Success by Default?[†]

Agustín Rayo[*]

I argue that Neo-Fregean accounts of arithmetical language and arithmetical knowledge tacitly rely on a thesis I call [*Success by Default*]—the thesis that, in the absence of reasons to the contrary, we are justified in thinking that certain stipulations are successful. Since Neo-Fregeans have yet to supply an adequate defense of [*Success by Default*], I conclude that there is an important gap in Neo-Fregean accounts of arithmetical language and knowledge. I end the paper by offering a naturalistic remedy.

1. Stipulation

We are in charge of our own linguistic behavior. Given a predicate and a set of sentences, we are free to stipulate that the predicate is to be used in such a way that each of the sentences turns out to be true. Of course, the stipulation might fail: there is, in general, no guarantee that it will have the effect of rendering its definiendum meaningful in such a way that every sentence in the set turns out to be true.

Consider, for example, the stipulation that the predicate '... is a zapp' is to be used in such a way that both 'Jones is a zapp' and 'Jones is not a zapp' turn out to be true.[1] No choice of extension for '... is a zapp' would allow it to satisfy the constraint imposed by our stipulation. The stipulation is therefore unsuccessful. We attempted to endow '... is a zapp' with meaning in such a way that both 'Jones is a zapp' and 'Jones is not a zapp' turn out to be true, and we failed.

Sometimes it is up to the world to determine whether a stipulation is successful. Suppose we stipulate that the predicate '... is a yapp' is to be used in such a way that both 'everyone in room 102 is a yapp' and 'my

[†] Many thanks to Bob Hale, Stewart Shapiro, Crispin Wright and two anonymous referees.

[*] Department of Logic and Metaphysics, University of St Andrews, St Andrews KY16 9AL, Scotland. ar29@st-andrews.ac.uk

[1] Here and throughout I assume that a stipulation of this kind does not have the effect of altering the logical form of the sentences involved, or the meaning of expressions other than its definienda.

©2003 PHILOSOPHIA MATHEMATICA (3) Vol. 11 (2003), pp. 305–322.

uncle Hector is not a yapp' turn out to be true. If the world does not cooperate, and my uncle Hector happens to be in room 102, then no choice of extension for 'yapp' can satisfy the constraint imposed by our stipulation. The stipulation is therefore unsuccessful. We attempted to endow '. . . is a yapp' with meaning in such a way that both 'everyone in room 102 is a yapp' and 'my uncle Hector is not a yapp' turn out to be true, and we failed.

For P a predicate and Ω a finite set of sentences, let $\langle P, \Omega \rangle$ be the stipulation that P is to be used in such a way that each of the sentences in Ω turns out to be true. We shall say that $\langle P, \Omega \rangle$ is *satisfiable* just in case some (empty or non-empty) extension[2] E is such that every sentence in Ω is true when E is the extension of P. Clearly, $\langle P, \Omega \rangle$ can only be satisfiable if Ω is logically consistent. But the converse is not true, since the world may impose constraints on satisfiability which go beyond logical consistency, as in the example about my uncle Hector. An especially interesting case of unsatisfiability without inconsistency arises for stipulations $\langle P, \Omega \rangle$ and $\langle P^*, \Omega^* \rangle$ when Ω and Ω^* are individually consistent but jointly incompatible (for instance, Ω might imply that the universe is finite and Ω^* that the universe is infinite). In such a scenario, at most one of the stipulations can be satisfiable. We will return to this issue below, when we discuss the so-called Bad Company Objection.

One final piece of notation: say that $\langle P, \Omega \rangle$ is a *successful* stipulation just in case it has the effect of rendering P meaningful in such a way that every sentence in Ω turns out to be true.

Suppose $\langle P, \Omega \rangle$ is a successful stipulation. It follows from the definition of success that every sentence in Ω is true. But the sentences in Ω cannot be true unless there is some (empty or non-empty) extension E such that every sentence in Ω is true when E is the extension of P.[3] So $\langle P, \Omega \rangle$ can only be a successful stipulation if it is satisfiable.

The conclusion that a stipulation can only be successful if it is satisfiable is not a claim about explanatory or epistemic priority. It does not imply, for example, that knowledge of the satisfiability of a stipulation must precede knowledge of its success. Our conclusion is simply that a stipulation must

[2] Talk of extensions is introduced for expository purposes only. Sentences of the form 'There is an extension E such that ψ (or every sentence in Ω) is true when E is the extension of P' are to be replaced by the Ramsey sentence corresponding to ψ (or the conjunction of sentences in Ω), that is, the result of substituting variables of the appropriate type for all occurrences of P in ψ (or the conjunction of sentences in Ω) and binding the new variables with an initial existential quantifier. When P is a first-level predicate, second-order quantification is required; when P is a second-level predicate, third-order quantification is required. The machinery described in Rayo and Yablo [2001] may be suitable for such purposes. In either case, I assume that higher-order quantifiers are allowed to take empty values.

[3] In other words, the sentences in Ω cannot be true unless the Ramsey sentence corresponding to their conjunction is. See footnote 2.

be either satisfiable or unsuccessful.

Nor does our conclusion presuppose any particular story about how it is that successful stipulations render their definienda meaningful. For instance, it is compatible with the view that stipulation is best understood on the model of reference fixing, so that $\langle P, \Omega \rangle$ is somehow tantamount to affixing the one and only 'possible meaning' satisfying the constraints imposed by Ω to P. But our conclusion is also compatible with the (more plausible) view that a stipulation is successful in virtue of its potential for establishing a linguistic *practice*, and that meaningfulness is constrained by practice.[4] All that is required to get our conclusion off the ground is the observation that a sentence S containing a meaningful predicate P cannot be true unless there is some (empty or non-empty) extension E such that the sentence is true when E is the extension of P.

