13

On the Harmless Impredicativity of $N^=$ ('Hume's Principle')

CRISPIN WRIGHT

I. INTRODUCTION: DUMMETT'S OBJECTIONS TO 'NEO-FREGEANISM'

Michael Dummett's *Frege: Philosophy of Mathematics* and its famous precursor, *Frege: Philosophy of Language*, contrast sharply in their estimates of Frege's achievements in those respective subjects.[1] The guiding conviction of *Frege: Philosophy of Language* was that even admirers of Frege's writings have tended to a merely superficial appreciation of his contribution, that he is justly viewed as the father of analytical philosophy itself, the first philosopher to perceive the centrality of the philosophy of language, and especially the theory of meaning, in all philosophy, and the inventor of the tradition of systematic thought about meaning within which so much of important recent and contemporary philosophy belongs. By contrast, *Frege: Philosophy of Mathematics*—notwithstanding the kindest possible concluding accolade[2]—returns a predominantly negative verdict on Frege's accomplishment in relation to the specific questions which were most important to him. In particular, to the questions at which *Grundlagen* and *Grundgesetze* are primarily directed—the character of the subject-matter of number theory and real analysis and the nature of our knowledge about it—Frege's distinctive Platonist-cum-logicist answer was, according to Dummett, demonstrably 'catastrophically wrong'.[3]

My thanks to all who participated in the discussion at the Munich conference and especially to Bob Hale and Stewart Shapiro.

[1] Michael Dummett, *Frege: Philosophy of Language*, London, Duckworth, 1973; *Frege: Philosophy of Mathematics*, London, Duckworth, 1991.

[2] '[T]he greatest philosopher of mathematics yet to have written', *Frege: Philosophy of Mathematics*, 321.

[3] *Frege: Philosophy of Mathematics*, 292. The emphatic tone of this assessment is sustained in numerous passages in Dummett's text. Compare 'the central flaw in [Frege's] entire philosophy of arithmetic', p. 232; 'the big error . . . in [Frege's] philosophy of mathematics' and '[Frege's] philosophy of arithmetic was, indeed, fatally flawed', p. 321; 'The inconsistency of Frege's . . . system was not a mere accident [but] . . . an error affecting his entire philosophy', p. 223; 'the task at which Frege so lamentably failed', p. 317; and so on.

Crispin Wright, *On the Harmless Impredicativity of N= ('Hume's Principle')*. In: *The Philosophy of Mathematics Today,* Edited by: Matthias Schirn, Oxford University Press. © Oxford University Press 1998. DOI: 10.1093/oso/9780198236542.003.0014.

This of course has long been the usual view of the matter. Indeed, Frege himself was among the earliest to subscribe to it. The dominant opinion throughout this century has been that Russell's Paradox totally exploded Frege's approach, showing that the nature of numbers, and our knowledge of them, cannot be explained and justified in anything like Frege's way.

Thirteen years ago I published a short book[4] (henceforward *Frege's Conception*) which challenged this view, arguing that Frege's Platonism—the conception of numbers as specific objects providing the proper subject-matter of arithmetic, but given to us independently of any special 'intuition'—admitted of an epistemologically responsible, powerful defence along the very lines suggested by Frege's writing, and that—the paradox notwithstanding—(something to the purpose of) the logicism too was salvageable as a programme, at least as far as number theory was concerned, and maybe further.

It merits emphasis that Dummett's negative assessment is qualified by his conviction that, at some general level and in a way which I do not think he succeeds in making very clear, the logicist conception of mathematics is an insight,[5] and by his acceptance that the defence of Platonism, based on the Context Principle, which is offered in *Frege's Conception* is both true to the intent of *Grundlagen* and effective against the philosophical mind-set which has typically driven nominalists and reductionists of various stripes.[6] Given these broad sympathies, the question arises where specifically Dummett would take issue with the pro-Fregean lines of argument offered in *Frege's Conception*.

We are spared the need to speculate, for the latter book comes in for quite extensive criticism in *Frege: Philosophy of Mathematics*, with four principal points of abrasion emerging:

(i) In the chapter on 'Abstract Objects' in *Frege: Philosophy of Language*, and others of his earlier writings,[7] Dummett contended that Frege's use of the Context Principle to justify the ascription of reference to numerical singular terms, while correct as a counter to the nominalist *denial* of the reality of abstract objects, serves to vindicate only a 'thin' (contrast: *realistic*) conception of reference for such terms—a conception therefore at odds with Frege's full-blooded realism about the items for which they stand. The coherence of this idea of Dummett's was questioned in some detail in *Frege's Conception*.[8] In his new book Dummett repeats the claim—or some-

[4] Crispin Wright, *Frege's Conception of Numbers as Objects*, Aberdeen University Press/Humanities Press, 1983.

[5] See e.g. *Frege: Philosophy of Mathematics*, 308–9.

[6] See e.g. pp. 155–6, 181–3, 207–8, 231, 236, and 309–10.

[7] See also for instance Michael Dummett, *The Interpretation of Frege's Philosophy*, London, Duckworth, 1981, 424–7 and 452–7.

