like this: the metaphysical realist offers an allegedly intended model of their theory; Putnam applies one of his theorems to generate a new model

---

[51] Putnam 1983c, p. 1. Putnam does not say much against the alternative 'platonist' form of realism, but his distaste for it is clear.

[52] These are discussed more fully in Putnam 1981, Chapter 2 and mentioned briefly at the end of Models and Reality (p. 24).

[53] This is the most fraught of the three: Putnam's claim cannot be a theorem of ZF, assuming it is consistent, because it entails the existence of a model (and hence the consistency) of ZF. Bays 2001 argues that this is a serious problem for Putnam, and that his other model-theoretic arguments suffer a similar flaw. Button 2011 responds (on behalf of the first two kinds of construction) that all Putnam needs is the conditional claim that *if* a theory has a model *then* it will also have unintended models of the relevant sort, and that conditional theorems to this effect are provable in weak metatheories. Kanamori 2018 extends this style of argument to the constructivization argument.

that makes the same theory true but which, by the realist's lights, has different representational content. If these kinds of models cannot be ruled out as unintended, then the moderate realist picture is in trouble: our thought and talk is not in a position to discriminate between the possibilities that they appear to express.

Putnam's argument that moderate realists cannot rule these models out as unintended is based on a certain metasemantic picture to which he thinks they are committed. Meaning and reference are fixed by 'the total use of the language', which includes (i) theoretical constraints (constraints on interpretation that arise from the sentences we accept or perhaps are disposed to accept in some idealized 'limit' of inquiry) and (ii) operational constraints (arising from connections between the use of language and empirical observations).[54] But, Putnam thinks, neither sort of constraint can rule out the putatively unintended models that his theorems generate.

His argument for this claim is somewhat unclear; he *appears* to be thinking that a model satisfies all operative theoretical and operational constraints if it assigns the right truth-values to sentences.[55] This is closely related to the notorious 'just more theory' maneuver that Putnam uses to respond to the thought that additional constraints, e.g. a causal theory of reference, might help the metaphysical realist. The response (Putnam 1983c, pp. 17–8) is that any further constraints can be regarded only as more sentences to be added to the theory that the relevant models must make true. This may be the most controversial move of the paper: the most prominent criticism is that it begs the question against the realist, who will claim that to impose e.g. a causal condition on reference is to impose an *external* condition, not merely to add more sentences to the theory, up for reinterpretation.[56]

Putnam's arguments against realism were intended to work *globally*. But, for basically the reasons I have just outlined, they have not attracted many adherents. However, mathematics is a special case: mathematical objects are, on standard views, abstract and causally inert. So even if external constraints on reference work as a response to Putnam in general, they do not seem to help in the mathematical case, and so a *local* version of the argument nevertheless has considerable force.[57] Anyone attracted to moderate realism consequently still faces the serious challenge of explaining, on naturalistic grounds, how *mathematics* in particular gets its content. This cluster of issues has received considerable discussion in recent years. A particularly pressing version of the challenge arises for arithmetic. It is very natural to think that there is an intended model (or isomorphic class of models), whose domain is the ω-sequence of natural numbers $\mathbb{N} = \{0, 1, 2, \dots\}$, each obtained from 0 by iterating the successor

---

[54] Putnam 1983c, p. 4.

[55] See the discussion of the interpretation 'OP' in Putnam 1983c, p. 13.

[56] See for instance Devitt 1983 and Lewis 1984; though see also Button 2013, Chapter 4 for defense.

[57] Although for a contrary view, see Soysal 2020.

function finitely many times. But, using the Löwenheim-Skolem theorem or otherwise, all kinds of non-standard models satisfying the first-order axioms of arithmetic can be cooked up. In this context, Putnam's challenge becomes that of explaining how these non-standard models can be excluded as interpretations of our mathematical vocabulary. And, of course, similar problems arise for other parts of mathematics such as set theory.