Finally, the conclusion that a stipulation can only be successful if it is satisfiable does not rely on any particular story about the sorts of constraints that a stipulation must impose on the use of a predicate before it can be regarded as endowing the predicate with meaning. For instance, our conclusion does not hinge on the issue of whether ⟨'glub', 'Some apples are glub'⟩ suffices to render 'glub' meaningful. Nor does it hinge on any particular story about how it is that the following stipulations can be successful in the absence of martians,

⟨ 'xapp', 'something is a xapp iff it is a *blue* martian'⟩
⟨ 'wapp', 'something is a wapp iff it is a *green* martian'⟩,

or about how it is that, in the absence of martians, these stipulations might have the effect of rendering 'xapp' and 'wapp' semantically inequivalent. Regardless of how these questions are answered, it will remain the case that a sentence S containing a meaningful predicate P cannot be true unless there is some (empty or non-empty) extension E such that the sentence is true when E is the extension of P.

2. The Neo-Fregean Program

HP is the following sentence:

$$\forall X \forall Y (\mathbf{N}(X) = \mathbf{N}(Y) \leftrightarrow X \approx Y).$$ [5]

When '$\mathbf{N}(X)$' is read 'the number of the Xs', HP might be read:

[4] See Horwich [1997], [1998], chapter 6, and Hale and Wright [2000].

[5] We take '$X \approx Y$' to abbreviate a second-order formula expressing one-one correspondence between the objects falling under 'X' and the objects falling under 'Y'. So far we have only considered stipulations in which the definiendum is a predicate. In keeping with this, let us select an unused atomic predicate '$\mathbf{P}(X, y)$', and take '$\Phi(\mathbf{N}(X))$' to abbreviate '$\Phi(\iota y[\mathbf{P}(X, y)])$', where '$\iota x[\psi(x)]$' is a Russellian definite description (that is, '$\Phi(\iota x[\psi(x)])$' abbreviates '$\exists x \forall y((\psi(y) \leftrightarrow x = y) \wedge \Phi(x))$'). Nonetheless, for expositional purposes, we will proceed as if '$\mathbf{N}$' were atomic. For a defense of second-level predicates such as '$\mathbf{P}$' see Rayo [2002a].

> For arbitrary Xs and Ys, the number of the Xs is identical to the number of the Ys just in case the Xs are in one-one correspondence with the Ys.

Let us say that $\langle \mathbf{N}, \text{HP} \rangle$ is the stipulation that '**N**' is to be used in such a way that HP turns out to be true. Friends of the Neo-Fregean Program, championed by Bob Hale and Crispin Wright, subscribe to the following claim:

[*Success*]
$\langle \mathbf{N}, \text{HP} \rangle$ is a successful stipulation.

If they are able to justify [*Success*], Neo-Fregeans will be in a position to give an attractive account of arithmetical language, that is, an account of how it is that our arithmetical terminology might be rendered meaningful in such a way that the standard arithmetical axioms turn out to be true. For, if [*Success*] holds, then $\langle \mathbf{N}, \text{HP} \rangle$ renders '**N**' meaningful in such a way that HP is true. This is especially interesting in light of the fact that, when 'number' is interpreted in terms of '**N**'[6] the Dedekind-Peano Axioms are deducible in second-order logic from HP (together with suitable definitions).[7]

The Neo-Fregean case for [*Success*] is controversial. Although there are different reasons for concern, here we will focus on the fact that $\langle \mathbf{N}, \text{HP} \rangle$ is only successful if it is satisfiable, and that it is only satisfiable if there are infinitely many objects (by Frege's Theorem).[8]

Whether or not satisfiability is an issue depends on who Neo-Fregeans take to be their target audience. A *modest* version of the Neo-Fregean Program is directed towards someone who explicitly grants the existence of infinitely many objects. An *ambitious* version of the Neo-Fregean Program, on the other hand, is directed towards someone who does not explicitly grant that the universe is infinite. Thus, modest and ambitious Neo-Fregeans differ in their dialectical situation. Whereas modest Neo-Fregeans are in a dialectical situation which allows them to take the satisfiability of $\langle \mathbf{N}, \text{HP} \rangle$ for granted, ambitious Neo-Fregeans are not.

In the case of modest Neo-Fregeans, it is conceded immediately that $\langle \mathbf{N}, \text{HP} \rangle$ is satisfiable.[9] So, unless they are not entitled to [*Success*] for some other reason, they may carry on with the Neo-Fregean account of arithmetical language mentioned above. It is important to note that the

[6] This rests on the assumption that it is permissible to interpret 'number' in terms of '**N**'. But see Heck [1997a].

[7] The result is known as *Frege's Theorem*. It was originally proved by Frege (by making what was later shown to be a non-essential use of Basic Law V), and more recently rediscovered and cleaned up from contradiction by Crispin Wright. See Frege [1893/1903] and Wright [1983].

[8] In particular, we will ignore the so-called Caesar Problem. For discussion, see Hale and Wright [2001b].

[9] Assuming global choice, it is easy to show that HP is satisfiable in any infinite domain. (If the domain is set-sized, the standard axiom of choice will do.)

modest version of the Neo-Fregean Program is not trivial, since an account of arithmetical language does not follow immediately from the claim that mathematical objects exist.