[8] See *Frege's Conception* ch. 2, sect. x.

thing very close to it[9]—insisting that the required distinction between two conceptions of reference is valid, and that the Fregean has no right to any but a 'non-realistic' notion in application to contextually defined singular terms.[10]

(ii) The core of the Fregean defence of a Platonist account of a region of discourse is that a referential construal of the relevant range of abstract singular terms may be explained and justified by *abstraction*—by laying down equivalences like $N^=$, that (where 'Nx:Fx' stands for the number of things which are F):

> For all F, G: Nx:Fx = Nx:Gx $\leftrightarrow$ F and G are equinumerous (i.e. may be put into one–one correspondence)

or the corresponding principle for the identity of directions,

> For all lines, a, b: the direction of a = the direction of b $\leftrightarrow$ a and b are parallel.

In the view of the Fregean, such a principle, stipulating that the truth-conditions of its two sides are to coincide, allows us to pass from a prior understanding of the type of context featured on its right-hand side to a grasp of the type of context which features on the left. Naturally, no one would deny that we can do that if the left-hand side is taken merely as a notational variant, an *unstructured* equivalent of the right. What is distinctive about the Fregean perspective is that we are to take it, in our passage across the biconditional, that the syntax of the left-hand side is *just what it seems*, and that the familiar expressions it contains—in the case of the direction principle, for example, the symbol for identity and the embedded names of lines—are to have their customary meaning. The epistemology of the abstract objects spoken of in contexts of the kind typified on the left-hand side is thus taken to be just that of the presumed (relatively) unproblematic states of affairs depicted by the kind of context featuring on the right.

[9] It is moot how close. Dummett writes (p. 191):

> 'Ultimately, Wright fails to find this intermediate view'—that attributed to Dummett in *Frege's Conception*—'coherent: he doubts if there is any tenable position between the austere'—eliminative reductionist—'and robust'—realist—'interpretations. As concerns contextual definitions, properly so called, I shall here maintain an intermediate view, perhaps one more austere than that which Wright had in mind. I shall however spend no time in discussing either how faithfully Wright represents the views I expressed in *Frege: Philosophy of Language*, or how far those I advance here diverge from them.'

This is regrettable. Clarity would have been best served if Dummett had taken the trouble to respond to the *Frege's Conception* criticisms, and to define any respects in which the new 'intermediate' view represents a change of mind. For an attempt to do the work for him, and argument that the new view is no stronger, see Bob Hale's 'Dummett's Critique of Wright's Attempt to Resuscitate Frege', in *Philosophia Mathematica*, (3) vol. 2 (1994), 122–47; sect. 4: 133–8.

[10] See *Frege: Philosophy of Mathematics*, 189–99, 234–5, and 239 for development of this line of criticism.

Is this line so much as coherent? Dummett apparently argues that it is not—or at least that anything recognizable as a Fregean notion of sense must be a casualty of it. For Frege, sense determines reference. So if the senses of the two sides of, say, the direction principle are the same, they must purport the same references. The Fregean Platonist must therefore acknowledge the presence of reference to directions on the right-hand sides of instances of the direction principle, even though it has to be possible—in his view—fully to understand the sense of such clauses while wholly unaware of the concept of direction. And that is to violate an important element in the *transparency* of Fregean sense: that it must be transparent to one who grasps the sense of an expression what references to what kinds of objects its featuring in (or being itself) a true sentence would involve.[11]

(iii) The third area of disagreement concerns the 'Caesar problem'—the indeterminacy which Frege thought afflicted $N^=$ as an account of numerical identity, as manifested in its incapacity to decide statements of the form,

Nx:Fx = Julius Caesar,

or more generally, any statement of the form,

Nx:Fx = q,

where 'q' is a term not given in the form, 'Nx:Fx', nor introduced as an abbreviation for such a term. As is familiar, it was in effect the Caesar problem that moved Frege to identify the cardinal numbers with extensions—a futile response, even prescinding from the inconsistency of his theory of extensions, since the problem must arise for extensions too. I absolutely agree that, for a variety of reasons, one of which will be very salient later, it is essential to find some constructive response to the difficulty. But in Dummett's view, the defence of a Fregean philosophy of arithmetic offered in *Frege's Conception* depends on an inadequate solution to the Caesar problem, indeed one actually canvassed and rejected by Frege himself.[12]

[11] For elaboration, see *Frege: Philosophy of Mathematics*, 168–79 and 194–5.

[12] Specifically, Dummett interprets the *Frege's Conception* proposal as a licence to treat 'the way in which the object q is introduced as a property of q, which it is not' (*Grundlagen* §67; *Frege: Philosophy of Mathematics*, 160). This is a bad misreading of what is actually suggested at pp. 113–17 of *Frege's Conception*. In particular, the proposal, so interpreted would, as Dummett emphatically observes, have the effect that numbers cannot be classes—since they are not introduced as such—whereas I explicitly remarked (*Frege's Conception*, 115) that, pending further refinement, the effect of the proposal would seem to be precisely the opposite. Maybe Dummett attributed that to mere confusion. In any case, what *is* presupposed by the *Frege's Conception* proposal is not what Dummett—and Frege—object to but rather that the way in which it is (correctly) introduced may teach one something of the properties of a *concept*, and *thereby* something about objects falling under it. I will return to this briefly in Sect. V below. For further discussion of Dummett's misreading, see Hale's 'Dummett's Critique of Wright's Attempt to Resuscitate Frege', 130–3.