Adding new (first-order) axioms will not obviously help, because the expanded theory will also have non-standard models. A different and much-discussed approach attempts to meet the challenge by appealing to model-theoretic categoricity theorems. A theory is categorical if any two of its models are isomorphic. By the Löwenheim-Skolem theorem, no first-order theory with an infinite model is categorical. But the situation for some theories formulated using second-order logical resources (i.e. quantification into predicate position) is different. Second-order arithmetic, for instance, is categorical in the sense that any two *full* models – in which the interpretation of the second-order quantifiers ranges over *all* subsets of the domain – are isomorphic.[58] But as Putnam points out (Putnam 1983c, p. 23), second-order theories also have so-called Henkin models, on which second-order quantifiers range over some restricted set of subsets of the domain; and categoricity results fail if Henkin models are considered.[59] So, it seems, an appeal to categoricity theorems requires some prior justification for the exclusion of Henkin models and the restriction to full models – a challenge of comparable difficulty to the original problem.

One family of responses appeals to 'internal' categoricity results that do not require ascending to a stronger metalanguage. The idea is that arithmetic or set theory can, in a sense, characterize its own subject matter from within: for instance, any two structures satisfying the Peano axioms with open-ended induction are provably isomorphic in a theory that combines them. Such results aim to show that the determinacy of arithmetic does not depend on a prior grasp of 'the' standard model, but rather emerges from the inferential resources implicit in arithmetical practice itself. Whether internal categoricity theorems can fully meet Putnam's challenge remains contested and a topic of much discussion.[60]

Another approach, more in the spirit of Putnam's operational constraints, is to enlist our capacities to refer to physical objects as a means of ruling out non-standard models. The most prominent proposal of this kind, Field 2001, Chapter 12, contends

---

[58] Second-order Zermelo-Fraenkel set theory is *quasi-isomorphic* in the sense that its models are either isomorphic or one is isomorphic to a proper initial segment of the other.

[59] See also Weston 1976 and the appendix to Field 2001, Chapter 12.

[60] For discussion of internal categoricity and its philosophical significance, see McGee 1997, Lavine Unpublished, Field 2001, Chapter 12, Parsons 2007, Chapter 8, Button and Walsh 2018, Part B, Maddy and Väänänen 2023, Picollo and Waxman 2023.

that *if* we can refer to a physical ω-sequence, non-standard models can be ruled out (roughly, we could express a notion of finiteness in terms of 'corresponding one-to-one with some member of the sequence').[61] Of course, Field's approach is hostage to the cosmological assumption that there is an empirical ω-sequence, and to the semantic assumption that we can refer to it; presumably the Putnam of Models and Reality would take exception to the latter.[62] Many years later (Auxier, Hahn, and Anderson 2016, Reply to Chapter 2) Putnam responded to Field, endorsing in effect an application of the 'just more theory' move: he criticizes Field's cosmological assumptions for *using* the notion of finiteness and so being susceptible to reinterpretation in a non-standard model.[63] Those who are not sympathetic to the 'just more theory' response have a natural reply: Field's point is that, assuming that things in the world are a certain way (and we have certain referential capacities), that imposes further *external* constraints – not merely new sentences added to our theory – that rule out non-standard models. Admittedly, if things in the world are not that way (and Putnam's challenge cannot be resolved by other means), then we could not express the possibility that they are. But it is unclear how this undermines Field's claim.

Let me finally turn to Putnam's 'internal realist' view as a response to the model-theoretic challenge of Models and Reality. Putnam spent considerable effort developing the view in the late 1970s and 1980s; my treatment here will be confined to a brief outline and some remarks about its impact in the philosophy of mathematics. I emphasize that this terrain is very murky indeed.

As I understand it Putnam's central move is the rejection of the picture of semantics, assumed by most of the previous discussion, on which explaining the meaning of a language amounts to showing how its use determines a particular intended model:

> The language, on the perspective we talked ourselves into, has a full program of use; but it still lacks an interpretation ... To speak as if this were my problem, "I know how to use my language, but, now, how shall I single out an "interpretation?" is to speak nonsense. Either the use already fixes the "interpretation" or nothing can. Putnam 1983c, p. 24.

In place of the model-theoretic picture of semantics, Putnam expresses sympathy for a 'non-realist' semantics that he attributes to Dummett, according to which language-mastery is to be explained not in terms of truth-conditions (truth in the intended model) but in terms of verification conditions (in the case of mathematics,

---

[61] For a refined version of this approach that appeals only to the *physical possibility* of an ω-sequence, see Berry 2021.

[62] Field's suggestion is intriguingly anticipated by Putnam 1975h, p. 25 and rejected on the grounds that it makes mathematics a posteriori.