Let us now turn to the case of ambitious Neo-Fregeans. A sentence that has been stipulated to be true is not, in general, guaranteed to be true.[10] For if the stipulation does not succeed, the sentence is either false or meaningless. Since ambitious Neo-Fregeans are not automatically granted the claim that $\langle \mathbf{N}, \text{HP} \rangle$ is satisfiable, they are not automatically granted the claim that $\langle \mathbf{N}, \text{HP} \rangle$ is successful and, hence, they are not automatically granted the claim that HP is true. They can only be permitted to proceed with their account of arithmetical language once an argument for [*Success*] has been supplied. And, since the ambitious Neo-Fregean's addressee does not explicitly grant the existence of infinitely many objects, the argument for [*Success*] must be *infinity-independent*, that is, it must not rely on premises that explicitly presuppose the existence of infinitely many objects.

To my knowledge, Neo-Fregeans have not provided such an argument. But they sometimes make it sound as if they think something like the following thesis is true:[11]

[*Success by Default*]

In the absence of reasons for thinking that a stipulation fails to meet the adequacy conditions in a certain set S, we are justified in thinking that it is successful.

Before saying more about S, and about the status of [*Success by Default*], it is important to note the following: if Neo-Fregeans manage to provide an argument for [*Success by Default*], and if they manage to show that there are no reasons for thinking that $\langle \mathbf{N}, \text{HP} \rangle$ fails to meet the adequacy conditions in S, then they will have provided an argument for the view that we are justified in thinking that $\langle \mathbf{N}, \text{HP} \rangle$ is successful, and will have done so without appealing to infinity assumptions. They can then go on to offer their account of arithmetical language.[12]

[*Success by Default*] is far from constituting an uncontroversial principle. It can only be accepted once the following subtasks have been carried out:

1. A specification of the conditions in S must be given.
2. The general view that there is such a thing as default justification must

[10] It is for expository purposes only that I speak of the stipulation that a sentence is to be true, instead of speaking of the stipulation that the definiendum is to be used in such a way that the sentence turns out to be true. Analogous expository simplifications are appealed to throughout the paper.

[11] In a talk entitled 'Implicit definition and abstraction' (The University of St Andrews, October 30, 1999), Stewart Shapiro discussed a principle similar to [*Success by Default*].

[12] An argument for the view that we are *justified* in thinking that $\langle \mathbf{N}, \text{HP} \rangle$ is successful is not the same as an argument for the view that $\langle \mathbf{N}, \text{HP} \rangle$ is successful, but it is presumably good enough.

be defended. In other words, a general argument must be provided to the effect that there are circumstances under which it is sufficient for a belief to be justified that there be no reasons to the contrary.

3. An argument must be provided to the effect that default justification is appropriate in the cases that are of interest to the Neo-Fregean.

Much discussion in the Neo-Fregean literature has centered upon subtask 1. The main adequacy condition in S is the requirement that the stipulation in question involve an abstraction principle, that is, a principle of the form

$$\forall\alpha\forall\beta(\Sigma(\alpha) = \Sigma(\beta) \leftrightarrow R(\alpha, \beta)),$$

for R an equivalence relation.[13] Unfortunately, this requirement will not do on its own. There are a number of abstraction principles giving rise to stipulations that Neo-Fregeans cannot regard as successful. This is what has come to be known as the *Bad Company Objection*. The simplest example of a problematic abstraction principle is Frege's Basic Law V, which is inconsistent and, hence, unsatisfiable. Neo-Fregeans must therefore add a requirement of consistency to S. But consistency cannot be the end of the matter. Heck [1992], Boolos [1997], and Weir [forthcoming], among others, have set forth an array of consistent but pairwise incompatible abstraction principles. Since [*Success by Default*] can only be taken to provide support for one or the other of a pair of consistent but incompatible abstraction principles, Neo-Fregeans are forced to postulate additional adequacy conditions. Recent efforts include conservativeness,[14] modesty,[15] stability,[16] and irenicity.[17] But Neo-Fregeans have yet to find a fully satisfactory set of adequacy conditions.

Subtask 2 is an issue that Wright and others have addressed in the literature on skepticism,[18] mostly in connection with perceptual beliefs. It

[13] Alternatively, the stipulation might involve a pair of *rules* leading from each side of an abstraction principle to the other.

[14] Roughly, an abstraction principle is conservative if it entails nothing new about objects other than the referents of terms in the principle's left-hand side. See Wright [1997] and Shapiro and Weir [1999].

[15] According to appendix 1 of Wright [1999], 'an abstraction [principle] is Modest if its addition to any theory with which it is consistent results in no consequences—whether proof- or model-theoretically established—for the ontology of the combined theory which cannot be justified by reference to its consequences for its own abstracts.' Here an abstract is the referent of a term in the left-hand side of the relevant abstraction principle.

[16] An abstraction principle is stable if, for some cardinal κ, it is true at all and only those cardinalities greater than or equal to κ. See Fine [1998], Shapiro and Weir [1999], and Weir [forthcoming].

[17] An abstraction principle is irenic if it is conservative and compatible with any conservative abstraction principle. See Weir [forthcoming].

[18] See Wright [2000a], Wright [2000b], Wright [forthcoming], and, for a related proposal, Pryor [2000].

is argued, for example, that, in the absence of reasons to the contrary, I am now justified in thinking that here is a hand and here is another.

There is, however, a gap with respect to subtask 3. As far as I know, this is an issue that has never been addressed in print, and there is no obvious way of exporting the conclusions about default justification in the case of perception to the realm of stipulations. The failure to address—or even explicitly consider—subtask 3 constitutes a very serious omission in the Neo-Fregean literature.