(iv) Finally, Dummett contends that

The inconsistency of [*Grundgesetze*] was not a mere accident (though a disastrous one) due to carelessness of formulation. [Frege] discovered, by August 1906, that it could not be put right within the framework of the theory, that is, with the abstraction operator as primitive and an axiom governing the condition for the identity of value-ranges: but the underlying error lay much deeper than a misconception concerning the foundations of set theory. It was an error affecting his entire philosophy.[13]

And, if so, one inherited by *Frege's Conception*. For Dummett, the inconsistency of Basic Law V points to a fundamental and irremediable flaw in the whole conception of how abstract objects may be given to us which permeates Frege's work, and my book.

Each of these four claims involves, if cogent, a fundamental line of criticism, and deserves treatment in detail. Considerations of space demand some restriction of focus, however, and—while strenuously disagreeing with each of them—I therefore elect here to concentrate on that to which Dummett himself appears to attach the greatest importance, the fourth.[14]

II. BAD COMPANY?

A strict definition should provide for the elimination from any context in which it occurs of the expression, or type of expression it defines. $N^=$, and Basic Law V, are not, by this criterion, strict definitions. Numerical singular terms, as introduced via $N^=$, and singular terms standing for extensions or, more generally, for courses of values, as introduced by Basic Law V, will not always be eliminable via the paraphrases licensed by those two principles. The reason is that predicates (open sentences) containing singleton occurrences of singular terms of those kinds will be wanted to feature among those to which the cardinality operator, and its analogue for courses of values (the abstraction operator), may permissibly be applied. It is, of course, vital for Frege's purposes that this be so, since his proof of the infinity of the number series depends upon its being allowable to number numbers which precede a given number, and hence to bind with the cardinality operator predicates, like

$y = Nx: x \neq x$,

[13] *Frege: Philosophy of Mathematics*, 223

[14] Each of the four lines of objection is usefully addressed in Bob Hale's paper cited above. The second objection was anticipated and treated in my 'Field and Fregean Platonism', in A. D. Irvine (ed.), *Physicalism in Mathematics*, Dordrecht, Kluwer, 1990, 73–93, sect. 4.

which feature occurrences of that operator which $N^=$ provides no means to eliminate.[15]

One effect of this consideration is to the Fregean's advantage. Even were Dummett to succeed in making out a contrast between 'thin' and robust, or 'realistic' singular reference, and in showing that contextually definable abstract singular terms qualify only for the former, it would still remain to be demonstrated that numerical singular terms, introduced via $N^=$, are in the same boat—and only if they are would Dummett's claim be justified that Frege's Platonism about number, to the extent that it is sustained by the Context Principle, cannot provide a sufficient foundation for his arithmetical realism.[16]

But of course there is a concomitant disadvantage: since $N^=$ does not provide a strict *definition* of every necessary and legitimate kind of use of terms formed using the cardinality operator, the question becomes urgent whether and how $N^=$ is really satisfactory as an *explanation* of that range of numerical singular terms. No doubt it would be too simple to assume that whatever is not a strict definition of a symbol is not a satisfactory explanation of it either. Still, definition, whether explicit or contextual, is one readily intelligible form of explanation. If $N^=$ succeeds as an explanation of uses of the cardinality operator for which it provides no strict definition, we need to hear a story about how.

Here is the point of entry of Dummett's fourth criticism—what he presents as the most serious and fundamental criticism of Frege's entire approach. In Dummett's view, $N^=$ does *not* provide the necessary explanation, and the general model it incorporates, of how the use of a range of abstract singular terms, and with it our conception of an associated population of mathematical objects, may be explained and justified, is altogether misconceived. Why does he think this?

Dummett writes:

> [I]f the context principle, as expounded by Wright, is enough to validate the 'contextual' method of introducing the cardinality operator, it must be enough to validate a similar means of introducing the [class] abstraction operator.

15 Just to elaborate a little. The statement

(a) $Nx{:}x\neq x = Nx{:}x\neq x$

—the statement that the number of things which are not self-identical is identical to itself—can be transformed via $N^=$ into a theorem of pure second-order logic with identity, to the effect that there is a one–one correspondence between the instances of the concept, *not self-identical*, and themselves. But now consider:

(b) $(\exists y)(y = Nx{:}x\neq x)$.

This is, of course, an immediate consequence of (a) in any logic suitable for the Fregean's purpose. But its singleton occurrence of the numerical operator is one which the resources afforded by $N^=$ provide no means to eliminate.

16 Dummett is under no illusion about this, of course (*Frege: Philosophy of Mathematics*, 205).

This is why the mere fact that, on his view, it is unnecessary to define the cardinality operator in terms of classes or of value-ranges does not entitle Wright to ignore the problem of the abstraction operator. For Frege's method of introducing [the latter] was, notoriously, *not* in order. It rendered his system inconsistent; and that inconsistency forced him eventually to acknowledge that his entire enterprise had failed. If the context principle, as stated by Wright, was sound, there could have been no inconsistency . . . we may take [Wright] as concerned to vindicate, by appeal to the context principle, a method of introducing the cardinality operator which Frege did not in fact adopt: namely, by laying down the criterion of the identity for numbers [$N^=$], and supplementing it by some solution to the Julius Caesar problem . . . To all appearances, this would exactly resemble Frege's method of introducing the abstraction operator in *Grundgesetze.* In this case, we therefore have three options: to reject the context principle altogether; to maintain it, but declare that it does not vindicate the procedure Wright has in mind; and to formulate a restriction upon it that distinguishes the cardinality operator from the abstraction operator. Wright does none of these things: he maintains the context principle in full generality, understood as he interprets it, and defends the appeal to it to justify ascribing a reference to numerical terms, considered as introduced in the foregoing manner, without stopping to explain why an apparently similar method of introducing value-range terms should have led to contradiction. He owes us such and explanation.[17]