[63] Putnam attributes the point to Bueno 2005.

canonically: proof).[64] Dummett's development of this idea, famously, involved rejecting classical logic and embracing intuitionistic logic in its place. Putnam by contrast advocates what he calls a *liberalized* version of the intuitionist standpoint: one which accepts classical *reasoning* even while rejecting classical (model-theoretic) semantics.

The result, Putnam thinks, is to deflate questions of 'intended models' altogether.[65] Non-realist semantics allows us to say that, trivially, a model that gives 'natural number' a non-standard extension is unintended: we recognize this 'from the description through which it is given'; 'models are not lost noumenal waifs' but 'have names from birth'.[66] The best interpretation of these remarks, in my view, is that the very methods we use to construct a non-standard model M (such the Löwenheim-Skolem theorem) allow us to establish claims like 'M's natural numbers has elements that cannot be obtained by finitely many iterations of the successor operation from 0'; and since the natural numbers do *not* have such elements, that is enough to conclude that M is unintended.

Does this line of thought solve the challenge with which Putnam began? It certainly seems right that once one takes the step of rejecting the model-theoretic picture of meaning, concerns about how we pick out the intended model lose their force. (It is worth pointing out that Putnam's preferred, more-or-less verificationist, alternative to the model-theoretic picture is not the only alternative; other views in the broadly inferentialist, use-theoretic tradition, have largely the same benefit). Of course, it is a major question whether or not to take that step. There are also other, closely related, challenges about the metasemantics of mathematics. To my mind, the most pressing involve using Gödelian incompleteness results – results about the limits of *proof*, not our ability to pick out models – to argue that mathematics is indeterminate (in a sense that does not essentially use model-theoretic notions).[67] How exactly *these* challenges should be developed, and whether they can be resolved by a view in the ballpark of Putnam's internal realism is a topic of much recent discussion.[68] But there I had better

---

[64] Putnam does not cite a particular paper but presumably he is thinking of Dummett 1978a. The remarks about models having names from birth are highly reminiscent of Dummett 1978b, p. 191.

[65] Field (Auxier, Hahn, and Anderson 2016, Chapter 2) reads his view as close to deflationism about truth.

[66] Putnam 1983c, p. 25.

[67] Another contribution of Putnam's, highly relevant to these arguments but which unfortunately I do not have space to engage with in the depth it deserves, involves his rejection of Gödelian anti-mechanism. In Putnam 1975d he influentially responds to an argument of Nagel and Newman 1958 (an early version of later arguments made famous by Lucas and Penrose) claiming to show, from Gödel's first incompleteness theorem, that the human mind cannot be modeled by a Turing machine. Putnam's response is that this is a misapplication of the theorem: at best the theorem allows, for a given theory T, the formulation of a sentence, G, such that *if* T is consistent, G is true; but for all it says, the antecedent of that conditional cannot be discharged. Many, myself included, take Putnam's response to be decisive. See also Putnam 1994a for a later response to Penrose. For contemporary discussions of this cluster of issues see Horsten and Welch 2016.

[68] See e.g. Button and Walsh 2018, Warren and Waxman 2020, Button 2022, and Picollo and Waxman

stop.[69]

References

Auxier, Randall E., Lewis E. Hahn, and Douglas R. Anderson (2016). *The Philosophy of Hilary Putnam*. Chicago, IL, USA: Open Court.

Azzouni, Jody (2023). "What is PA + Con(PA) About, and Where?" *Synthese* 202.3, pp. 1–22.

Baghramian, Maria, ed. (2012). *Reading Putnam*. New York: Routledge.

Barrett, Thomas William (2020). "On Putnam's Proof of the Impossibility of a Nominalistic Physics". *Erkenntnis* 88.1, pp. 1–28.

Barton, Neil (2024). *Iterative Conceptions of Set*.

Bays, Timothy (2001). "On Putnam and His Models". *Journal of Philosophy* 98.7, p. 331.

Benacerraf, Paul (1965). "What Numbers Could Not Be". *Philosophical Review* 74.1, pp. 47–73.

Benacerraf, Paul and Hilary Putnam (1983). *Philosophy of Mathematics Selected Readings*. Cambridge University Press, 1983.

Berry, Sharon (2021). "Physical Possibility and Determinate Number Theory". *Philosophia Mathematica* 29.3, pp. 299–317.