3. Conditional Stipulation

Neo-Fregeans are keen to point out that there is a certain sense in which $\langle \mathbf{N}, \text{HP} \rangle$ is a conditional stipulation. They argue that

> what is stipulated to be true is ... a (universally quantified) *biconditional*, so that what is done is to fix the truth-*conditions* for identities linking **N**-terms. The truth-*value* of instances of HP's left-hand side is never itself a matter of direct stipulation—if any identities of the form '$\mathbf{N}(F) = \mathbf{N}(G)$' are true, that is always the product of two factors: their truth-conditions, as given by the stipulation, together with the independently constituted and, in the best case, independently ascertainable truth of corresponding instances of HP's right-hand side. The existence of referents for **N**-terms is therefore never part of what is stipulated ...[19]

Thus, all that is required for $\langle \mathbf{N}, \text{HP} \rangle$ to succeed is the creation of an appropriate *concept*. What *extension* that concept turns out to have is a further issue, which depends on how matters stand in the world. It is not something immediately decided by the stipulation.

This invites a comparison between $\langle \mathbf{N}, \text{HP} \rangle$ and a stipulation such as

⟨ 'xapp', 'something is a xapp iff it is a blue martian'⟩.

For one might think that the latter succeeds in assigning a concept to 'xapp' independently of how matters stand in the world. If there are no blue martians, the concept will have an empty extension; if there are blue martians, the concept will have a non-empty extension. But the success of the stipulation is in no way dependent on whether or not there are any blue martians.

Similarly, it might be argued, $\langle \mathbf{N}, \text{HP} \rangle$ succeeds in assigning a concept to **N** independently of how matters stand in the world. But in this case it is facts about one-one correspondence, rather than facts about martians, that determine what shape the concept's extension must take.

It is not clear that insistence on the conditional character of $\langle \mathbf{N}, \text{HP} \rangle$ is of any help to Neo-Fregeans in the present context. For if there are only

[19] Hale and Wright [2000], pp. 315–316. I have substituted 'HP' for 'abstraction' and '**N**' for '§'.

finitely many objects, then the stipulation isn't satisfiable.[20] There is no extension E (empty or not) such that HP is true when E is the extension of $\mathbf{N}$.[21] And in the absence of such an extension there can be no concept satisfying the constraints imposed by HP.[22] By contrast, it is plausible to think that the stipulation pertaining to martians is satisfiable no matter what the circumstances might be, even though the delivered concept might turn out to have an empty extension.

Our earlier point remains. Since ambitious Neo-Fregeans have not been granted the claim that the universe is infinite, they cannot automatically be granted the claim that, as a result of setting forth $\langle \mathbf{N}, \mathrm{HP} \rangle$, HP is true. They can only be permitted to proceed with their account of arithmetical language once an argument for [*Success*] has been provided. A defense of [*Success by Default*] might yield the needed argument, but no such defense has been given.

4. The Abstraction Thesis

In addition to [*Success*], Neo-Fregeans subscribe to an intriguing claim.[23] It is based on the idea that semantic content may be 'conceptualized' or 'carved up' in different ways. They hold that, as a result of stipulating HP to be true, $\ulcorner \mathbf{N}(F) = \mathbf{N}(G) \urcorner$ comes to have the same content as $\ulcorner F \approx G \urcorner$, only 'carved up' in a different way. Call this the *Abstraction Thesis*.

An ambitious version of the Neo-Fregean program is targeted towards someone who does not explicitly grant that the universe is infinite. But, armed with the Abstraction Thesis, Neo-Fregeans might wish to argue that even an agent who has no *explicit* belief to the effect that the universe is infinite is *implicitly* committed to there being infinitely many objects. For, were she to 'reconceptualize' appropriate facts about one-one correspondence in a purportedly finite scenario, she would come to see that the scenario cannot be finite after all.

It should be noted, however, that the Abstraction Thesis is a *strengthening* of [*Success*], since it constitutes a special case of the view that, as

[20] Unless global choice holds, the situation might be worse still. I know of no proof to the effect that HP is satisfiable in any infinite domain which does not rely on some version of global choice (or standard choice if the domain of set-sized). (All of this assumes that the second-order quantifiers are allowed to take an 'empty' value. If they are not, then HP is satisfiable in a finite domain, but is of no use to the Neo-Fregean because Frege's Theorem fails. For discussion, see Shapiro and Weir [2000].)

[21] In other words, if there are finitely many objects, then the Ramsey sentence corresponding to HP is false.

[22] It is crucial that the concept satisfy the constraints imposed by HP. For it is of no use to the Neo-Fregean if $\langle \mathbf{N}, \mathrm{HP} \rangle$ delivers a concept unless HP is *true* when $\mathbf{N}$ expresses that concept.

[23] See Hale and Wright [2000], pp. 148–150, Wright [1997], pp. 207–209, Hale [1997] and [1999], and Wright [1983]. The original idea is due to Frege [1884], §64.

a result of setting forth $\langle \mathbf{N}, \mathrm{HP} \rangle$, HP is meaningful and true. So the Abstraction Thesis is in no less need of justification than [*Success*]. In the absence of further argumentation, it is therefore illegitimate to deploy the Abstraction Thesis at this stage in the dialectic.

In fact, commitment to the Abstraction Thesis can only exacerbate the problem of arguing for [*Success*]. It is already a problem for ambitious Neo-Fregeans to show that $\langle \mathbf{N}, \mathrm{HP} \rangle$ is a successful stipulation when all they expect it to deliver is a material equivalence. Matters can only be worse if they also expect it to deliver sameness of content.

5. Knowledge

In addition to their account of arithmetical language, Neo-Fregeans hope to provide an account of arithmetical knowledge. In particular, they might like to tell the following story about how an agent—let us call her Hero—could come to acquire a warrant for the belief that there are infinitely many objects:

The Warrant-Acquisition Story

Step A

As a result of stipulating HP to be true, Hero comes to have a warrant for the belief that HP is true.

Step B

On the basis of her warrant for the belief that HP is true, Hero may go through a proof of Frege's Theorem and thereby acquire a warrant for the belief that there are infinitely many objects.