Two things about this. First, it should not seem at all obvious that I *do* owe any such explanation. Dummett here writes as though nothing at all were known about the likelihood of similar inconsistency afflicting $N^=$. But in fact, as Boolos and Burgess have independently observed,[18] the system which results from superimposing $N^=$ onto a suitable foundation of higher-order logic is consistent if classical analysis is. Dummett is right that there is a structural similarity between Frege's own procedure with Basic Law V and the way I proposed that a Fregean might deploy $N^=$; but the question is why—when there is relative security about the consistency of $N^=$—the fate of Basic Law V should nevertheless be thought to cast a pall of suspicion over its status. So far from my needing to explain why, in the teeth of the contradiction, it might still be justifiable to deploy $N^=$ in the fashion proposed, the onus is rather upon the critic to explain why, in the teeth of the presumable consistency of $N^=$, it is not.

Second, it anyway isn't true that I didn't 'stop to explain' why an apparently similar manner of introducing value-range terms should have

[17] *Frege: Philosophy of Mathematics*, 188–9.

[18] See the discussion at pp. 6–10 of George Boolos's 'The Consistency of Frege's *Foundations of Arithmetic*', in J. Thomson (ed.), *On Being and Saying: Essays in Honor of Richard Cartwright*, Cambridge, Mass., MIT Press, 1987, 3–20, and John Burgess's review of *Frege's Conception* in *Philosophical Review*, 93 (1984), 638–40; also Harold Hodes's 'Logicism and the Ontological Commitments of Arithmetic', *Journal of Philosophy*, 81 (1984), 123–49—(see remarks at 138 concerning an equivalent of $N^=$ involving branching quantifiers). For a detailed proof, see the first appendix to Boolos and Heck, '*Die Grundlagen der Arithmetik*, §§82–3', in this volume.

led to contradiction. On the contrary, a perfectly good explanation of that is provided by the comparative discussion[19] of the derivation of Russell's Paradox from Basic Law V and the outcome of the corresponding reasoning from $N^{=}$. As is familiar, in order to obtain Russell's Paradox from Basic Law V:

$$(\forall F)(\forall G)(\{x: Fx\} = \{x: Gx\} \leftrightarrow (\forall x)(Fx \leftrightarrow Gx)),$$

we first derive:

$(\forall F)(\exists y)\ y = \{x: Fx\}$—'naïve' abstraction for extensions—

and then take F to be the concept of non-self-membership, which we may express as:

$$(\forall G)(z = \{x: Gx\} \rightarrow -Gz),$$ [20]

thus arriving at Russell's class,

$$\{z: (\forall G)(z = \{x: Gx\} \rightarrow -Gz)\}.$$

Call this class **r** and suppose it satisfies the condition on its own members, i.e.:

$$(\forall G)(\mathbf{r} = \{x: Gx\} \rightarrow -G\mathbf{r})$$

and take for G that very condition. Then, since $\mathbf{r} = \{x: Gx\}$, we have $-G\mathbf{r}$; so **r** does not satisfy that condition. Hence: $(\exists G)(\mathbf{r} = \{x: Gx\}\ \&\ G\mathbf{r})$. But the condition which makes this existential statement true must, by Basic Law V, be coextensive with that by reference to which **r** was defined. Hence **r** must, after all, satisfy the latter. And so on, back and forth.

When the paradox is presented in this rather cumbersome fashion, rather than, as is normal, in terms of a primitive symbol of class membership, it is evident both why there might seem to be the prospect of a similar antinomy with $N^{=}$, and why that prospect is not fulfilled. Again, we may first use $N^{=}$ to obtain

$(\forall F)(\exists y)(y = Nx{:}Fx)$—'naïve' abstraction for numbers—

and then take as F a precisely analogous condition to Russell's, viz.

$$(\forall F)(z = Nx{:}Fx \rightarrow -Fz),$$

thus arriving at the potentially rogue number:

$$Nz{:}\ (\forall F)(z = Nx{:}Fx \rightarrow -Fz)$$

[19] *Frege's Conception*, 155–6

[20] No particular importance attaches to this way of characterizing the notion.

—the number of numbers which fall under none of the concepts of whose instances they are the number. Call this number **s.** We now ask whether **s** falls under its own defining condition, i.e. whether

$$(\forall F)(\mathbf{s} = Nx{:}Fx \rightarrow -F\mathbf{s})$$

If so, then, taking F as that very condition, we immediately have that it fails to do so. So far, so bad, as it were. But can we complete the other half? We have:

$$(\exists F)(\mathbf{s} = Nx{:}Fx \ \&\ F\mathbf{s}),$$

but now the next step fails us. We cannot infer, from the fact that **s** falls under *some* condition of whose instances it is the number, that it falls under its originally defining condition above. We cannot do so precisely because, in contrast with the situation of classes, where we have that

$$(k = \{x{:}\ Fx\} \ \&\ k = \{y{:}\ Gy\}) \rightarrow (Fk \leftrightarrow Gk),$$

we clearly cannot take it generally that:

$$(k = Nx{:}Fx \ \&\ k = Ny{:}Gy) \rightarrow (Fk \leftrightarrow Gk).$$

All that results, then, is a proof that **s** falls under some concept of whose instances it is the number. That, so far from being paradoxical, is surely true for all cardinal numbers generally, finite as well as infinite.[21]

