Steps Towards a Minimalist Account of Numbers

Thomas Schindler
University of Bristol, UK
thomas.schindler1980@gmail.com

This paper outlines an account of numbers based on the numerical equivalence schema (NES), which consists of all sentences of the form '$x.Fx = n$ if and only if $\exists^n x\, Fx$', where is the number-of operator and $\exists^n$ is defined in standard Russellian fashion. In the first part of the paper, I point out some analogies between the NES and the T-schema for truth. In light of these analogies, I formulate a minimalist account of numbers, based on the NES, which strongly parallels the minimalist (deflationary) account of truth. One may be tempted to develop the minimalist account in a fictionalist direction, according to which arithmetic is useful but untrue, if taken at face value. In the second part, I argue that this suggestion is not as attractive as it may first appear. The NES suffers from a similar problem to the T-schema: it is deductively weak and does not enable the derivation of any non-trivial generalizations. In the third part of the paper, I explore some strategies to deal with the generalization problem, again drawing inspiration from the literature on truth. In closing this paper, I briefly compare the minimalist to some other accounts of numbers.

1. Introduction

This paper outlines an account of numbers based on what I will call the *numerical equivalence schema* (or NES for short):

$$x.Fx = \mathbf{n} \leftrightarrow \exists^n x\, Fx$$

The schematic letter '**n**' is a place-holder for the numeral of the natural number *n*. I assume that numerals are formed via some canonical recursive procedure, for example, from '**0**' and the successor function symbol. The locution '$\exists^n x\, Fx$' is defined recursively in standard Russellian fashion:

$$\exists^0 x\, Fx =_{Df} \neg \exists x\, Fx$$

$$\exists^1 x\, Fx =_{Df} \exists x\, (Fx \wedge \forall y\, (Fy \rightarrow x = y))$$

$$\vdots$$

$$\exists^{n+1} x\, Fx =_{Df} \exists x\, (Fx \wedge \exists^n y\, (Fy \wedge y \neq x))$$

$$\vdots$$

Mind, Vol. 00 . 0 . 2021

doi:10.1093/mind/fzab060

I will call the expression '$\exists^n$' a numerically definite quantifier. The schematic letter 'F' may be replaced by any open formula in the language under consideration. (In the first three sections, we will deal with first-order languages only. In §4, we will consider second-order languages as well.) '' is a variable-binding operator which, applied to an open formula 'Fx', yields a singular term '$x.Fx$' in which 'x' is bound.

The NES is formulated in a regimented or scientific language in the sense of Quine (1957). For ease of communication, I occasionally write 'there are n Fs' instead of '$\exists^n x\, Fx$' and 'the number of Fs is $\mathbf{n}$' instead of '$x.Fx = \mathbf{n}$' and so on. Importantly, however, I shall not be concerned with linguistic analysis of the ordinary language expressions 'There are n Fs' or 'The number of Fs is n' but with theory construction.[1]

According to Quine's (1948) criterion of ontological commitment, the ontological commitments incurred by the left- and right-hand side of the NES differ: the occurrence of 'n' in '$\exists^n x\, Fx$' is syncategorematic and disappears upon analysis. The expression '$\mathbf{n}$' in '$x.Fx = \mathbf{n}$' is a singular term. It is written in boldface in order to remind us of that difference.

In §2 I point out some analogies between the NES and the T-schema for truth. In light of these analogies, I formulate a minimalist account of numbers, based on the NES, that strongly parallels the minimalist (deflationary) account of truth. One may be tempted to develop the minimalist account in a fictionalist direction, according to which arithmetic is useful but untrue, if taken at face value. In §3 I argue that this suggestion is not as attractive as it may first appear. The NES suffers from a problem similar to the T-schema: it is

[1] In particular, I don't claim that '$\exists^n x\, Fx$' reveals the 'true logical form' of 'there are n Fs'. Understood as linguistic analysis, this is arguably incorrect because it assigns different logical forms to different number words. Similarly, some have argued that the occurrence of 'nine' in the ordinary language expression 'the number of planets is nine' is not a singular term (Moltmann 2013), while others claim that it is, but denotes something other than a number (Snyder 2017). This view, if correct, would seem to undermine so-called easy arguments for the existence of numbers, that is, arguments to the effect that the existence of numbers can be inferred from seemingly innocuous observations (for example, that there are eight planets) via some rules that are analytic of our ordinary language concepts (for example, the number of planets is eight if and only if there are eight planets). From a Quinean perspective, easy arguments are not a good idea to begin with. Ordinary language is a messy thing and incurs too many commitments. When we are in the business of 'limning down the true and ultimate structure of reality', we need to appeal to the 'austere canonical notation for science' (Quine 1960, pp. 225, 221).

deductively weak, and does not enable the derivation of any non-trivial generalizations. In §4 I explore some strategies to deal with the generalization problem, again drawing inspiration from the literature on truth. Finally, in §5 I briefly compare the minimalist to some other accounts of numbers.

2. The applicability of arithmetic

Tennant (1997) takes the derivability of all instances of the NES as an adequacy condition on any theory of numbers, in the same way as the derivability of all instances of the T-schema is an adequacy condition on any theory of truth. That is, a theory of numbers cannot be content with characterizing the numbers purely in terms of their internal structure, but needs (in some way) to take their applicability into account. It is the NES that

> connects use of the term-forming operator 'the number of … s' with our use of quantifiers, negation and identity in our numerically non-committal, ordinary discourse about concrete things. (Tennant 1997, p. 317)

Similarly, Wright claims that the fact that Hume's Principle entails the principle called N^q—which is a slight variant of the NES[2]—is sufficient to enforce 'the interpretation of Fregean arithmetic as genuine arithmetic, and not merely a theory which can be interpreted as such'. According to Wright, this is so because 'any doubt on the point has to concern whether the definition of the arithmetical primitives which Frege offers … [is] adequate to the ordinary *applications* of arithmetic' (Wright 1999, p. 18, emphasis in original).

But how exactly does the NES account for the applicability of arithmetic? In order to see this, let us briefly consider a question raised by Dummett (1991, p. 99): how can we explain the fact that we can appeal to the equation '$\mathbf{5} + \mathbf{2} + \mathbf{0} = \mathbf{7}$' in order to justify that there were seven animals in the field from the fact that there were five sheep, two cows, and no other animals there?

Assume there were five sheep, two cows, and no other animals in the field (formulated in terms of numerically definite quantifiers). And assume that $\mathbf{5} + \mathbf{2} + \mathbf{0} = \mathbf{7}$. Given the NES, we infer from the first premiss that the number of sheep $= \mathbf{5}$, that the number of cows $=$

[2] N^q differs from the NES in that the numeral '**n**' is replaced by Frege's definition of that term; for instance, '**0**' is replaced by '$x.(x \neq x)$'.

2, and that the number of other animals $= \mathbf{0}$. Next, given suitable definitions—for example, that $x.Fx + x.Gx = x.(Fx \vee Gx)$, provided that the *F*s and *G*s are mutually disjoint—we conclude that the number of sheep, cows and other animals $= \mathbf{5} + \mathbf{2} + \mathbf{0}$. Now the second premiss—the arithmetical equation—implies that the number of sheep, cows and other animals $= \mathbf{7}$. Finally, another application of the NES yields the desired conclusion, namely, that there were seven animals in the field.

Thus the NES acts as a bridge principle: given some premisses involving numerically definite quantifiers, and given some body of (purely) arithmetical truths, the NES enables us to utilize the arithmetical truths to draw consequences involving numerically definite quantifiers.

To be sure, one can derive the claim that there were seven animals from the assumption that there were five sheep, two cows, and no other animals *without* appealing to the arithmetical equation. But appealing to the equation makes the derivation shorter and simpler, as has been observed by, for example, Putnam (1967) and Field (2016, ch. 2). Similarly, the introduction of a truth predicate into one's language gives rise to such 'speed-up' phenomena (Fischer 2014).

We can pursue this analogy between our truth talk and our number talk a bit further. An instance of the T-schema

(T) $\mathrm{T}(\ulcorner A \urcorner) \leftrightarrow A$

contains an expression, *A*, on both sides of the biconditional. On the right-hand side, *A* is *used*, whereas on the left-hand side, *A* is *mentioned*. On the right-hand side, the expression occurs in a syntactic position that our ordinary (first-order) quantifiers are incapable of generalizing (namely, that of a sentence). On the left-hand side, it occurs as part of a longer expression—a name of *A*—that our ordinary quantifiers are capable of generalizing. Given the equivalence between the left- and right-hand sides, we are able to quantify (indirectly) into a position that our ordinary quantifiers are incapable of—resulting in a massive increase of our expressive power.

Now compare (T) to the NES. An instance of the NES

$$x.Fx = \mathbf{n} \leftrightarrow \exists^n x\, Fx$$

contains a certain expression, a numeral, on both sides of the biconditional. The occurrence on the right-hand side is syncategorematic and not open for quantification. On the left-hand side, the numeral occurs in argument position and is therefore open for generalization. Given the equivalence between the left- and right-hand sides, we are able to quantify (indirectly) into a position that our ordinary quantifiers are incapable of—resulting in a massive increase of our expressive power.