Step A of the Warrant-Acquisition Story presupposes that $\langle \mathbf{N}, \mathrm{HP} \rangle$ is a successful stipulation. This means that ambitious Neo-Fregeans cannot be entitled to it until an infinity-independent argument for [*Success*] has been provided. Modest Neo-Fregeans do not face this problem because they are explicitly granted the existence of infinitely many objects. But even if step A is assumed to be unproblematic, we shall see that modest Neo-Fregeans are not entitled to step B unless they are able to show that Hero could acquire a special kind of warrant for the belief that HP is true: Hero's warrant for the belief that HP is true must be *infinity-independent*; in other words, it must not require an antecedent warrant for the belief that the universe is infinite.

I will argue that, in the absence of an infinity-independent warrant, step B of the Warrant-Acquisition Story begs the question: it involves a case of *warrant-transmission failure*. To see what I have in mind, consider the following example. Suppose Jones goes to the Zoo and, after observing an animal in a pen, forms the belief that the animal is a zebra. Suppose, moreover, that Jones's belief is warranted, even though her evidence does not, by itself, constitute a warrant for the belief that the animal is not a cleverly disguised mule.[24] Finally, suppose Jones carries out the following argument:[25]

The Zebra Argument

The animal in the pen is a zebra.

If the animal in the pen is a zebra, then it is not a cleverly disguised mule.

Therefore, the animal in the pen is not a cleverly disguised mule.

Jones's evidence does not, by itself, constitute a warrant for the belief that the animal is not a cleverly disguised mule. And it seems clear that going through the Zebra Argument shouldn't help, even though the argument is valid and Jones is warranted in believing its premises. Merely by deploying the argument, she should not be able to *acquire* a warrant for the argument's conclusion. Whenever an agent faces such a predicament we will say that there is a case of warrant-transmission failure. (It is worth noting that what is at issue here is not *closure*. We are not concerned with the issue of whether knowledge is closed under known implication.)

Accounts of warrant-transmission failure have been recently offered by Crispin Wright and Martin Davies,[26] and criticized by Jim Pryor.[27] Fortunately, the point which will be made here is quite general. It does not hinge on the details of Wright's and Davies's proposals, nor on the effect of Pryor's criticism. For present purposes all that is required is the observation that the following is a *sufficient* condition for warrant-transmission failure:

[*Transmission*]

Epistemic warrant cannot be transmitted from the premisses of a valid argument to its conclusion if, for one of the premises, the warrant for that premise requires an antecedent warrant for the conclusion.

Let us return to the Neo-Fregean's Warrant-Acquisition Story. Step B is the claim that, on the basis of her warrant for believing that HP is true, Hero may go through a proof of Frege's Theorem and thereby acquire a warrant for believing that there are infinitely many objects. Suppose that Hero carries out the following argument:

The Neo-Fregean Argument

HP is true.

[24] On a relevant-alternatives view, this may be the case if the evidence does not suffice to *rule out* the possibility that the animal is a cleverly disguised mule. For recent expositions of this sort of view, see De Rose [1995] and Lewis [1996]

[25] The *Zebra Argument* first appeared in Dretske [1970], but it was not originally set forth as an instance of warrant-transmission failure. The observation that it can be used to exemplify warrant-transmission failure is due, I believe, to Crispin Wright.

[26] See, for instance, Davies [2000] and Wright [2000a].

[27] See 'Is Moore's argument an example of transmission-failure?'. An online draft is available at http://www.people.fas.harvard.edu/~jpryor/papers/.

If HP is true, then there are infinitely many objects (by Frege's Theorem).

Therefore, there are infinitely many objects.

What about warrant-transmission failure? Could Hero come to acquire a warrant for believing in the infinity of the universe on the basis of this argument? Not unless her warrant for the belief that HP is true is infinity-independent. If Hero's warrant for the belief that HP is true requires an antecedent warrant for the belief that the universe is infinite, then [*Transmission*] diagnoses warrant-transmission failure, and modest Neo-Fregeans are not entitled to step B of the Warrant-Acquisition Story.

Modest Neo-Fregeans are therefore only in a position to defend the Warrant-Acquisition Story if they are able to argue that Hero could acquire an infinity-independent warrant for the belief that HP is true. To my knowledge, Neo-Fregeans have not provided such an argument, although they sometimes make it sound as though they think that something along the lines of [*Success by Default*] would do the job.

We have seen that, in the absence of an infinity-independent argument for [*Success*], ambitious Neo-Fregeans are not in a position to defend their account of arithmetical language. We have also seen that, in the absence of an infinity-independent warrant for the belief that HP is true, modest Neo-Fregeans are not in a position to defend the Warrant-Acquisition Story. For even if step A is assumed to be unproblematic, step B is guilty of a warrant-transmission failure. The moral is clear. Neo-Fregeans must either mount a defense of [*Success by Default*], or find an alternative argument for the needed conclusions.

6. Boolos

In discussing the Neo-Fregean program, George Boolos often set forth the following complaint:

> Properly understood, [HP] only *seems* to be analytic; seeing how it can be put to work reveals it as synthetic if either analytic or synthetic.[28]

In particular, Boolos thought that HP is revealed as synthetic (if analytic or synthetic) when one sees that it implies the existence of infinitely many objects.[29] Call this the *Boolos Complaint*. In this section we will see how the Boolos Complaint relates to the arguments we offered in the preceding section.

Boolos [1990b] contains an early formulation of the Boolos Complaint.[30] But its clearest manifestation occurs in Boolos [1997], where after conceding

[28] Boolos [1993], p. 233. My emphasis. See also Boolos [1990b], p. 214, and Boolos [1997]. Page references are to Boolos [1998].