A simple, head-on response to Dummett's demand that a restriction be formulated to distinguish the case of the cardinality operator from that of the class-abstraction operator is merely to exploit this difference in the obvious way. If:

$$(\forall F)(\forall G)\ \Sigma x{:}Fx = \Sigma x{:}Gx \leftrightarrow Fx \approx Gx,$$

is a proposed Fregean abstraction whose right-hand side expresses an equivalence relation on *concepts*, the question may be put whether that relation is or is not such that a Σ-object's falling under *some* concept, F, of which it is Σx: Fx entails that it falls under *every* F of which it is Σx:Fx. Say that the relation is *Russellian* if so. Then if, but only if, such a relation is Russellian will it be possible to confound the mooted principle by an application of the reasoning which leads to Russell's Paradox. So the demanded restriction is merely that the relevant equivalence relation should not be Russellian. Coextensiveness violates this restriction; one–one correspondence does not.[22]

[21] Actually, not quite all. Zero is evidently an exception. But it would seem to be the only exception, all other numbers falling under *some* concept to which they belong. Thus, as J. L. Bell has observed to me in discussion, in a domain consisting just of cardinal numbers, **s** is in fact Frege's number 1—the number belonging to the concept, $x = 0$. (Elsewhere, where κ is the number of non-numbers, **s** will be $\kappa + 1$.)

[22] Someone who thinks such a restriction would be 'ad hoc' (though I am not very clear

III. DUMMETT'S REAL OBJECTION

These reflections address the letter of Dummett's quoted complaint. But he would undoubtedly complain they are mere skirmishing in the context of the point he primarily intends. They would be *apropos* if all that could be amiss with an explanatory schema of this general kind was inconsistency. But explanations can have less dramatic shortcomings.[23]

Dummett's thinking about the issues here is strikingly of the kind which Wittgenstein set himself against in his notorious discussions of inconsistency and the paradoxes in the *Remarks on the Foundations of Mathematics*—thinking which would have it that Russell's Paradox is merely the symptom of an *underlying* disorder, whose insidious workings can be present even in cases where that especially distressing symptom is not

what the force of the complaint would be) may be interested to think through how any second-order abstraction on a Russellian relation is at the service of a development in some ways akin to *Cantor's Paradox*; that is, will call for objects which, on pain of contradiction, may not be reckoned to occur within an original (putatively all-inclusive) domain. Suppose to the contrary that it is possible to abstract on a Russellian relation, R, in such a way that, for each concept F defined over the domain, an object, z, can be found within the domain whose properties are those of the relevant abstract, ΣF. Then there should be a one–one map, M, containing all and only pairs such that, for each concept F and the corresponding equivalence class of concepts under R, {F, G, ...}, one term is a z, lying within the domain, whose properties are exactly those of ΣF and the other is {F, G, ...}. Consider a particular such pair, ⟨z: {F, G, ...}⟩. Since each of F, G, etc., have the same Σ-abstract, z must have the properties of each of ΣF, ΣG, etc. So it follows from the Russellian character of R that z instantiates all or none of the concepts, F, G, etc. That immediately suggests the Cantorian thought. Form the concept C instantiated just by each z which instantiates none of the concepts in its M-correlate. C will be an element of an equivalence class under R, {C, . . , . .}, which, by hypothesis, will have an M-correlate, k, in the original domain. Does k instantiate C? If it does, then, by the definition of C, it instantiates none of the concepts in its M-correlate. C is one such concept; so k does *not* instantiate C. But in that case the Russellian character of the abstractive relation ensures that k instantiates no other concept in {C, . . , . .}. So k instantiates none of the concepts in its M-correlate, and *is* therefore an instance of C . . . . . . Thus no element of the domain can play the role of ΣC.

Clearly it will not do to attempt a Fregean abstraction on an equivalence relation which is *inflationary* (to borrow Kit Fine's useful term) in this way: one which, on any domain of objects, generates more equivalence classes of concepts than there can be corresponding objects in the domain. That is a restriction whose point could have been appreciated before we had any inkling of Russell's Paradox.

[23] An important line of objection to the neo-Fregean perspective on $N^=$, about which I shall say nothing in this paper, branches off at this point, based on the observation that there are principles of the same general type as $N^=$ and Basic Law V which although, unlike the latter, consistent, are inconsistent with the former. An example would be the abstraction on the second-order equivalence relation: 'F and G are such that finitely many things are one or the other but not both.' The system generated by adding that principle to second-order logic has finite but no infinite models. And a problem is therefore raised along the lines: what allows $N^=$ to rank as (something like an) analytic principle at the expense of such competitors which seem, formally, perfectly analogous and no less legitimate?

This objection is original to and was developed by George Boolos in his 'The Standard of Equality of Numbers', in G. Boolos (ed.), *Meaning and Method: Essays in Honor of Hilary Putnam*, Cambridge, Cambridge University Press, 1990, 261–77. It is discussed in detail in my 'The Philosophical Significance of Frege's Theorem', in Richard Heck, Jr. (ed.), *Language, Thought, and Logic,* Oxford, Oxford University Press, 1997, pp. 201–44.

present.[24] For Dummett, Frege's whole conception, as embodied in the Context Principle, of what it takes to justify and explain the use of a range of singular terms is disfigured by a flaw, at work in the generation of the contradiction, but also present in cases, like $N^=$, where a contradiction is not forthcoming.