That 'true' allows us to generalize sentence places is widely acknowledged. According to minimalists or deflationists about truth, there's actually nothing more to 'true' than that (a) the introduction of a truth predicate satisfying the T-schema enables us to quantify into positions of our language that our ordinary quantifiers are not capable of (namely, that of a sentence), and (b) all truth-theoretic facts—for instance, 'A conjunction is true if and only if both conjuncts are true'—can be explained on that basis (possibly plus some other explanatory principles that have nothing specifically to do with truth). This account of truth is minimal in so far as it claims that the T-schema (in conjunction with some non-truth-theoretic principles) is sufficient to explain all facts about truth. And it is deflationary in the sense that our truth talk fulfils a mere quasi-logical role. The truth predicate, in virtue of being a predicate, may pick out a property, but in applying the truth predicate to a sentence, our primary concern is not to attribute that property to the sentence, but to generalize on the position of that sentence.

Now, our comparison of the NES with the T-schema conjures a simple question: is it possible to develop an analogous minimalist or deflationary account of numbers based on the NES as well? What I have in mind is something along the following lines:

(α) The introduction of number terms satisfying the NES enables us to quantify into positions of our language that our ordinary quantifiers are not capable of (that is, the 'index' of a numerically definite quantifier).

(β) All arithmetical facts—for instance, '$\mathbf{5} + \mathbf{2} + \mathbf{0} = \mathbf{7}$', 'There are infinitely many primes'—can be explained on that basis (possibly together with some other explanatory principles that have nothing specifically to do with numbers).

This account of numbers is minimal in so far as it claims that the NES (in conjunction with some non-arithmetical principles) is sufficient to explain all facts about numbers. And it is deflationary in the sense that our number talk fulfils a mere quasi-logical role. The primary role of number talk, then, is not to make claims about certain abstract objects, even if they exist, but to enable us to generalize on positions of our language that our ordinary quantifiers are incapable of.

Frege (1903) famously said that 'applicability cannot be an accident' (*Grundgesetze* II, §89). Indeed, one reason why the minimalist account would be attractive is because it can provide us with an answer to the question, 'Can we give an account of arithmetic that does not make it depend for its truth on the way the [concrete] world is? And if so, what constrains the world to conform to arithmetic?' (Potter 2000, p. 1). Very roughly, the answer provided by the minimalist account is this: by (β), all truths of arithmetic can be explained on the basis of the NES. But the NES does not depend on the way the concrete world is. In particular, it doesn't depend on how many concrete objects there are, nor on what (natural) properties they exemplify. Therefore, the truths of arithmetic do not depend on the way the concrete world is. Moreover, since all truths of arithmetic can be explained on the basis of the NES, and the NES just is what connects our number talk with our discourse about the concrete world, it's not surprising that arithmetic is applicable to the concrete world. Admittedly, this explanation needs to be fleshed out, but I hope it is suggestive enough. At any rate, my aim in this paper is not to answer Potter's question, but to develop the minimalist account a little further.

3. The ontological commitments of deflationism

The idea of basing an account of numbers on some connection between numbers and numerically definite quantifiers is not entirely novel. Field (2016, ch. 2) hints at the idea. Hodes (1990) has developed a technically sophisticated account according to which numbers are fictions that encode certain cardinality quantifiers. Yablo (2005) has proposed an intriguing account of numbers, closely resembling the minimalist one, according to which numbers are representative aids that enable us to express infinite conjunctions and disjunctions of statements about the concrete world (very much like the truth predicate).

These accounts have been proposed in a fictionalist spirit, and it might be tempting to develop the minimalist account in a fictionalist direction as well. In order to clarify the minimalist account a bit further, and to distinguish it from these other accounts, I will explain now why I do not find the combination of minimalism and fictionalism very attractive. Now, the label 'fictionalism' is used in radically different ways. I will use it here for the view that arithmetic, despite being useful, is untrue if taken at face value. Drawing again on an analogy with the deflationary account of truth, this view can be motivated as follows.

Deflationists about truth like to point out that we would have no need for a truth predicate were we able to use a language with infinite conjunctions and disjunctions. For let $\mathcal{L}$ be some first-order language, let $A_1, A_2, \ldots$ be a list of all sentences of $\mathcal{L}$, and assume that there is a singular term $\ulcorner A_i \urcorner$ for each sentence A_i available in $\mathcal{L}$. Then we can define a truth predicate (for $\mathcal{L}$) by the following infinite disjunction:

$$(1) \quad \mathrm{T}(x) \equiv_{Df} \bigvee_i (x = \ulcorner A_i \urcorner \wedge A_i)$$

Given some extra assumptions to be specified shortly, (1) entails all instances of the T-schema for sentences of $\mathcal{L}$. We need not attribute too much theoretical significance to this observation: 'the cute point about infinite languages isn't really essential' Putnam (1978, p. 16).[3] But it serves rather well to *illustrate* the deflationary view: truth, although of a certain logical or mathematical complexity, is not 'an ordinary sort of property' (Horwich 1998, p. 2), but rather a logical or quasi-logical one.

One could express a similar view about numbers: were we able to use a language with infinite conjunctions and disjunctions, we would have no need for (singular) number terms either. Very roughly, the idea is that we can analyse '2 + 4 = 6', say, as the infinite conjunction

$$(2) \quad \bigwedge_{F,G} (\exists^2 x\, F(x) \wedge \exists^4 x\, G(x) \wedge \exists^0 x\, (F(x) \wedge G(x)) \rightarrow \exists^6 x\, (F(x) \vee G(x)))$$

This motivates a view according to which the worldly content of a purportedly arithmetical truth is just a logical truth, and numbers are only brought in as representational aids that allow us to 'express the infinitely many facts in finite compass' (Yablo 2005, p. 94), very much like the truth predicate:

> Just as truth is an essential aid in the expression of facts not about truth (there is no such property), perhaps numbers are an essential

[3] It is often said that the truth predicate allows us to express infinite conjunctions and disjunctions, but this is perhaps not the best way to spell out its logical function; see Halbach (1999) and Picollo and Schindler (2018a, 2018b).

aid in the expression of facts not about numbers (there are no such things). (Yablo 2005, p. 95)

Thus numbers are merely a convenient speed-up fiction: arithmetic is not literally true, but we can assign correctness conditions to arithmetical statements such that a statement of arithmetic (that is ordinarily taken to be true) is correct if and only if all conjuncts (one disjunct) in the corresponding infinite conjunction (disjunction) are (is) true. We can explain the utility of arithmetic without invoking the existence of numbers.[4]

There is a problem with Yablo's assignment of correctness conditions: it only works on the assumption that there are infinitely many objects (cf. Rayo 2008, p. 399). If there are only *n* objects, every statement of the form '$\mathbf{n} + \mathbf{m} = \mathbf{k}$' will be classified as correct, because the antecedent in each conjunct in 3 will be false and the whole conjunction vacuously true. But given that there are other ways to state the correctness conditions of arithmetical claims (Balaguer 2009), I don't consider this as a fatal objection to the fictionalist thesis.

We have motivated the fictionalist account of numbers by comparing it to the deflationary account of truth. It may seem now that the analogy between truth and numbers breaks down: truth can be eliminated in infinitary logic, while numbers can't. This, however, is not the case. Ironically, truth is not ontologically innocent, and cannot be eliminated in infinitary logic without appropriate existence assumptions. In order to see this, let us return to definition (1). Contrary to what is often insinuated (for example, Blackburn and Simmons 1999, pp. 13–14; Yablo 2005, p. 113 n. 16), instances of the T-schema do *not* follow logically from (1), but require suitable existence assumptions.

No doubt, the mere formulation of the T-schema already involves the singular terms $\ulcorner A_1 \urcorner$, $\ulcorner A_2 \urcorner$, . . ., and in a classical context all singular terms denote. But classical logic does not entail that there is more than one object, even in the presence of an infinite number of singular terms. It is logically possible that all these terms denote one and the same object. Interestingly, the T-schema is not conservative over logic, because it entails that there is more than one object. This was first observed by Halbach (2001b, p. 179). Consequently, definition (1) cannot entail all instances of the T-schema unless one adds suitable existence assumptions.

[4] I should note that Yablo himself claims to be an agnostic, rather than a nominalist about numbers (Yablo 2005, p. 111).