[29] See Boolos [1993], p. 233. Boolos also mentions the 'surmountable' difficulty that HP implies a number corresponding to all the objects there are.

[30] See p. 212.

that 'at first glance, HP might certainly seem analytic',[31] Boolos argues as follows:

> One person's *tollens* is another's *ponens*, and Wright happily regards the existence of infinitely many objects, and indeed, that of a Dedekind infinite concept, as analytic, since they are logical consequences of what he takes to be an analytic truth [namely, HP]. He would also regard the existential quantification of HP (over the positions occupied by '**N**') as analytic. But what guarantee have we that there is such a function from concepts to objects as HP and its existential quantification claim there to be? ... If there is such a function then it is quite reasonable to think that whichever function '**N**' denotes, it maps non-equinumerous concepts to different objects and equinumerous ones to the same object, and this moreover because of the meaning of '**N**', the number-of sign or the phrase 'the number of'. But do we have any analytic guarantee that there is a function that works in the appropriate manner?[32]

It is clear from context that Boolos thought the final question should be answered in the negative.

In addition, Boolos suggested that 'there may be some analytic truths in the vicinity of HP with which it is being confused'.[33] In particular, he suggested 'very tentatively and playing along' that the conditional whose consequent is HP and whose antecedent is HP's existential quantification might be regarded as analytic.[34] This, he thought, might help explain why HP seems analytic, even though it is not.

Taking all of this on board, it seems to me that the Boolos Complaint can be reasonably reconstructed as follows. Boolos was willing to grant that our grasp of the meaning of '**N**' is enough by itself to supply us with a warrant for believing the conditional whose consequent is HP and whose antecedent is HP's existential quantification. But he did not think that our grasp of the meaning of '**N**' is enough, by itself, to supply us with a warrant for believing HP. In order for our grasp of the meaning of '**N**' to supply us with a warrant for believing HP, we need an antecedent warrant for the belief that HP's existential quantification is true. One way of acquiring an antecedent warrant for the belief that HP's existential quantification is true is by acquiring a warrant for the belief that the universe is infinite, and showing that HP's existential quantification follows from the infinity of the universe (given a suitable choice principle). If other ways of acquiring an antecedent warrant for the belief that HP's existential quantification is true are set aside, then it follows that, in order for our grasp of the meaning

[31] Boolos [1997], p. 303.

[32] See Boolos [1997], p. 306. Here and in what follows I have substituted ''**N**'' for Boolos's 'octothorpe'.

[33] P. 306.

[34] P. 306.

of '**N**' to supply us with a warrant for believing HP, we need an antecedent warrant for believing that the universe is infinite. In other words, our grasp of the meaning of '**N**' does not supply us with an *infinity-independent* warrant for believing that HP is true.

If correct, the Boolos Complaint leads immediately to a conclusion about warrant-transmission failure. For suppose that, as a result of grasping the meaning of '**N**', we come to acquire a warrant for the belief that HP is true. If the Boolos Complaint is along the right lines, then our warrant for the belief that HP is true required an antecedent warrant for the belief that the universe is infinite. So, by [*Transmission*], we cannot use the Neo-Fregean Argument to transmit warrant from our belief that HP is true to the conclusion that the universe is infinite.

Assessing the merits of the Boolos Complaint (and the ensuing Neo-Fregean reaction)[35] is a delicate matter, and would take us too far afield. For present purposes it suffices to distinguish between the Boolos Complaint and the argument set forth in the preceding section. In the preceding section we argued for the view that, if modest Neo-Fregeans are to be in a position to defend the Warrant-Acquisition Story, they need an infinity-independent warrant for the belief that HP is true. The Boolos complaint, on the other hand, is the view that a certain kind of warrant is not infinity-independent, namely, the kind of warrant supplied by our grasp of the meaning of '**N**'.

A thesis like [*Success by Default*] might eventually give Neo-Fregeans a powerful tool for addressing the sort of challenge posed by the Boolos Complaint. But not as things stand. Not enough has been done to justify the claim that [*Success by Default*] is true.

7. An Alternative

So far, we have argued that the Neo-Fregean case for [*Success*] relies on [*Success by Default*], and that [*Success by Default*] has not received proper justification. Fortunately, there is a different way of defending [*Success*]. Instead of relying on [*Success by Default*], one might provide an independent argument for the infinity of the universe.

The purpose of this final section is to provide such an argument. The idea is as follows. Independently of our account of arithmetical language—our story about *how* it is that arithmetical vocabulary might be rendered meaningful in such a way that the standard arithmetical axioms turn out to be true—there are good reasons for thinking that the standard arithmetical axioms *are* true. But it follows from the truth of arithmetic that the universe is infinite. So, independently of our account of arithmetical language, there are good reasons for thinking that the universe is infinite.

To defend the claim that we have good reasons for believing in the truth of arithmetic, I rely on the following epistemological principle:

[35] See, for instance, Hale and Wright [2000].

[*Nat*]

We are justified in believing an overall theory of the world to the extent that it is at least as successful as any rival by the lights of our current scientific standards (including criteria such as empirical adequacy, simplicity and strength),

and argue as follows:

The Naturalistic Argument

Arithmetic is part of our best theory of the world (by the lights of current scientific standards).

By [*Nat*], we are justified in believing that arithmetic is true insofar as it is part of our best overall theory of the world.

Hence, we are justified in believing that arithmetic is true.[36]

The Naturalistic Argument delivers the desired conclusion, but it might give rise to various sources of concern. I will try to address them in turn.

1. *Indispensability*

   Is the Naturalistic Argument an indispensability argument?