What is this flaw? It appears to be Dummett's view that explanatory principles like $N^=$ and Basic Law V, taken in the way intended, fail to be explanatory because *circular*—because they presuppose the very conceptual resources they are supposed to explain. It is very familiar that there is no deriving Russell's Paradox unless the rogue set is taken to lie within the range of the universal quantifier by reference to which, in a step of naïve abstraction, it is defined—we have to be able to pass from

$$(\forall x)x \in \mathbf{r} \leftrightarrow -(x \in x)$$

to

$$\mathbf{r} \in \mathbf{r} \leftrightarrow -(\mathbf{r} \in \mathbf{r}).$$

In the context of the derivation of the paradox licensed by Basic Law V, the point emerges as the reflection that the courses of values, or more specifically extensions, for which Basic Law V enables us to introduce singular terms, are taken to lie within the range of the universal quantifier which occurs on its right-hand side. And this feature is, of course, also present in $N^=$: the fully explicit expansion of the statement that there is a one–one correspondence between F and G naturally involves both first-order and higher-order quantification, and the referents of the numerical singular terms which $N^=$ enables us to introduce are themselves taken to lie within the relevant range of first-order quantification.

In Dummett's view, it seems, this impredicativity imports a vicious circularity: it is viciously circular to attempt to define a term-forming operator by reference to a domain of quantification which is intended to include the referents of the very expressions which that operator enables us to form. An analogue of Basic Law V which was purified of this feature would not allow the derivation of the paradox—so to that extent, the feature is responsible for the paradox. But, so apparently thinks Dummett, the feature is objectionable even in the consistent case.[25]

[24] See e.g. *Remarks on the Foundations of Mathematics*, pt. III, §80. I am not, of course, suggesting that one should take Dummett's concern lightly purely on this account!

[25] Here are some illustrative quotations (all from *Frege: Philosophy of Mathematics*):

> Suppose it [N=] had been presented as an axiom governing the cardinality operator, taken as primitive, as Axiom V governs the abstraction operator in *Grundgesetze*. . . . So considered, it would have been as objectionably impredicative as the analogous specification for value-ranges: for the truth of any statement of identity between numbers would depend on the extensions of two predicates defined on a domain which included the cardinal numbers, and whose composition the axiom was supposed to be playing an essential part in determining. (p. 226)

For Fregean Platonism, it suffices, in order to justify the claim that a certain class of expressions possess objectual reference, to observe first that their syntactic function is that of ordinary, uncontroversial singular terms;[26] and second, that they so function in statements which, in the light of some canonical explanation of their truth-conditions, we are justified in claiming to be true. Now Dummett, as I read him, does not intend to question this broad conception of the matter—on the contrary, he regards it as incorporating a decisive insight against what he himself regards as the misbegotten forms of nominalism that have benighted discussions of mathematical existence in this century. What Dummett *does* hold is that the force of this defence of Platonism stands in need of qualification in response to the following dilemma. Of any higher-order Fregean abstraction we may ask: Is the field of the relation cited on its right-hand side to include concepts under which may fall the very items to which the terms introduced on the left-hand side purportedly refer? If not, then the equivalence may constitute or contribute towards what is genuinely a contextual *definition* of the relevant terms; but the notion of reference which we will thereby be entitled to apply to them will be at most the etiolated or 'thin' notion with which Dummett wants to contrast a full-blooded, 'realistic' counterpart. But if, on the other hand, the relation does include such concepts within its field, then, according to Dummett, we have not yet met the conditions for the Fregean Platonist defence to be activated: specifically,

> The context principle allows us to ascribe a reference to mathematical terms provided that we have fixed the truth-conditions of sentences in which they occur;

> Frege leapt to the conclusion that the basis for introducing any new range of abstract terms must consist in the determination of the truth-conditions of identity statements involving them. In a certain sense, this was not far from the truth. It led, however, to the root confusion that allowed him to believe that he could simultaneously fix the truth-conditions of such statements and the domain over which the individual variables were to range. This belief was a total illusion. . . . we have *first* to attain a grasp of the intended domain of the individual variables: it is only after that that we can so much as ask after the meanings of the primitive non-logical symbols. (p. 232)

> [Frege] had, however, fatally overlooked the circularity of the whole procedure: that of specifying the criterion of identity in terms of the truth of sentences of the theory, and, more generally, that of attempting simultaneously to specify the domain and the application of the primitive predicates to its elements. (p. 233)

Also see esp. the passage from p. 236 quoted in the text below.

[26] 'Proof-theoretic' might be a little happier than 'syntactic', since it is the behaviour of the terms in inferential contexts which is important. Dummett's pioneering discussion of how to characterize this notion of a singular term is in *Frege: Philosophy of Language*, ch. 4. Bob Hale points up certain difficulties for Dummett's proposals in 'Strawson, Geach and Dummett on Singular Terms and Predicates', in *Synthese*, 42 (1979). The issue is the focus of §ix of *Frege's Conception* and of ch. 2 of Hale's *Abstract Objects*, Oxford, Blackwell, 1987. Linda Wetzel responds to Hale in 'Dummett's Criteria for Singular Terms', *Mind*, 99 (1990), 239–54. Hale responds to Wetzel and to other difficulties in two recent papers: 'Singular Terms (1)' in Matthias Schirn (ed.), *Frege: Importance and Legacy*, Berlin, Walter de Gruyter, 1996, 438–57; and 'Singular Terms' in B. McGuinness and G. Oliveri, eds., *The Philosoply of Michael Dummett*, Dordrecht, Kluwer, 1997, pp. 17–44.