— (2022). *A Logical Foundation for Potentialist Set Theory*. Cambridge University Press.

Bueno, Otávio (2005). "On the Referential Indeterminacy of Logical and Mathematical Concepts". *Journal of Philosophical Logic* 34.1, pp. 65–79.

Burgess, John P. (2018). "Putnam on Foundations: Models, Modals, Muddles". In: *Hilary Putnam on Logic and Mathematics*. Ed. by John Burgess. Springer Verlag, pp. 129–143.

Burgess, John P. and Gideon Rosen (1997). *A Subject with No Object: Strategies for Nominalistic Interpretation of Mathematics*. New York: Oxford University Press.

Button, Tim (2011). "The Metamathematics of Putnam?s Model-Theoretic Arguments". *Erkenntnis* 74.3, pp. 321–349.

— (2013). *The Limits of Realism*. Oxford: Oxford University Press UK.

— (2021). "Level Theory, Part 2: Axiomatizing the Bare Idea of a Potential Hierarchy". *Bulletin of Symbolic Logic* 27.4, pp. 461–484.

— (2022). "Mathematical Internal Realism". In: *Engaging Putnam*. Ed. by Sanjit Chakraborty and James Ferguson Conant. De Gruyter, pp. 157–182.

2023.

[69] I am extremely grateful to Sharon Berry, Sanjit Chakraborty, Tom Donaldson, Alex Lebrun, Lavinia Picollo, Zeynep Soysal, Jared Warren, and an anonymous referee for helpful comments and discussion which considerably improved the paper.

Button, Tim (2025). "The Iterative Conception of Function and the Iterative Conception of Set". In: *The Palgrave Companion to the Philosophy of Set Theory*. Ed. by Carolin Antos, Neil Barton, and Giorgio Venturi. Springer Nature Switzerland.

Button, Tim and Sean P. Walsh (2018). *Philosophy and Model Theory*. Ed. by Sean Walsh and Wilfrid Hodges. Oxford, UK: Oxford University Press.

Carnap, Rudolf (1937). *Logical Syntax of Language*. London, Routledge.

Chakraborty, Sanjit and James Ferguson Conant, eds. (2022). *Engaging Putnam*. Berlin, Germany: De Gruyter.

Clarke-Doane, Justin (2020). *Morality and Mathematics*. Oxford University Press.

Colyvan, Mark (2001). *The Indispensability of Mathematics*. Oxford, England: Oxford University Press.

Cook, Roy T. and Geoffrey Hellman, eds. (2018). *Hilary Putnam on Logic and Mathematics*. Springer Verlag.

Devitt, Michael (1983). "Realism and the Renegade Putnam: A Critical Study of Meaning and the Moral Sciences". *Noûs* 17.2, pp. 291–301.

Dummett, Michael (1978a). "The Philosophical Basis of Intuitionistic Logic". In: *Truth and other enigmas*. Ed. by Michael Dummett. Harvard University Press, pp. 215–247.

— (1978b). "The Philosophical Significance of Gödel's Theorem". In: *Truth and other enigmas*. Ed. by Michael Dummett. Harvard University Press, pp. 215–247.

Feferman, Solomon (1960). "Arithmetization of metamathematics in a general setting". *Fundamenta mathematicae* 49.1, pp. 35–92.

— (1992). "Why a Little Bit Goes a Long Way: Logical Foundations of Scientifically Applicable Mathematics". *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1992, pp. 442–455.

Field, Hartry (1989). *Realism, Mathematics & Modality*. New York, NY, USA: Blackwell.

— (2001). *Truth and the Absence of Fact*. New York: Oxford University Press.

— (2016). *Science Without Numbers*. Oxford University Press.

Floyd, Juliet and Hilary Putnam (2000). "A Note on Wittgenstein's 'Notorious Paragraph' About the Gödel Theorem". *Journal of Philosophy* 97.11, pp. 624–632.

Hamkins, Joel David and Øystein Linnebo (2022). "The Modal Logic of Set-Theoretic Potentialism and the Potentialist Maximality Principles". *Review of Symbolic Logic* 15.1, pp. 1–35.

Hellman, Geoffrey (1989). *Mathematics Without Numbers: Towards a Modal-Structural Interpretation*. Oxford, England: Oxford University Press.