   No. By [*Nat*], we are justified in believing that arithmetic is true insofar as arithmetic is a *part* of our best theory of the world, whether or not it is an indispensable part. This is important because there are reasons for doubting that mathematics is indispensable for the development of the natural sciences.[37]

   A mathematical theory might be part of our best overall theory of the world even if there is a nominalistic alternative, as long as the alternative is less successful by the lights of our current scientific standards (for instance, the nominalistic alternative might result in a loss of simplicity).

2. *Arithmetic as a part of our best overall theory of the world*

   How do we know that arithmetic is part of our best overall theory of the world?

   Penelope Maddy has recently set forth a defense of mathematical naturalism—the view that the methods of mathematicians should not be subject to non-mathematical criticism.[38] If mathematical naturalism is true, then all it takes for arithmetic to be part of our best overall theory of the world is that it be sanctioned by the methods implicit in the practice of mathematicians. And, of course, arithmetic is sanctioned by the methods of mathematicians if anything is.

   It is important to note, however, that the Naturalistic Argument does not presuppose mathematical naturalism. All we need is the assumption

[36] More precisely, what the argument delivers is the view that we are justified in accepting an overall theory of the world of which arithmetic is a part. In what follows I ignore this distinction for the sake of simplicity.

[37] See Field [1980] and [1982]. See also Rayo [2002b].

[38] See Maddy [1997].

that our best theories of the *natural* world are part of our best overall theory of the world. With this assumption in place, the widespread use of mathematical machinery within the natural sciences gives us good (though defeasible) evidence for the view that arithmetic is part of our best overall theory of the world. (The evidence is defeasible because it is possible—though implausible—that practicing scientists do not fully live up to the standards implicit in their practice when they employ mathematical machinery.)

3. *An objection to* [*Nat*]
   The inference from the claim that an overall theory of the world is at least as successful as any rival by the lights of our current scientific standards to the claim that we are justified in taking the theory to be true is non-trivial. It is open to the objection that no reason has been given to suppose that we are lucky enough to live in a world in which reality plays by the rules imposed by our current scientific standards. Why should we suppose, for example, that simplicity is a guide to truth?
   This is an objection that must be taken seriously. But it is a general problem in epistemology. For it threatens the inference from the claim that an arbitrary theory, mathematical or not, meets our current scientific standards to the claim that we are justified in taking the theory to be true. In particular, it threatens the inference—accepted by most philosophers—from the claim that a *physical* theory meets our current scientific standards to the claim that we are justified in taking the theory to be true.[39] Insofar as it leads to general skepticism about our best scientific theories, we should sooner be wary of the objection than give up [*Nat*].
4. *Anti-Realism*
   Consider the result of substituting 'accepting' for 'believing' in [*Nat*]:
   [*Nat**]
   > We are justified in *accepting* an overall theory of the world to the extent that the theory is at least as successful as any rival by the lights of our current scientific standards.

   If accepting a theory and believing it to be true come to the same thing, then [*Nat*] and [*Nat**] come to the same thing. But someone uncomfortable with scientific realism may wish to argue that the acceptance of a scientific theory requires *less* than belief in the theory's truth. For example, one might wish to argue that the acceptance of a scientific theory requires only belief in the theory's empirical adequacy, as in van Fraassen [1980]. [*Nat**] would then deliver the conclusion that we are justified in believing our best overall theory of the world to be empirically adequate, but not the conclusion that we are justified in believing

[39] See Burgess and Rosen [1997], I.A.2.c.

the overall theory to be true.

The Naturalistic Argument will go through even if [*Nat*] is replaced by [*Nat**]. But it will yield the result that we are justified in *accepting arithmetic*—in whatever sense of 'acceptance' is appropriate—rather than the result that we are justified in believing arithmetic to be true. Thus, if the Naturalistic Argument is to supply the foundation for a realist account of arithmetical language, one cannot rest content with less than [*Nat*]. This is a potential problem, but only for those who are, at the same time, realists about mathematics and uncomfortable enough about scientific realism to give up [*Nat*].

5. *Circularity*

One might worry that our use of the Naturalistic Argument is problematic as a defense of the claim that $\langle \mathbf{N}, \mathrm{HP} \rangle$ is a successful stipulation. For the Naturalistic Argument rests on the assumption that we are already in possession of a developed overall theory of the world, which includes arithmetic. There is therefore room for the complaint that we have no story to tell about how an agent with no prior arithmetical knowledge could come to know that $\langle \mathbf{N}, \mathrm{HP} \rangle$ is a successful stipulation. Note, however, that an agent with no prior arithmetical knowledge might begin by laying down $\langle \mathbf{N}, \mathrm{HP} \rangle$ without knowing whether her stipulation has been successful. If she then sets out to do science, she might discover that she can produce an overall scientific theory better than any other (by the lights of her current scientific standards) if she makes use of certain consequences of HP, the truth of which would imply the infinity of the universe. This would justify her in thinking that her original stipulation was successful after all.

8. Conclusion

Neo-Fregeans must provide a defense of [*Success*]. A thesis like [*Success by Default*] might eventually yield one way of addressing the problem, but so far it has not received proper justification. Fortunately, there is an alternative. The Naturalistic Argument gives us independent reasons for thinking that the universe is infinite.

References

BOGHOSSIAN, P., and C. PEACOCKE, eds. [2000]: *New Essays on the* A Priori. Oxford: Clarendon Press.

BOOLOS, G., ed. [1990a]: *Meaning and Method: Essays in Honor of Hilary Putnam*. Cambridge: Cambridge University Press.

——— [1990b]: 'The standard equality of numbers', in Boolos [1990a], pp. 261–278. Reprinted in Boolos [1998], pp. 202–219.

——— [1993]: 'Whence the contradiction?', *Aristotelian Society Supplementary Volume* **67**, 213–233. Reprinted in Boolos [1998], pp. 220–236.