—and, Dummett should have added, we are justified in supposing that some of those truth-conditions are met—

> but Frege was *completely mistaken* [my emphasis] about how we can go about fixing such truth-conditions.[27]

And again

> What must be disputed . . . is Frege's—and Wright's—idea of what is sufficient for determining the truth-conditions of sentences containing terms for newly introduced kinds. Impredicative specification of the conditions for the truth of identity statements involving one or two such terms is *not* sufficient, contrary to Frege's belief and to that of his disciple [*sic*!] Crispin Wright. It fails to fix truth-conditions for all sentences containing the new terms, when these terms are formed by attaching an operator to a predicate or functional expression; and it fails to do so because of the lack of an independent specification of the domain, which it attempts, but fails, to circumscribe simultaneously with its determination of the truth-conditions of sentences containing the new terms.[28]

The crux of Dummett's fourth objection, then, is the question: Does the necessary impredicativity of $N^=$—necessary if, as Frege's strategy for proving the infinity of the number series requires, it is to be possible to apply the cardinality operator to concepts under which numbers fall—does this impredicativity prevent $N^=$ from counting as an adequate explanation of the numerical terms whose use it is supposed to explain and legitimate?

IV. MOTIVATIONS FOR THE WORRY

Ever since the concern first surfaced in the wake of the paradoxes, discussion of the issues surrounding impredicativity—when, and under what assumptions, are what specific forms of impredicative characterizations and explanations acceptable—has been signally tangled and inconclusive. It seems to me that, despite his repeated attempts at formulation, it remains extremely difficult to fill out Dummett's specific concern here in convincing detail. The vicious circularity of a purported explanation ought to consist in its unavoidable presupposition, on the part of someone who is to follow it, of the very conceptual resources which following it is supposed to impart. By that standard, so I contend, $N^=$, though indeed ineliminably impredicative as the Fregean must intend it, is not viciously circular. I shall canvass two, as it seems to me, defective lines of thought which might suggest otherwise; the final section will then offer considerations which, if sound, point the way to a settlement in the Fregean's favour.

[27] *Frege: Philosophy of Mathematics*, 234. [28] Ibid. 236.

(i)

Some of Dummett's formulations—that just quoted, for example[29]—suggest a train of thought along the following lines:

> A statement involving quantifiers is determinate in content only if the range of quantification is fixed in advance. Hence it is only if we have fixed in advance what the range of the objectual quantifiers in fully explicit expansions of the right-hand sides of instances of $N^=$ is to be, that the left-hand-side statements, whose use that principle purportedly explains, can have any clear meaning. Yet for someone so far innocent of the concept of natural number, that domain can be satisfactorily characterized only if no natural numbers are included in it. It follows that $N^=$ can work as an explanation only if constrained to be predicative—and hence useless for Frege's purposes.

One response would be as follows. Suppose we grant that the determinacy in content of a quantified statement demands some prior specification of the range of its quantifiers. Even so, it is by no means evident that such a specification cannot be given in the kind of case which interests us without introducing a circle of presupposition. It depends on what *kind* of specification of the domain is supposed to be necessary. Someone innocent of the concept, *tiger*, for instance, may nevertheless possess the concept, *feline quadruped*; and for such a person there need, presumably, be no indeterminacy in the content of quantified statements whose range of quantification is understood to be precisely the class of feline quadrupeds. Analogously, the line of thought outlined could therefore presumably be met if we could disclose some more *inclusive* sortal concept, which someone might have grasped who did not yet possess the notion of cardinal number, and which could be used to specify the range of the quantifiers in $N^=$.

However, candidates do not thrust themselves forward. Frege himself—at the time of his innocence before Russell's Paradox—would presumably have thought that we do indeed possess various such concepts, in the notion of *extension* (class), for instance, or indeed in the general notion of an *object*. But if that is to be the response to the objection, then a defence is now owing of the idea that there is a determinate totality of *all classes*, or of *all the objects that there are*, which may intelligibly be stipulated as the range for certain uses of quantification.

I don't think that is the way for the Fregean to go. It cannot be a satisfying response to the outlined train of thought that it may always be answered by the citation of a determinate domain of sufficient generality—if only, at the limit, the domain of 'all objects'—if one has any feeling at all for the Russellian thought[30] that it is a deep misconception of certain very

[29] For additional relevant references to earlier writings of Dummett's, see Richard Cartwright's paper cited in n. 32 below.

[30] Russell's thought is taken up by Dummett in his notion of an 'indefinitely extensible'

general ranges of objects to think of them as constituting totalities—*domains* in the intended sense—at all.

What should immediately be said is, rather, that it is not and *cannot* in general be a prerequisite for a quantified statement to be determinate—or determinate enough—in content that the range of its quantifiers should have been specified in advance. The target of first-order predicate logic, for instance, is the codification of statement forms whose instances will hold true of *every* determinate non-empty domain, independently of its further detail. Such statements are clearly determinate in content if we are able to recognize that they have this feature; but the feature is precisely one that they possess *independently* of the detail of any specified domain of quantification—and their content accordingly cannot depend on the prior specification of an intended domain.