Horsten, Leon and Philip Welch, eds. (2016). *Godel's Disjunction: The Scope and Limits of Mathematical Knowledge*. Oxford, England: Oxford University Press.

Kanamori, Akihiro (2018). “Putnam’s Constructivization Argument”. In: *Hilary Putnam on Logic and Mathematics*. Ed. by Roy T. Cook and Geoffrey Hellman. Springer Verlag, pp. 235–247.

Koellner, Peter (2009). “Truth in Mathematics: The Question of Pluralism”. In: *New Waves in Philosophy of Mathematics*. Ed. by Otávio Bueno and Oystein Linnebo. Palgrave Macmillian, pp. 80–116.

Laan, Boaz D (Apr. 2025). “How Should We Understand the Modal Potentialist’s Modality?” *Philosophia Mathematica* 33.2, pp. 242–252.

Lavine, Shaughan (Unpublished). *Skolem Was Wrong*.

Lebrun, Alex (manuscript). “Putnam on Mathematics and Ontological Commitment”.

Leng, Mary (2010). *Mathematics and Reality*. Oxford: Oxford University Press.

Lewis, David (1984). “Putnam?s Paradox”. *Australasian Journal of Philosophy* 62.3, pp. 221–236.

Liggins, David (2008). “Quine, Putnam, and the ?Quine?Putnam? Indispensability Argument”. *Erkenntnis* 68.1, pp. 113–127.

Linnebo, Øystein (2013). “The Potential Hierarchy of Sets”. *Review of Symbolic Logic* 6.2, pp. 205–228.

— (2018a). “Putnam on Mathematics as Modal Logic”. In: *Hilary Putnam on Logic and Mathematics*. Ed. by Geoffrey Hellman and Roy Cook. Springer Verlag, pp. 249–267.

— (2018b). *Thin Objects: An Abstractionist Account*. Oxford.

Maddy, Penelope (1988). “Believing the Axioms. I”. *Journal of Symbolic Logic* 53.2, pp. 481–511.

Maddy, Penelope and Jouko Väänänen (2023). *Philosophical Uses of Categoricity Arguments*. Elements in the Philosophy of Mathematics. Cambridge University Press.

McGee, Vann (1997). “How We Learn Mathematical Language”. *Philosophical Review* 106.1, pp. 35–68.

Melia, Joseph (2000). “Weaseling Away the Indispensability Argument”. *Mind* 109.435, pp. 455–480.

Nagel, Ernest and James Roy Newman (1958). *Godel’s Proof*. Ed. by James Roy Newman. New York, NY, USA.

Parsons, Charles (2007). *Mathematical Thought and its Objects*. New York: Cambridge University Press.

Picollo, Lavinia and Daniel Waxman (2023). “Internalism and the Determinacy of Mathematics”. *Mind* 132.528, pp. 1028–1052.

— (2025). “Arithmetical Pluralism and the Objectivity of Syntax”. *Noûs* 59.2, pp. 372–391.

Pincock, Christopher (2011). *Mathematics and Scientific Representation*. Oxford and New York: Oxford University Press USA.

Putnam, Hilary (1956). "Mathematics and the Existence of Abstract Entities". *Philosophical Studies* 7.6, pp. 81–88.

— (1968). "Is Logic Empirical?" *Boston Studies in the Philosophy of Science* 5.

— (1971). *Philosophy of Logic*. Ed. by Stephen Laurence and Cynthia Macdonald. New York, NY, USA: Routledge.

— (1975a). "An examination of Grunbaum's philosophy of geometry". In: *Philosophical Papers: Volume 1, Mathematics, Matter and Method*. New York: Cambridge University Press.

— (1975b). "It Ain't Necessarily So". In: *Philosophical Papers: Volume 1, Mathematics, Matter and Method*. New York: Cambridge University Press.

— (1975c). "Mathematics Without Foundations". In: *Philosophical Papers: Volume 1, Mathematics, Matter and Method*. New York: Cambridge University Press.

— (1975d). "Minds and Machines". In: *Philosophical Papers: Volume 2, Mind, Language and Reality*. New York: Cambridge University Press.

— (1975e). *Philosophical Papers: Volume 1, Mathematics, Matter and Method*. New York: Cambridge University Press.