——— [1997]: 'Is Hume's principle analytic?', in Richard Heck [1997b], pp. 245–261. Reprinted in Boolos [1998], pp. 301–314.

——— [1998]: *Logic, Logic and Logic*. Cambridge, Mass.: Harvard University Press.

BURGESS, J., and G. ROSEN [1997]: *A Subject With No Object*. New York: Oxford University Press.

DAVIES, M. [2000]: 'Externalism and armchair knowledge', in Boghossian and Peacocke [2000], pp. 384–414.

DE ROSE, K. [1995]: 'Solving the skeptical problem', *Philosophical Review* **104**, 1–52.

DRETSKE, F. [1970]: 'Epistemic operators', *Journal of Philosophy* **67**, 1007–1023.

FIELD, H. [1980]: *Science Without Numbers*. Oxford: Basil Blackwell and Princeton: Princeton University Press.

——— [1982]: 'Realism and anti-realism about mathematics', *Phil. Topics* **13**, 45–69. Reprinted in Field [1989], pp. 53–78.

——— [1989]: *Realism, Mathematics and Modality*. Oxford: Basil Blackwell.

FINE, K. [1998]: 'The limits of abstraction', in Schirn [1998], pp. 503–629.

FREGE, G. [1884]: *Die Grundlagen der Arithmetik*. English translation J. L. Austin, *The Foundations of Arithmetic*. Evanston, Ill.: Northwestern University Press, 1980.

——— [1893/1903]: *Grundgesetze der Arithmetik*. Vol. 1 (1893), Vol. 2 (1903). English translation Montgomery Furth, *The Basic Laws of Arithmetic*. Berkley and Los Angeles: University of California Press, 1964.

HALE, B. [1997]: '*Grundlagen* §64', *Proceedings of the Aristotelian Society* **97**, 243–261. Reprinted in Hale and Wright [2001a], pp. 91–116.

——— [1999]: 'Arithmetic reflection without Intuition', *Aristotelian Society Supplementary Volume* **73**, 75-98.

HALE, B., and C. WRIGHT [2000]: 'Implicit definition and the *a priori*', in Boghossian and Peacocke [2000], pp. 286–319. Reprinted in Hale and Wright [2001a], pp. 117–150.

——— [2001a]: *The Reason's Proper Study: Essays towards a Neo-Fregean Philosophy of Mathematics*. Oxford: Clarendon Press.

——— [2001b]: 'To bury Caesar. . . ', in Hale and Wright [2001a], pp. 335–396.

HECK, R. [1992]: 'On the consistency of second-order contextual definitions', *Noûs* **26**, 491–494.

——— [1997a]: 'Finitude and Hume's principle', *Journal of Philosophical Logic* **26**, 589–617.

———, ed. [1997b]: *Language, Thought and Logic*. Oxford: Clarendon Press.

Horwich, P. [1997]: ‘Implicit definition, analytic truth and *a priori* knowledge’, Noûs **31**, 423-440.

——— [1998]: *Meaning*. Oxford: Clarendon Press.

Lewis, D. [1996]: ‘Elusive knowledge’, *Australasian Journal of Philosophy* **74**, 549–567.

Maddy, P. [1997]: *Naturalism in Mathematics*. Oxford: Clarendon Press.

Pryor, J. [2000]: ‘The skeptic and the dogmatist’, *Noûs* **34**, 517–549.

Rayo, A. [2002a]: ‘Word and objects’, *Noûs* **36**, 436–464.

——— [2002b]: ‘Frege’s unofficial arithmetic’, *Journal of Symbolic Logic* **67**, 1623–1638.

Rayo, A., and S. Yablo [2001]: ‘Nominalism through de-nominalization’, *Noûs* **35**, 74-92.

Schirn, M., ed. [1998]: *The Philosophy of Mathematics Today*. Oxford: Clarendon Press.

Shapiro, S., and A. Weir [1999]: ‘New V, ZF and abstraction’, *Philosophia Mathematica* (3) **7**, 293–321.

——— [2000]: ‘“Neo-logicist” logic is not epistemically innocent’, *Philosophia Mathematica* (3) **8**, 160–189.

van Fraassen, B. C. [1980]: *The Scientific Image*. Oxford: Clarendon Press.

Weir, A. [forthcoming]: ‘Neo-Fregeanism: An embarrassment of riches’, *Notre Dame Journal of Formal Logic*.

Wright, C. [1983]: *Frege’s Conception of Numbers as Objects*. Aberdeen: Aberdeen University Press.

——— [1997]: ‘The philosophical significance of Frege’s theorem’, in Heck [1997b], pp. 201–244.

——— [1999]: ‘Is Hume’s principle analytic?’, *Notre Dame Journal of Formal Logic* **40**, 6–30. Reprinted in Hale and Wright [2001a], pp. 307–332.

——— [2000a]: ‘Cogency and question-begging: Some reflections on McKinsey’s paradox and Putnam’s proof’, *Philosophical Issues* **10**, 140–163.

——— [2000b]: ‘Replies’, *Philosophical Issues* **10**, 201–219.

——— [forthcoming]: ‘Wittgensteinian certainties’, in D. McManus, ed., *Wittgenstein and Scepticism*. London: Routledge.

Abstract. I argue that Neo-Fregean accounts of arithmetical language and arithmetical knowledge tacitly rely on a thesis I call [*Success by Default*]—the thesis that, in the absence of reasons to the contrary, we are justified in thinking that certain stipulations are successful. Since Neo-Fregeans have yet to supply an adequate defense of [*Success by Default*], I conclude that there is an important gap in Neo-Fregean accounts of arithmetical language and knowledge. I end the paper by offering a naturalistic remedy.