In order fully to appreciate Halbach's observation, we need to fill in some background. Axiomatic truth theories consist of axioms for truth formulated over a base theory that contains a sufficiently rich syntax theory that informs us about the properties of the specific objects that truth is ascribed to—say, sentences or codes thereof. Let $\mathcal{L}$ be a first-order language, the language of the base theory, and let $\mathcal{L}_T$ extend $\mathcal{L}$ with a monadic predicate T, for truth. Typically, $\mathcal{L}$ will contain all the vocabulary required to express certain syntactic properties of expressions of $\mathcal{L}_T$; in particular, $\mathcal{L}$ will contain a quote name $\ulcorner\sigma\urcorner$ for each expression σ of $\mathcal{L}_T$. Then a base theory is simply a recursively axiomatized theory formulated in $\mathcal{L}_T$ containing a syntax theory for $\mathcal{L}_T$. Typically, only logical or syntactic principles containing T are derivable in the base theory, but no truth-theoretic principles. An axiomatic truth theory is simply the result of adding truth-theoretic principles to the base theory—say, instances of the T-schema or Tarski-style compositional principles.

As mentioned above, the mere formulation of the T-schema already involves the singular terms $\ulcorner A_1\urcorner, \ulcorner A_2\urcorner, \ldots$, but we need some non-logical axioms to discriminate between the $\ulcorner A_i\urcorner$s. This is where the base theory comes into play. Typically, the latter will allow us to prove certain syntactic facts, for instance, that the result of concatenating $\ulcorner\neg\urcorner$ with $\ulcorner A\urcorner$ is $\ulcorner\neg A\urcorner$. Moreover, where σ, τ are distinct expressions, the base theory will normally allow us to prove that $\ulcorner\sigma\urcorner \neq \ulcorner\tau\urcorner$.

What Halbach observed was that the T-schema (restricted to sentences of the base language $\mathcal{L}$) *logically* entails the existence of at least two objects; that is, even in the absence of a base theory, the T-schema has existential consequences. Let A_i be a sentence of the base language. The T-schema implies $\mathrm{T}(\ulcorner A_i\urcorner) \leftrightarrow A_i$ and $\mathrm{T}(\ulcorner\neg A_i\urcorner) \leftrightarrow \neg A_i$. This entails in classical logic $\neg(\mathrm{T}(\ulcorner A_i\urcorner) \leftrightarrow \mathrm{T}(\ulcorner\neg A_i\urcorner))$, which by Leibniz's law implies that $\ulcorner A_i\urcorner \neq \ulcorner\neg A_i\urcorner$.[5]

Halbach's observation is interesting because it shows that although the T-schema is often taken to be analytic or definitional of truth, it has existential consequences, very much like Hume's Principle. (Of course, compared to Hume's Principle, the existential consequences of the T-schema are rather modest.) Moreover, since the T-schema

[5] In an interesting paper, Heylen and Horsten (2017) show that when the disquotational theory of truth is formulated over negative free logic, one can still obtain existential consequences.

entails that there is more than one object, definition (1) cannot entail the T-schema unless one adds suitable existence assumptions.[6]

We can significantly strengthen Halbach's observation. We have a need not only for a truth predicate, but also for a satisfaction predicate. For, just as the notion of truth answers a general need to generalize on sentence positions in our language, so the notion of satisfaction answers a general need to generalize on predicate places in our language (Parsons 1974). Interestingly, the existential consequences of a satisfaction predicate are much more substantial than that of a truth predicate. The satisfaction predicate, $\mathrm{Sat}(x, y)$, is governed by the following schema:

$$\text{(S)} \quad \forall x\,(\mathrm{Sat}(x, \ulcorner F_i \urcorner) \leftrightarrow F_i(x))$$

where $F_1, F_2, \ldots$ is an enumeration of all formulae with one free variable of the base language. As before, we can show that even without the base theory, the satisfaction schema has existential consequences. This time, however, we can demonstrate that schema (S) logically entails the existence of *infinitely* many objects.

To see this, let's single out a sequence of distinguished formulae $G_0, G_1, G_2, \ldots$ from our formulae $F_1, F_2, F_3, \ldots$ as follows:

$$G_0(x) \equiv_{Df} x \neq x$$
$$G_1(x) \equiv_{Df} x = \ulcorner G_0 \urcorner$$
$$G_{n+1}(x) \equiv_{Df} \bigvee_{i \leqslant n} x = \ulcorner G_i \urcorner$$

There is no object x such that $x \neq x$. On the other hand, there is (as a matter of logic) exactly one object x such that $x = \ulcorner G_0 \urcorner$. Thus (S) implies that $\ulcorner G_0 \urcorner$ does *not* satisfy the formula $\ulcorner G_0 \urcorner$ (because no object is non-self-identical) whereas there is exactly one object (namely, $\ulcorner G_0 \urcorner$) satisfying $\ulcorner G_1 \urcorner$. In other words, we have $\neg\,\mathrm{Sat}(\ulcorner G_0 \urcorner, \ulcorner G_0 \urcorner)$ and $\mathrm{Sat}(\ulcorner G_0 \urcorner, \ulcorner G_1 \urcorner)$. But this implies, by Leibniz's law, that $\ulcorner G_0 \urcorner \neq \ulcorner G_1 \urcorner$. Repeating this argument, we can show for all $n \neq m$ that $\ulcorner G_n \urcorner \neq \ulcorner G_m \urcorner$.

Although I won't go into any details, let me mention (for later reference) that the case of *uniform* disquotation is even more interesting. Uniform disquotation consists of all instances of the following schema

[6] Although the T-schema only entails the existence of at least two objects, in order to derive the T-schema from (1) one needs to presuppose much more than just the existence of two objects. The existence of two objects is a necessary, but not a sufficient condition.

$$\forall x_1 \ldots \forall x_n (\mathrm{T}(\ulcorner F(\dot{x}_1, \ldots, \dot{x}_n)\urcorner) \leftrightarrow F(x_1, \ldots, x_n))$$

where the dots above the variables on the left-hand side indicate the presence of certain substitution functions that (on the intended interpretation) replace the free variables by certain canonical names for some objects in our domain (see Halbach 2014 for details). If one picks all instances of uniform disquotation for formulae not involving the truth predicate (that is, formulae of $\mathcal{L}$), then the resulting disquotational theory relatively interprets the arithmetical theory commonly known as $\mathcal{L}$.[7] Although R is deductively rather weak, it is non-trivial, allowing for the numeral-wise representation of all primitive recursive functions, and is therefore essentially undecidable. If one chooses as legitimate instances of uniform disquotation all positive formulae (that is, formulae in which the truth predicate never occurs in the scope of an odd number of negation signs), then the resulting disquotational theory even interprets Robinson arithmetic Q. For a proof of these results, see Schindler (2018).

These observations show that the notions of truth, satisfaction and number cannot simply be eliminated in infinitary logic. However, given appropriate existence assumptions and appropriate logical resources, the relevant concepts can be defined and their laws deduced. Deflationism is not deflationary because it denies the existence of certain objects, say, truth-bearers or numbers. Deflationism is deflationary in virtue of claiming that talk about truth and numbers serves a mere quasi-logical role.

Let us now return to the question whether we should interpret the statements of arithmetic literally or not. Where $^\star A^\star$ is to be understood as 'it is to be imagined that A', Yablo formulates the NES (roughly) as follows: $^\star x.Fx = \mathbf{n}^\star$ if and only if $\exists^n x\, Fx$. But surely we cannot reformulate the T-schema as follows: $^\star\mathrm{T}(\ulcorner A \urcorner)^\star$ if and only if A. If truth ascriptions are not taken literally, then there is simply no reason why we shouldn't ascribe truth to statements of arithmetic. (And similarly for satisfaction.) However, taken literally, the T-schema and schema (S) logically entail the existence of infinitely many objects. Indeed, the fact that the existence of infinitely many objects is *logically* entailed is quite irrelevant. Truth ascriptions make little sense if the objects to which truth is ascribed are not already understood as objects of a particular kind, say, sentences or

[7] For a definition of the theories R and Q, see Monk (1976). For the notion of relative interpretation, see Tarski, Mostowdki and Robinson (1953).

propositions.[8] And it's only natural to assume that there are infinitely many of them, and that they are abstract.

A die-hard nominalist will of course insist that truth is attributed to utterances or other concrete objects, rather than propositions or sentences. I have little to say that could change the mind of someone of that persuasion. But I don't find this position very attractive: it doesn't lead to any decent theory of truth. As truth theorists, we typically want to say things like 'If x and y are sentences, then the conjunction of x and y is true if and only if x is true and y is true'. And it's quite cumbersome, if not impossible, to express this principle if reference to sentence types is not admitted.

I have argued that truth ascriptions and satisfaction claims need to be interpreted literally, and that the notions of truth and satisfaction presuppose or entail the existence of abstract objects. Consequently, we cannot be fictionalists about all abstract objects. This still leaves open the possibility that we be fictionalists about particular kinds of abstract objects, say, numbers. I have no real argument against fictionalism about numbers. But once we admit abstract objects of some sort or another into our domain of quantification, the rationale for a fictionalist treatment of arithmetic—namely, avoiding commitment to abstract objects—is considerably lessened.