— (1975f). "Philosophy of Physics". In: *Philosophical Papers: Volume 1, Mathematics, Matter and Method*. New York: Cambridge University Press.

— (1975g). "The Analytic and Synthetic". In: *Philosophical Papers: Volume 2, Mind, Language and Reality*. Ed. by Hilary Putnam. Cambridge University Press, pp. 33–69.

— (1975h). "The Thesis That Mathematics is Logic". In: *Philosophical Papers: Volume 1, Mathematics, Matter and Method*. New York: Cambridge University Press.

— (1975i). "Truth and Necessity in Mathematics". In: *Philosophical Papers: Volume 1, Mathematics, Matter and Method*. New York: Cambridge University Press.

— (1975j). "What is Mathematical Truth?" In: *Philosophical Papers: Volume 1, Mathematics, Matter and Method*. New York: Cambridge University Press.

— (1981). *Reason, Truth and History*. New York: Cambridge University Press.

— (1983a). "Analyticity and Apriority: Beyond Wittgenstein and Quine". In: *Philosophical Papers: Volume 3, Realism and Reason*. New York: Cambridge University Press.

— (1983b). "Equivalence". In: *Philosophical Papers: Volume 3, Realism and Reason*. New York: Cambridge University Press.

— (1983c). "Models and Reality". In: *Philosophical Papers: Volume 3, Realism and Reason*. New York: Cambridge University Press.

— (1983d). *Philosophical Papers: Volume 3, Realism and Reason*. New York: Cambridge University Press.

— (1987). *The Many Faces of Realism*. Open Court.

Putnam, Hilary (1994a). *The Best of All Possible Brains?* URL: `https://archive.nytimes.com/www.nytimes.com/books/97/04/27/nnp/17540.html?scp=32&sq=brain%20rules&st=cse`.

— (1994b). *Words and Life*. Ed. by James Conant. Cambridge: Harvard University Press.

— (2006). "Indispensability Arguments in the Philosophy of Mathematics".

— (2012). "Indispensability Arguments in the Philosophy of Mathematics". In: *Philosophy in an Age of Science: Physics, Mathematics, and Skepticism*.

— (2014). URL: `http://putnamphil.blogspot.com/2014/12/the-modal-logical-interpretation-and.html`.

Roberts, Sam (2019). "Modal Structuralism and Reflection". *Review of Symbolic Logic* 12.4, pp. 823–860.

Scambler, Chris (2021). "Can All Things Be Counted?" *Journal of Philosophical Logic* 50.5, pp. 1079–1106.

Shapiro, Stewart (2000). *Philosophy of Mathematics: Structure and Ontology*. Oxford University Press USA.

Sober, Elliott (1993). "Mathematics and Indispensability". *Philosophical Review* 102.1, pp. 35–57.

Soysal, Zeynep (2020). "Descriptivism About the Reference of Set-Theoretic Expressions: Revisiting Putnam?s Model-Theoretic Arguments". *The Monist* 103.4, pp. 442–454.

Studd, J. P. (2013). "The Iterative Conception of Set: A (Bi-)Modal Axiomatisation". *Journal of Philosophical Logic* 42.5, pp. 1–29.

Tarski, Alfred, Andrzej Mostowski, and Raphael M. Robinson (1953).

Warren, Jared (2015). "Quantifier Variance and the Collapse Argument". *Philosophical Quarterly* 65.259, pp. 241–253.

— (2017). "Quantifier Variance and Indefinite Extensibility". *Philosophical Review* 126.1, pp. 81–122.

— (2020). *Shadows of Syntax: Revitalizing Logical and Mathematical Conventionalism*. New York, USA.

Warren, Jared and Daniel Waxman (2020). "A Metasemantic Challenge for Mathematical Determinacy". *Synthese* 197.2, pp. 477–495.

Weston, Thomas (1976). "Kreisel, the Continuum Hypothesis and Second Order Set Theory". *Journal of Philosophical Logic* 5.2, pp. 281–298.

Wright, Crispin (2000). "Neo-Fregean Foundations for Real Analysis: Some Reflections on Frege's Constraint". *Notre Dame Journal of Formal Logic* 41.4, pp. 317–334.

Wright, Crispin and Bob Hale (2001). *The Reason's Proper Study: Essays Towards a Neo-Fregean Philosophy of Mathematics*. Oxford: Clarendon Press.