Given that the notion of satisfaction already commits us to the existence of infinitely many syntactic objects, one may attempt to reduce numbers to syntactic objects instead.[9] This is especially appealing in light of the fact that the notion of satisfaction already enables us to interpret fragments of arithmetic. On the other hand, one may think that we have a much clearer grasp of numbers than of syntactic objects: it's no coincidence that axiomatic theories of truth typically use a theory of arithmetic as syntax theory. Thus, one may also attempt to reduce syntactic objects to numbers. Or perhaps we can take both of them as objects *sui generis*. I don't think deflationism should be combined with a 'heavyweight' realism about abstract objects either. This leaves essentially two options. One is to take abstract objects to be 'thin objects' roughly in the sense of Schiffer (2003) or Linnebo (2012). Another option is to distinguish between quantifier commitment and ontological commitment, and reject Quine's thesis that the

[8] See Horsten (2011, p. 82) for a similar remark.

[9] See Parsons (2008) for a revival of Hilbert's attempt to ground our knowledge of numbers in our knowledge of finite strings of strokes.

former entails the latter, as Azzouni (2004) does. But this is a topic for another paper. My intention here is merely to illustrate why I don't find the combination of minimalism and fictionalism very attractive.

4. From the NES to the theory of arithmetic

In this section, I will explore some options to argue in favour of the minimalist's claim (β), namely, that all arithmetical facts can be explained on the basis of the NES (possibly in conjunction with other explanatory principles that have nothing specifically to do with numbers). I don't consider any of the proposals that follow to be definitive, nor do I even consider them to provide a complete map of the possible options. My rather modest aim is simply to show that the prospects of arguing in favour of (β) are not as dim as one may initially think.

To begin with, let us first note that the NES logically implies that $\mathbf{m} \neq \mathbf{n}$ whenever '**m**' and '**n**' are distinct numerals. This follows from an argument similar to those considered in the previous section.

Next, it's possible to show that all true numerical equations and inequalities (involving addition and multiplication) can be derived from (a slightly more general version of) the NES in first-order logic, given suitable definitions. By numerical equations and inequalities, I mean (negated) identity statements between number terms containing no free or bound variables.

Let ^{2} be an operator that, applied to an open formula '$F(x, y)$', yields a singular term '${}^{2}xy.F(x,y)$' in which the variables 'x' and 'y' are bound.[10] Addition and multiplication can then be defined as follows:

$$x.Fx + x.Gx =_{Df} {}^{2}xy.((Fx \wedge y = 0) \vee (Gx \wedge y = 1))$$

(In words: $x.Fx + x.Gx$ is the number of pairs in the set $\{(x,0)|Fx\} \cup \{(x,1)|Gx\}$.)

$$x.Fx \cdot y.Gy =_{Df} {}^{2}xy.(Fx \wedge Gy)$$

(In words: $x.Fx \cdot y.Gy$ is the number of pairs in the set $\{(x,y)|Fx \wedge Gy\}$.)

[10] Alternatively, we could conceive of '' as a variadic operator, an operator that accepts a variable number of arguments. For our purposes, the choice is essentially immaterial, but I prefer the fixed-arity solution because variadic operators require some subtle changes in the underlying logic.

Since we formulated the NES in the introduction for formulae with one free variable only, we need to amend this detail. This is unproblematic: numerically definite quantifiers allow us to count not only the number of objects falling under some unary predicate, but also the number of pairs falling under a binary predicate, the number of triplets falling under some ternary predicate, and so on. Thus, where F is a formula with two free variables, we stipulate:

$$\exists^{0}(x,y)\,F(x,y) =_{Df} \neg\exists x \exists y\, F(x,y)$$
$$\exists^{1}(x,y)\,F(x,y) =_{Df} \exists x \exists y (F(x,y) \wedge \forall u \forall v (F(u,v) \rightarrow u = x \wedge v = y))$$
$$\exists^{2}(x,y)\,F(x,y) =_{Df} \exists x_1 \exists y_1 \exists x_2 \exists y_2 ((x_1 \neq x_2 \vee y_1 \neq y_2) \wedge F(x_1,y_1) \wedge F(x_2,y_2) \wedge \forall u \forall v (F(u,v) \rightarrow (u = x_1 \wedge v = y_1) \vee (u = x_2 \wedge v = y_2)))$$
$$\vdots$$

where we may read '$\exists^{n}(x,y)\,F(x,y)$' as 'There are (exactly) n pairs satisfying F'. (Obviously the definition doesn't require us to recognize pairs as an extra ontological category.) Then we can extend the NES as follows:

$$^{2}xy.F(x,y) = \mathbf{n} \leftrightarrow \exists^{n}(x,y)\,F(x,y)$$

(In words: the number of pairs satisfying F is $\mathbf{n}$ if and only if there are n pairs satisfying F.) Similarly, we can extend the NES to govern formulae with three free variables and more.

The above definitions merely entitle us to flank the addition and multiplication sign by terms of the form $x.Fx$. But we want to state equations and inequalities involving numerals. This can easily be achieved. For every $n > 0$, let $H_n(x)$ be shorthand for '$x = 1 \vee \ldots \vee x = \mathbf{n}$', and let $H_0(x)$ be shorthand for '$x \neq x$'. Then we stipulate that

$$\mathbf{n} + \mathbf{m} =_{Df} x.H_n(x) + x.H_m(x)$$

and similarly for multiplication.

Given these definitions, it's quite straightforward to prove the aforementioned result: *all true numerical equations and inequalities (involving addition and multiplication) can be derived from the NES in first-order logic.*

It's not too hard to see that we can derive more arithmetical truths from the NES once we expand our logical resources. For example, assume we avail ourselves of all *predicative* instances of the second-order comprehension axiom schema

$$\exists X \, \forall y \, (Xy \leftrightarrow \sigma)$$

where X is not free in σ, and σ must not contain any bound predicate variables (whence the label 'predicative'). Now we can define, say,

$$\mathbf{n} \leqslant \mathbf{m} \equiv_{Df} \exists X \, (\mathbf{n} + x.Xx = \mathbf{m})$$

and prove $\mathbf{n} \leqslant \mathbf{m}$ whenever that is true. For example, since we can prove $2 + 2 = 4$ and $2 = x.H_2(x)$, it follows that $2 \leqslant 4$. Given *impredicative* instances of second-order comprehension, it's also possible to derive some general truths, for example, that zero is not the successor of any number.[11]

However, most universal generalizations—even very elementary ones, such as that no number is its own successor—*cannot* be derived from the NES (even if one adds mathematical induction to the NES).[12]

The minimalist about truth finds herself in a similar predicament. The T-schema doesn't allow us to prove any non-trivial generalizations of the form 'For all sentences *x*, if *Fx* then *x* is true', but only their instances (Halbach 1999, Proposition 1).[13] Essentially, this is due to the compactness of first-order logic.[14] This leaves, I think, three options, all of which have been explored in the literature: (A) we

[11] As Heck (2000, p. 166) observes, the statement that zero is not the successor of any number follows from Frege's definition of successor and the statement that $x.Fx = 0 \leftrightarrow \exists^0 xFx$, which is an instance of the NES.

[12] In order to see this (in the first-order case, for simplicity), let M be a non-standard model of PA, and let $a \in M$ be some non-standard number. We construct a model N whose domain D is the interval $[0, a]$. For $x, y \in D$, let $S^N(x) = Min^M(S^M(x), a)$ and $x +^N y = Min^M(x +^M y, a)$ and $x \times^N y = Min^M(x \times^M y, a)$. Using induction in M, one can see that induction holds in N. However, injectivity of S does not hold in N. Thanks to Fedor Pakhomov for their help.

[13] More precisely, the following holds. Let B be some theory formulated in the base language and B^* the result of augmenting B with all instances of the T-schema (restricted to sentences of the base language). Let 'Fx' be a formula of the base language, and let M be a model of B such that infinitely many objects in M satisfy 'Fx'. (Assume furthermore that 'Fx' applies only to objects that are codes of sentences in the sense of M, and that M contains infinitely many objects that are not codes of sentences.) Then B^* does not prove that $\forall x \, (Fx \to Tx)$.

[14] *Proof sketch.* Let B, B^*, M and 'Fx' be as specified in the previous note 13 above. Assume that B contains an individual constant for every element in the domain of M and let c be a fresh constant. Let B^{**} be the result of augmenting B^* by the axioms 'Fc', '$\neg Tc$' and '$c \neq a$' for every constant a of B. It can be shown that every finite subset of B^{**} has a model, whence by

could appeal to a non-compact logic; (B) we could appeal to additional axioms; or (C) we could try to account for general facts in a non-deductive way.

Let me just briefly mention examples of types A and B. Horwich (1998, p. 137) once considered accounting for truth-theoretic generalizations by appealing to something like the ω-rule. This has provoked some criticism owing to its infinitary character: it is a rule that cannot actually be applied in practice. However, how successful this objection is will depend on what the appeal to the ω-rule is supposed to establish. If our goal is to give an account of how we actually arrive at generalizations, the account is obviously wrong. Thus Raatikainen (2005, p. 176) says that 'even if the rule would in theory entail the desired generalizations about truth, we human beings would never reach any of these generalizations. … But certainly we want ourselves to be able to reach, and in real life we do reach, such generalizations'. However, I find it hard to believe that this is what Horwich was proposing. I take it that his goal was not to give an account of how we actually reach such generalizations, but simply to show that they hold. Viewed in this way, it's not so clear that Horwich's appeal to the ω-rule misses its purpose, because the latter is obviously truth-preserving and does imply the desired generalizations.

Nevertheless, one might object that this is of little use. Hasn't the deflationist claimed that the disquotational theory of truth provides a complete account of truth? In order to address this objection, it is useful to distinguish—following Azzouni (2006, ch. 1.8)—between a *theory of 'true'* and a *theory of truth*, where the former is a (philosophical) theory about the word 'true' and its role in our language, and the latter an axiomatic theory characterizing the property of truth or extension of 'true'.[15] The deflationary account of truth is, first and foremost, a theory of 'true'. It maintains that 'true' is governed or implicitly defined by the T-schema, that the sole reason for having 'true' in our language is to fulfil a certain logical function, that it fulfils this function in virtue of obeying the T-schema, and that all other facts about truth can be explained on that basis (possibly in conjunction with other facts). This is entirely consistent with the claim that there are certain generalizations involving 'true' which aren't derivable from the T-schema but which we would like to include in a systematic

compactness B^{**} itself has a model, N. Since B^{**} contains B^{*}, N is also a model of B^{*}. But since '$\forall x\,(Fx \rightarrow Tx)$' cannot be true in N, it cannot be provable in B^{*}.

[15] A similar distinction can be found in Horwich (1998, p. 36).

body of knowledge about the property of truth. Of course, the inclusion of such generalizations into our axiomatic theory requires some justification, and given the minimalist's claim that all facts about truth can be explained on the basis of the T-schema (possibly in conjunction with other facts), the justification needs to be based on the T-schema (possibly in conjunction with other facts). But crucially, the generalizations don't need to follow logically from the T-schema.

Now, if the ω-rule can be justified in a way that is acceptable from the deflationist's point of view—for instance, on purely mathematical grounds—then it seems that we can justify the adoption of some generalizations. To avoid a potential misunderstanding: the suggestion is not that the T-schema plus the ω-rule constitute the deflationist's theory of truth; rather, the theory of truth will contain certain truth-theoretic generalizations as axioms, and their adoption is justified by the ω-rule.

I don't wish to imply, by the way, that the ω-rule can be justified from the point of view of the deflationist about truth (or that it can't).[16] I merely wish to say that once we take Azzouni's distinction into account, some of the objections that have been raised against the appeal to the ω-rule lose their force. Horwich claims that his theory of truth is made up exclusively of instances of the T-schema. Now *that* is really problematic; but minimalists need not follow Horwich on this point.

Instead of appealing to a non-compact logic, we could also invoke additional explanatory principles. Of course, these additional principles cannot be of a truth-theoretic character. In a recent paper, Horsten and Leigh (2017) argue that when we're justified in holding a theory, then we are *entitled* (in the sense of Tyler Burge) to adopt reflection principles for that theory.[17] Moreover, they show that the T-schema, when coupled with reflection principles, yields the desired generalizations (see Halbach 2001a for a predecessor of this result). Thus while the desired generalizations don't follow logically from the theory of 'true', we have—if Horsten and Leigh are right—epistemic warrant for incorporating them into our axiomatic theory of truth.

[16] For example, Cieśliński (2017, p. 73) has argued that the justification of the ω-rule relies on a notion of truth that is not available to the deflationist.

[17] One may wonder, however, whether reflection principles are of a truth-theoretic character or not. I argued elsewhere that they are not (Picollo and Schindler 2021).

Let us now return to the minimalist about numbers. As in the case of truth, it will be useful to draw a distinction akin to Azzouni's distinction between a theory of 'true' and a theory of truth. So let us distinguish between a *theory of numerals* and a *theory of arithmetic*, where the former is a (philosophical) theory about numerals and their role in our language, and the latter a systematic body of knowledge about the natural numbers, such as first-order Peano arithmetic (PA). The minimalist account of numbers is a theory of numerals rather than a theory of arithmetic. The minimalist's theory of arithmetic doesn't consist (merely) of the NES; but whatever propositions the minimalist chooses as axioms for their theory of arithmetic, they need to be justified on the basis of the NES, possibly in conjunction with other explanatory principles that have nothing specifically to do with numbers. Again, there are (at least) three options: (A) we could appeal to a non-compact logic; (B) we could appeal to additional axioms (which are not of a number-theoretic nature); or (C) we could try to account for arithmetic in a non-deductive way.

Let us start with strategies of type A. Obviously, we can obtain all arithmetical truths if we appeal to the ω-rule. Earlier I considered the possibility that the truth deflationist might appeal to this rule, provided it can be justified on non-truth-theoretic grounds, for example, based on purely mathematical considerations. But if the rule is justified on mathematical grounds, then it cannot be used in the present context, on pain of circularity.

A more promising idea is to appeal to the notion of second-order logical consequence. By the latter, I don't mean the consequence relation given by the so-called Henkin semantics, but the one given by the so-called standard semantics, where the second-order quantifiers are always interpreted as ranging over the full power set of the first-order domain (cf. Shapiro 1991). As is well known, there's no effective proof system that is both sound and complete with respect to this semantics. Although the standard proof system for second-order logic doesn't allow us to *derive* the axioms of PA from the NES (see note 12), the axioms of PA—indeed, of PA2—are semantically *entailed* by the NES in second-order logic (with standard semantics).

In order to see this, note that the NES entails (in second-order logic, with standard semantics) the principle known as 'Finite Hume's Principle' (FHP), as I'll show in a moment. FHP, in turn, entails all theorems of PA2—indeed, the theorems of PA2 can be derived from FHP in the standard proof system of second-order logic (cf. Heck 1997). FHP states that if either X or Y is finite (that is, has finitely many things falling under it, where finiteness is expressed by a standard second-order formula), then $x.Xx = x.Yx$ if and only if the Xs and Ys are equinumerous (that is, if the Xs and Ys can be put into a one-one correspondence, where this is expressed by a standard second-order formula).

The proof that the NES entails FHP isn't very difficult, but involves a certain formal subtlety that requires some care. Let $\Delta(x, y)$ be the primitive recursive function defined on pairs of numbers and formulae such that $\Delta(0, \ulcorner F \urcorner) = \ulcorner \exists^0 x\, Fx \urcorner$, $\Delta(1, \ulcorner F \urcorner) = \ulcorner \exists^1 x\, Fx \urcorner$, and so on. And let $\Theta(x, \ulcorner F \urcorner)$ be the primitive recursive function defined on pairs of numerals and formulae such that $\Theta(0, \ulcorner F \urcorner) = \ulcorner x.Fx = 0 \urcorner$, $\Theta(1, \ulcorner F \urcorner) = \ulcorner x.Fx = 1 \urcorner$, and so forth. Now let $M \vDash$ NES, and assume without loss of generality that $M \vDash$ 'X is finite' (relative to some variable assignment). Then (since we're working with the standard semantics) X really *is* finite, and consequently there is exactly one x such that $M \vDash \Delta(x, \ulcorner X \urcorner)$. Fix x. Since M is a model of the NES, $M \vDash \Theta(x, \ulcorner X \urcorner)$. Now, if $M \vDash x.Xx = x.Yx$, then $M \vDash \Theta(x, \ulcorner Y \urcorner)$, and therefore $M \vDash \Delta(x, \ulcorner Y \urcorner)$ (because $M \vDash$ NES). Hence (since we're working with the standard semantics) $M \vDash$ 'X, Y are equinumerous'. Conversely, if $M \vDash$ 'X, Y are equinumerous', then (since X is finite) there must be an x such that $M \vDash \Delta(x, \ulcorner X \urcorner) \wedge \Delta(x, \ulcorner Y \urcorner)$. Since M is a model of the NES, we have $M \vDash x.Xx = x.Yx$. Thus the NES entails FHP.

If second-order logic is logic, then the result shows that the NES logically implies all arithmetical truths. This is good news. I said that one of the reasons why the minimalist account might be interesting is that it can provide us with an answer to the question, 'Can we give an account of arithmetic that does not make it depend for its truth on the way the [concrete] world is?' (Potter 2000, p. 1). If the truths of arithmetic follow logically from the NES, and the NES accounts for the applicability of arithmetic, then that question seems to be answered in the affirmative.

But even if the preceding result offers a *semantic* route from the NES to PA2, one may doubt whether it also yields an *epistemic* route. The problem here, unsurprisingly, is that the consequence relation of second-order logic (with standard semantics) is not effective, so that, for example, the second-order consequences of the NES are not recursively enumerable.

Earlier I contemplated using the ω-rule—which isn't effective either—to justify some truth-theoretic generalizations. This is not really a problem, because I contemplated this not on the grounds that we actually use the ω-rule to derive certain truth-theoretic generalizations, but to establish that these generalizations follow from the theory consisting of the T-schema and the ω-rule—and *that* is of course established in the meta-theory. This should be acceptable provided we can justify both the ω-rule and the meta-theory by considerations that are not truth-theoretic in nature.

Now, our proof above that the NES second-order entails FHP was carried out in a meta-theory as well, and clearly involved some arithmetical reasoning. I am sceptical that one can justify the meta-theory by a strategy of type A itself, though it might be possible to account for it on the basis of strategies of type C, to which I will turn shortly.

Another possible approach is to appeal to logics that are stronger than first-order logic but weaker than full second-order logic with standard semantics. As an anonymous referee points out, one may exploit, for example, the literature on cardinality logics in order to formulate principles that naturally generalize the NES.[18] For instance, where n is a natural number, let 'C_n' be a third-order predicate that applies to a second-order predicate 'P' just in case there are n many things falling under 'P'. (Semantically, 'C_n' denotes the set of all subsets of the domain whose cardinality is n.) Now we may reformulate the NES as

$$x.Xx = \mathbf{n} \leftrightarrow C_n(X)$$

and allow 'C_n' to be replaced by a quantifiable third-order variable 'V' ranging over such cardinality properties (that is, 'V' ranges over the set $\{C_0, C_1, C_2, \ldots\}$). As Hodes (1988) has shown, enriching first-order logic with variables ranging over such cardinality properties (together with some further vocabulary that I won't discuss here) increases its expressive power. They do *not* increase the expressive power of dyadic second-order logic with standard semantics (Hodes 1988, Observation 2.2; Hodes 1990, §3). This is no problem, because the idea here is to abandon the standard semantics for the second-order quantifiers anyway. Instead, we might adopt Henkin semantics. It is conceivable that the above generalization of the NES—or something in its vicinity—entails more arithmetical principles than the NES entails in first-order logic, while being formulated in a logic that, arguably, is epistemically more tractable than full second-order

[18] Again, such generalizations might be justified by a strategy of type C.

logic with standard semantics. However, a rigorous investigation of this matter has to be left open for future research.

Let us now briefly look at an example of strategy B. For instance, we may consider the use of reflection principles, as we did in the case of truth. Suppose we are given a truth theory strong enough to prove a (uniform) reflection principle for the NES, that is,

$$\forall x_1 \ldots x_n\, (\mathrm{Prov}_{\mathrm{NES}}(\ulcorner F(\dot{x}_1, \ldots, \dot{x}_n)\urcorner) \rightarrow F(x_1, \ldots, x_n))$$

Given some base theory sufficiently strong to formalize standard meta-theoretic arguments, we can show that an NES proof of '$\mathbf{n} + 0 = \mathbf{n}$' can be transformed into an NES proof of '$(\mathbf{n} + 1) + 0 = \mathbf{n} + 1$'. By the induction axioms of the base theory, $\forall x\, \mathrm{Prov}_{\mathrm{NES}}(\ulcorner \dot{x} + 0 = \dot{x}\urcorner)$, whence by reflection $\forall x\, (x + 0 = x)$, and so on for the other Peano axioms minus induction. (The induction axioms of PA could perhaps be obtained as a consequence of some appropriately chosen definition of the natural numbers. Linnebo 2009, p. 233 makes a move of this sort in his account of the natural numbers.)

The main obstacle to this proposal is that the truth theory yields the desired generalizations only if the base theory is sufficiently strong to formalize certain meta-theoretic arguments. In particular, the base theory must 'know' that, for every $\mathbf{n}$, NES $\vdash \mathbf{n} + \pm 0 = \mathbf{n}$. This requires that we can represent all primitive recursive functions and have some induction available. Frequently, truth theorists use a fragment of arithmetic as their base or syntax theory. But an arithmetical base theory based on the NES is simply too weak for that purpose.

In §3, I mentioned that disquotational axioms for truth enable us to interpret a certain amount of arithmetic. For example, uniform disquotation for positive formulae gives us Robinson arithmetic. Of course, this is not enough to put the reflection principles to work either, because we need some induction, which Robinson arithmetic lacks. But it's conceivable that some other disquotational principles will yield stronger fragments of arithmetic.[19] Alternatively, we could use a non-arithmetical base theory based on our knowledge of syntactic objects. I am not unsympathetic to these suggestions, but it's obvious that more work needs to be done before we can form a proper judgement about the (philosophical) feasibility of these suggestions. Since all such theories have to interpret a theory of arithmetic stronger than the NES, one must wonder

[19] For some recent work on how to utilize the notion of truth to do justificatory work in mathematics (although in a somewhat different context), see Fischer, Horsten and Nicolai (2021).

if there is any epistemological gain. I think that there is, in particular if these strategies are combined with a solution of type C.

Finally, let us look at a solution of type C. One account that one might offer here involves what Russell calls the 'regressive method' (Russell 1907; see Potter 2000, pp. 157–60 and Irvine 1989 for further discussion). Russell was a logicist: he aimed to reduce mathematics to 'logic'. Such a reduction would yield an organization of our knowledge, minimize the possibility of error, and lead to new results. Crucially, however, Russell did not think of the reduction as providing an epistemic foundation for mathematics: it doesn't justify mathematics or put it on a more secure footing. Instead, the laws of logic and mathematics were to be justified by the regressive method: Russell assimilates the method of mathematics to that of the ordinary sciences of observation, claiming that

> we tend to believe the premises because we see that their consequences are true, instead of believing the consequences because we know the premises to be true. But the inferring of premises from consequences is the essence of induction; thus the method in investigating the principles of mathematics is really an inductive method, and is substantially the same as the method of discovering general laws in any other science. (Russell 1907, pp. 273–4)

According to this idea, we start out with some indubitable, elementary truths—for example, $2 + 2 = 4$—as our data, just as the ordinary sciences start with some observational statements. General laws are then obtained by inductive methods, and tested against these data, thereby conferring some degree of probability to the laws.

> [S]elf-evidence is never more than a part of the reason for accepting an axiom, and is never indispensable. The reason for accepting an axiom, as for accepting any other proposition, is always largely inductive, namely that many propositions which are nearly indubitable can be deduced from it, and that no equally plausible way is known by which these propositions could be true if the axiom were false, and nothing which is probably false can be deduced from it. (Russell 1910, p. 251)

Something like the regressive method has, in fact, become quite common in the justification of the axioms of higher set theory. For instance, Gödel recommends that

> [E]ven in the case [a new axiom] has no intrinsic necessity at all, a probable decision about its truth is possible also in another way, namely, inductively by studying its 'success'. Success here means fruitfulness in consequences, in particular in 'verifiable' consequences, i.e., consequences demonstrable without the new axiom, whose proofs with the help of the new axiom, however, are considerably simpler and easier to discover ... (Gödel 1947, p. 477)

There is a sense in which we can do even better than Russell. Russell cannot use the regressive method to explain why we know that $2 + 2 = 4$, because this is one of the data we use to confirm the axioms we adopt (cf. Potter 2000, p. 160). On the view presently under consideration, our knowledge of the data can be explained in terms of the NES, which provides a (quasi-logicist) foundation for quantifier-free elementary arithmetic.

The regressive method fits nicely with the distinction between a theory of numerals and a theory of arithmetic. The idea here is that our grasp of arithmetic comes in two steps. We start with the NES, which gives us some kind of Hilbertian, quantifier-free arithmetic. We then 'ask for the fewest and simplest logical premises from which it can be deduced' (Russell 1907, p. 275) and in the process of doing so, come to accept, say, Peano's axioms.

We can also use the regressive method to justify the meta-theory needed to carry out our earlier derivation of FHP from the NES, or the truth theory required for proposal B. We can also defend FHP directly by an abductive argument: we come to believe FHP because we recognize it as the best explanation of the NES. After all, a characteristic way of justifying or explaining a statement is to put it under general rules or principles. The NES has infinitely many instances, dealing with each natural number separately. FHP goes some way towards unifying and explaining these instances by bringing them under a general law.

The suggestion here is not that we justify either the Peano axioms or our meta-theory or FHP or our truth theory at the expense of the others. As Parsons (2008, §54) has argued, the traditional picture of the self-evidence of the axioms of arithmetic is in an important way misleading. The evident character of the axioms of arithmetic derives in large part from the *dense network* in which they are embedded. Part of their evident character comes from the consequences one can draw from them; another part comes from the fact that the axioms of arithmetic can be derived from other theories, say, a theory based on FHP or reflection principles for the NES.

This concludes my survey of some of the possible options for arguing in favour of the minimalist claim (β) that all arithmetical facts can be explained on the basis of the NES (possibly in conjunction with other explanatory principles). Much of what I have said is programmatic and requires further development, but I hope I have made good on the claim that the prospects for a minimalist account of numbers are not as dim as one may initially think.

5. Comparison with other approaches

To close this paper, I will briefly compare the minimalist account of numbers with some similar approaches in the extant literature, to further clarify and demarcate the account presented here.

First, one should note the similarities and differences between the minimalist account and the account proposed and rejected by Frege (1884) in *Grundlagen* §56. There, Frege considers introducing number terms via contextual definitions. So let us define 'the number 0 applies to a concept *F*' as 'there are no *F*s', 'the number 1 applies to a concept *F*' as 'there is exactly one *F*', and, more generally, 'the number $n+1$ applies to a concept *F*' as 'there is an object *a* which falls under *F* and such that the number *n* applies to the concept "falling under *F* but not identical with *a*"'. These definitions are contextual, as they assign meaning to a numeral, not on its own, but only in the context of phrases such as 'the number *n* applies to the concept *F*'. Frege objected to these definitions on the ground that they don't allow us to pick out numbers as self-subsistent objects, that is, to regard numerals as singular terms. In the dialectical context of the *Grundlagen*, Frege's argumentation has often been found wanting (Dummett 1991, §9). However, I won't go into this debate here, as my interest is not in an exegesis of Frege but rather in a comparison of this proposal with the minimalist one. Both approaches are similar in so far as both of them try to introduce number terms by reference to numerically definite quantifiers. But importantly, the minimalist approach treats numerals as singular terms from the very beginning. The NES doesn't provide a contextual definition of numerals in terms of statements involving numerically definite quantifiers. Rather, it introduces numerals as singular terms, but constrains their use by linking them to the use of numerically definite quantifiers. An immediate consequence of this is that we are able to derive all true numerical equations and inequalities. By contrast, the contextual definitions do not enable us to ground all true inequalities, unless (modal) principles ensuring the (potential) infinity of the domain of discourse are added (cf. Potter 2000, pp. 69–72, 116).

Next, some mention should be made of Jeffrey's account of arithmetic, which he dubs *logicism lite* (Jeffrey 2002). In contrast to more traditional forms of logicism, logicism lite doesn't regard arithmetical laws as abbreviations of logically true formulae. But just as the laws of physics are counted as empirical because they answer to empirical data, arithmetical laws should be counted as logical because they answer to logical data. Jeffrey associates numerical equations and inequalities with deduction schemata that don't involve any arithmetical vocabulary. The former are counted as true or false depending on whether the associated deduction schemata are valid or not. For instance, the inequality '$5 \neq 7$' is counted as true because the inference from the empty premiss set to the conclusion '$\exists^5 x\, Fx \leftrightarrow \exists^7 x\, Fx$' is logically invalid.[20] The truth or falsity of complex arithmetical sentences is then explained in terms of their components, treating universal and existential generalizations as infinite conjunctions and disjunctions, respectively. Of course, we typically 'lack the power to convene [the tribunal of the logical data] and hear its judgement' (Jeffrey 2002, p. 450). Instead, we accept a number-theoretic statement if we can find a proof of it from some suitable axioms.

Jeffrey's account is congenial to the minimalist one. However, the two accounts differ in scope. While Jeffrey provides a comprehensive account of the truth conditions of (purely) arithmetical statements, his remarks on the function of number talk, the ontological status of numbers, the applicability of arithmetic, and the justification of number-theoretic generalizations are rather brief. Moreover, while the NES is fundamental to the minimalist account, Jeffrey makes no mention of it. Following Tennant (1997), we have identified the derivability of all instances of the NES as an adequacy condition on any satisfactory theory of numbers. The minimalist account of numbers is minimal in the sense that its core principle—the NES—is the weakest principle satisfying Tennant's adequacy condition: every other theory satisfying it must entail the NES. The minimalist account of numbers,

[20] One interesting thing about Jeffrey's truth conditions is that they are, in some sense, purely logical. However, it should be noted that even on Jeffrey's account, the truth of inequalities such as '$5 \neq 7$' presupposes the existence of objects: the inference from the empty premiss set to the conclusion '$\exists^5 x\, Fx \leftrightarrow \exists^7 x\, Fx$' is logically invalid if and only if there is a model of suitable cardinality which makes the conclusion false.

like the minimalist account of truth, is of particular interest because it claims that all arithmetical facts (all facts about truth) can be explained on the mere basis of Tennant's (Tarski's) adequacy condition plus other explanatory principles that have nothing specifically to do with numbers (truth). The main contribution of the present paper is to spell out that account in more detail, and to argue that it is indeed a feasible one.

Finally, a brief remark on Hume's Principle (HP) is in order. The NES is derivable from HP, as Wright (1999) has shown. HP is, of course, the cornerstone of the neo-Fregean project of showing how we can refer to numbers and have knowledge of the number-theoretic truths (for example, Hale and Wright 2001; Cook 2009). Frege's Theorem tells us that we can derive the axioms of PA2 from HP given Frege's definition of *zero, predecession*, and *natural number*.

Given that HP implies the NES, and all theorems of PA2, are there any reasons to prefer the minimalist account over the neo-Fregean one? Let me just mention three points on which the minimalist account seems to fare better.

First, there is the epistemic worry raised by Boolos (1997, p. 304) concerning the 'quite strong content' that HP possesses. Frege arithmetic (HP plus second-order logic) is equi-consistent with PA2, a theory in which most mathematical arguments can be carried out. Since we cannot be sure that Frege arithmetic is consistent, how can we claim that HP is analytic or available without significant epistemological presupposition? Here, the deductive weakness of the NES turns into an advantage: the arithmetic content of the NES is so elementary and obvious that it's hard to doubt it. Moreover, if the minimalist's theory of arithmetic—say, PA—is justified by the regressive method, it's clear why there is a degree of uncertainty.

Second, there is the concern about the universal number, *anti-zero*, which—according to Frege arithmetic—would be a number greater than all other numbers. Can we be sure that a number with such a property exists (Boolos 1997, pp. 313–14)? By contrast, the minimalist account makes no claims about the universal number.

Third, there is the question about surplus content: HP involves our modern Cantorian concept of cardinality (according to which there are *transfinite* cardinal numbers) and 'the conceptual resources required if one is to so much as recognize the coherence of this concept (let alone HP's truth) vastly outstrip the conceptual resources employed in arithmetical reason' (Heck 1997, p. 246). Again, no such objection can be mounted against the NES. The latter doesn't involve

Cantor's concept of cardinality, and its deductive power is so weak that it certainly doesn't outstrip the conceptual resources employed in arithmetical reasoning.

Of course, the neo-Fregeans have responded to these objections (for example, Wright 1999), and I certainly don't wish to imply that they spell doom for the neo-Fregean project. In fact, I am quite sympathetic to many aspects of the neo-Fregean approach. Nevertheless, if it's possible to develop an account that makes do with logically weaker and conceptually simpler assumptions, then we have all the reasons we need to explore that possibility.[21]

References

Azzouni, Jody 2004: *Deflating Existential Consequence: A Case for Nominalism*. Oxford: Oxford University Press.

—— 2006: *Tracking Reason: Proof, Consequence, and Truth*. Oxford: Oxford University Press.

Balaguer, Mark 2009: 'Fictionalism, Theft, and the Story of Mathematics'. *Philosophia Mathematica*, 17(2), pp. 131–62.

Blackburn, Simon, and Keith Simmons 1999: 'Introduction'. In Simon Blackburn and Keith Simmons (eds.), *Truth*, pp. 1–28. Oxford: Oxford University Press.

Boolos, George 1997: 'Is Hume's Principle Analytic?' Reprinted in his *Logic, Logic, and Logic*, pp. 301–14. Cambridge, MA: Harvard University Press, 1999. First published in Richard G. Heck, Jr. (ed.), *Language, Thought, and Logic: Essays in Honour of Michael Dummett*, pp. 245–62. Oxford: Oxford University Press. Page references are to the reprint.

Bueno, Otávio, and Øystein Linnebo (eds.) 2009: *New Waves in Philosophy of Mathematics. Houndmills*: Palgrave Macmillan.

[21] For valuable feedback on earlier drafts of this paper I thank Bahram Assadian, Timo Beringer, Tim Button, Marta Campa, Volker Halbach, Luca Incurvati, Øystein Linnebo, Lavinia Picollo, Lorenzo Rossi, Julian Schlöder, Matteo Zicchetti, audiences in Bristol and Amsterdam, and various anonymous referees. An early version of this paper was written while the author received support from the European Commission (grant agreement no. 792202) within the project *The Logical Function of Property Talk* (LOFUPRO). The present, substantially revised and expanded version was written while the author received support from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 803684) within the project *Truth and Semantics* (TRUST).

Cieśliński, Cezary 2017: *The Epistemic Lightness of Truth: Deflationism and Its Logic*. Cambridge: Cambridge University Press.

Cook, Roy T. 2009: 'New Waves on an Old Beach: Fregean Philosophy of Mathematics Today'. In Bueno and Linnebo 2009, pp. 13–34.

Dummett, Michael 1991: *Frege: Philosophy of Mathematics*. London: Duckworth.

Field, Hartry 2016: *Science without Numbers: A Defense of Nominalism*, 2nd edn. Oxford: Oxford University Press. 1st edn. Princeton, NJ: Princeton University Press, 1980.

Fischer, Martin 2014: 'Truth and Speed-Up'. *Review of Symbolic Logic*, 7(2), pp. 319–40.

Fischer Martin, Leon Horsten, and Carlo Nicolai 2021: 'Hypatia's Silence: Truth, Justification, and Entitlement'. *Noûs*, 55(1), pp. 62–85.

Frege, Gottlob 1884: *Die Grundlagen der Arithmetik: Eine logisch-mathematische Untersuchung über den Begriff der Zahl. Breslau: Wilhelm Koebner*. Reprinted with an English translation by J. L. Austin as *The Foundations of Arithmetic: A Logico-Mathematical Enquiry into the Concept of Number*. Oxford: Basil Blackwell, 1950.

—— 1903: *Die Grundgesetze der Arithmetik*, *Volume* II. Jena: Hermann Pohle. Selections translated into English by P. T. Geach and Max Black in *Translations from the Philosophical Writings of Gottlob Frege*, pp. 139–224. Edited by Peter Geach and Max Black. Oxford: Basil Blackwell, 1952.

Gödel, Kurt 1947: 'What Is Cantor's Continuum Problem?' Reprinted with revisions in Paul Benacerraf and Hilary Putnam (eds.), *Philosophy of Mathematics: Selected Readings*, 2nd edn., pp. 470–85. Cambridge: Cambridge University Press, 1983. First published in *American Mathematical Monthly*, 54(9), pp. 515–25. Page references are to the reprint.

Halbach, Volker 1999: 'Disquotationalism and Infinite Conjunctions'. *Mind*, 108(429), pp. 1–22.

—— 2001a: 'Disquotational Truth and Analyticity'. *Journal of Symbolic Logic*, 66(4), pp. 1959–73.

—— 2001b: '*How Innocent Is Deflationism?*' *Synthese*, 126(1/2), pp. 167–94.

—— 2014: *Axiomatic Theories of Truth*, rev. edn. Cambridge: Cambridge University Press. 1st edn. 2011.

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

Heck, Richard Kimberly 1997: 'Finitude and Hume's Principle'. Reprinted in Heck 2011, pp. 237–66. First published under the name 'Richard G. Heck, Jr.' in *Journal of Philosophical Logic*, 26(6), pp. 589–618. Page references are to the reprint.

—— 2000: 'Cardinality, Counting, and Equinumerosity'. Reprinted in Heck 2011, pp. 156–79. First published under the name 'Richard G. Heck, Jr.' in *Notre Dame Journal of Formal Logic*, 41(3), pp. 187–209. Page references are to the reprint.

—— 2011: *Frege's Theorem*. Oxford: Oxford University Press. Published under the name 'Richard G. Heck, Jr.'

Heylen, Jan, and Leon Horsten 2017: 'Truth and Existence'. *Thought*, 6(2), pp. 106–14.

Hodes, Harold 1988: 'Cardinality Logics, Part I: Inclusions between Languages Based on "exactly"'. *Annals of Pure and Applied Logic*, 39(3), pp. 199–238.

—— 1990: 'Where Do the Natural Numbers Come From?' *Synthese*, 84(3), pp. 347–407.

Horsten, Leon 2011: *The Tarskian Turn: Deflationism and Axiomatic Truth. Cambridge*, MA: MIT Press.

Horsten, Leon and Graham E. Leigh 2017: 'Truth Is Simple'. *Mind*, 126(501), pp. 195–232.

Horwich, Paul 1998: *Truth*, 2nd edn. Oxford: Basil Blackwell. 1st edn. 1990.

Irvine, A. D. 1989: 'Epistemic Logicism and Russell's Regressive Method'. *Philosphical Studies*, 55(3), pp. 303–27.

Jeffrey, Richard 2002: 'Logicism Lite'. *Philosophy of Science*, 69(3), pp. 447–51.

Linnebo, Øystein 2009: 'The Individuation of the Natural Numbers'. In Bueno and Linnebo 2009, pp. 220–38.

—— 2012: 'Metaontological Minimalism'. *Philosophy Compass*, 7(2), pp. 139–51.

Moltmann, Friederike 2013: 'Reference to Numbers in Natural Language'. *Philosphical Studies*, 162(3), pp. 499–536.

Monk, J. Donald 1976: *Mathematical Logic*. New York: Springer-Verlag.

Parsons, Charles 1974: 'Sets and Classes'. *Noûs*, 8(10), pp. 1–12. Reprinted in his *Mathematics in Philosophy: Selected Essays*, pp. 209–20. Ithaca, NY: Cornell University Press, 1983.

—— 2008: *Mathematical Thought and Its Objects*. Cambridge: Cambridge University Press.
Picollo, Lavinia, and Thomas Schindler 2018a: 'Deflationism and the Function of Truth'. *Philosophical Perspectives*, 32: *Philosophy of Language*, pp. 326–51.
—— 2018b: 'Disquotation and Infinite Conjunctions'. *Erkenntnis*, 83, pp. 899–928.
—— 2021: 'Is Deflationism Compatible with Compositional and Tarskian Truth Theories?' In Carlo Nicolai and Johannes Stern (eds.), *Modes of Truth: The Unified Approach to Truth, Modality, and Paradox*, pp. 41–68. Abingdon and New York: Routledge.
Potter, Michael 2000: *Reason's Nearest Kin: Philosophies of Arithmetic from Kant to Carnap*. Oxford: Oxford University Press.
Putnam, Hilary 1967: 'The Thesis that Mathematics Is Logic'. Reprinted in his *Mathematics, Matter and Method: Philosophical Papers, Volume* 1, pp. 12–42. Cambridge: Cambridge University Press, 1975. First published in Ralph Schoenman (ed.), *Bertrand Russell: Philosopher of the Century. Essays in His Honour*, pp. 273–303. London: Allen and Unwin.
—— 1978: 'Meaning and Knowledge'. The John Locke Lectures, 1976. In his *Meaning and the Moral Sciences*, pp. 1–80. London: Routledge.
Quine, W. V. 1948: 'On What There Is'. *Review of Metaphysics*, 2(5), pp. 21–38. Reprinted in his *From a Logical Point of View*, pp. 1–19. Cambridge, MA: Harvard University Press, 1953.
—— 1957: 'The Scope and Language of Science'. *British Journal for the Philosophy of Science*, 8(29), pp. 1–17. Reprinted in his *The Ways of Paradox and Other Essays*, pp. 215–32. New York: Random House, 1966.
—— 1960: *Word and Object*. Cambridge, MA: MIT Press.
Raatikainen, Panu 2005: 'On Horwich's Way Out'. *Analysis*, 65(3), pp. 175–7.
Rayo, Agustín 2008: 'On Specifying Truth Conditions'. *Philosophical Review*, 117(3), pp. 385–443.
Russell, Bertrand 1907: 'The Regressive Method of Discovering the Premises of Mathematics'. Read before the Cambridge Mathematical Club, 9 March 1907. In Russell 1973, pp. 272–83.
—— 1910: 'The Theory of Logical Types'. In Russell 1973, pp. 215–52. First published as '*La Théorie des Types Logiques*' *in Revue de Métaphysique et de Morale*, 18, pp. 263–301.

—— 1973: *Essays in Analysis*. Edited by Douglas Lackey. London: Allen and Unwin.

Schiffer, Stephen 2003: *The Things We Mean*. Oxford: Oxford University Press.

Schindler, Thomas 2018: 'Some Notes on Truths and Comprehension'. *Journal of Philosophical Logic*, 47(3), pp. 449–79.

Shapiro, Stewart 1991: *Foundations without Foundationalism: A Case for Second-Order Logic*. Oxford: Clarendon Press.

Snyder, Eric 2017: 'Numbers and Cardinalities: What's Really Wrong with the Easy Argument for Numbers?' *Linguistics and Philosophy*, 40(4), pp. 373–400.

Tarski, Alfred, Andrzej Mostowski, and Raphael M. Robinson 1953: *Undecidable Theories*. Amsterdam: North-Holland.

Tennant, Neil 1997: 'On the Necessary Existence of Numbers'. *Noûs*, 31(3), pp. 307–36.

Wright, Crispin 1999: 'Is Hume's Principle Analytic?' *Notre Dame Journal of Formal Logic*, 40(1), pp. 6–30. Reprinted in Hale and Wright 2001, pp. 307–32.

Yablo, Stephen 2005: 'The Myth of the Seven'. In Mark Eli Kalderon (ed.), *Fictionalism in Metaphysics*, pp. 88–115. Oxford: Clarendon Press.