The point is not, however, or not merely, that the contrary assumption is wrong in particular cases. Rather, it is obscure how it is even as much as coherent. If it is only in a context in which the domain of quantification has been independently determined that quantified statements can have determinate content, the obvious question to ask is: What are the *means* whereby that domain may or may not have been determined? It's hardly credible that a range of items might be specified without the use of generality of any kind. But then the demanded prior specification will *itself* involve further quantifications.[31] So a vicious regress is generated—vicious because the determinacy of any quantified statement will now depend upon an infinity of prior specifications.

Two points, then, are salient. First, the contention that an understanding of quantified statements depends quite generally upon a prior specification of a domain of quantification is neither true nor coherent. Second, as the case of first-order logical laws brings out, some counter-examples to this incoherent contention are of *unrestricted generality*. There simply are intelligible, unrestrictedly general thoughts—thoughts which are therefore available to thinkers who are capable of simultaneously grasping the idea that some of the objects within the range of the former may, if only temporarily, lie outwith their classificatory repertoire. Doubtless we could do with further explanation of such generality: of what it is to grasp such an unrestrictedly general thought, of whether classical forms of quantificational reasoning continue to be appropriate in such cases, and so on. But the phenomenon is unquestionable. And we ought therefore to be open to the idea

totality. See 'The Philosophical Significance of Gödel's theorem', *Ratio*, 5 (1963), 140–55 (repr. in M. Dummett, *Truth and Other Enigmas*, London, Duckworth, 1978, 186–201). The notion looms large in the concluding chapter of *Frege: Philosophy of Mathematics*, and in Dummett's 'What is Mathematics About?' in his anthology, *The Seas of Language*, Oxford, Clarendon Press, 1993, 429–45.

[31] Cf. Hale, 'Dummett's Critique of Wright's Attempt to Resuscitate Frege', at 143.

that the unrestricted, hence impredicative generality that features on the right-hand sides of (instances of) $N^=$ is an example of it.[32]

(ii)

A sympathizer with Dummett may want to persevere:

Let it be that it is incoherent to think that it is only after we have first specified a determinate domain that a quantified statement has any definite meaning. Nevertheless, it is undeniable that the truth-conditions of a quantified statement depend upon the range of objects to which its author intends it to answer. What has to be the case in order for an utterance of the sentence, 'All householders were worse off under the Poll Tax' to be true, depends on the intended range of 'All householders': conditions sufficing for the truth of the sentence when that range is restricted to householders in Britain cease to do so when it is widened to include Continental Europeans. It should be recognized, accordingly, that the truth-conditions of statements on the right-hand sides of instances of $N^=$, and *hence the meanings thereby assigned to their left-hand-side equivalents*, vary as a function of the range of the individual quantifiers on the right-hand sides. That sets up a simple dilemma for the Fregean: if that range is taken to embrace the referents—cardinal numbers—of the type of expression supposedly being explained, then no one innocent of the concept of number is yet in position to understand what that range is, nor therefore to latch on to the meanings intended for the left-hand-side statements. If, on the other hand, the range is restricted to entities of which someone who knows nothing of numbers can have some concept, then the truth-conditions of statements on the right-hand sides, and hence the meanings assigned to their left-hand-side equivalents, will not be the ones that the Fregean has in mind; and $N^=$ will fail to convey the intended information. In short: either $N^=$ fails to be intelligible to the intended audience, or it fails to assign the intended meanings to the statements it is supposed to explain.

In one respect this line of thought gets us no further. For at the first horn of the dilemma posed, it too makes the suspect assumption that quantification whose range includes objects of a certain kind can be intelligible only to someone who is *au fait* with the concept of that kind of object. But there is an additional weakness which is worth highlighting: the casual movement between the notions of meaning and truth-conditions. It is harmless to think of the truth-conditions of a quantified statement as varying in tandem with variation in the range of quantification only provided the point is precisely not confused with one about *meaning*. The contribution of its range of

[32] A very useful recent discussion of unrestrictedly general quantification, which came to my attention only after drafting this paper is Richard Cartwright's 'Speaking of Everything', *Noûs*, 28 (1994), 1–20. Cartwright's discussion is focused specifically on Dummett's and makes a cogent case for disentangling the question of whether and what sense is made by unrestrictedly general quantification from the question whether there can be any single all-inclusive objectual *domain* (construed in the usual model-theoretic sense of that term).

quantification to the purport of a quantified statement is best compared to the contribution effected by context—the location and time of its utterance, identity of the speaker, and so on—to the purport of a statement involving *indexicals*. There is correspondingly no harm in describing the truth-conditions of a sentence like 'I was here yesterday' as varying in the mouths of different speakers; but there is nevertheless a clear sense in which its *meaning* does not so vary but makes a uniform contribution, in conjunction with the context of utterance, towards determining whether an utterer speaks truly. So too with quantified statements: 'All householders were worse off under the Poll Tax' should be seen not as varying in meaning from context to context, according as the range of quantification varies, but as imposing a *uniform condition* whose being satisfied or not will depend on what the range of quantification is taken to be.

Now, the germane point is that, while the truth-conditions of the left-hand sides of instances of $N^=$ will indeed vary in response to variation in the domain of the objectual quantifiers involved on the right-hand sides, the most basic explanatory brief which $N^=$ has to discharge is one of which it is actually no part to convey such truth-conditions. That brief—akin to that of explaining the meaning of a sentence like 'I was here yesterday' in such a way as to convey an understanding of what would be said by it on any particular occasion of its use for variable speaker, place, and time—is to explain what condition a given population of objects has to meet if a particular 'Nx:Fx = Nx:Gx' is to be true of that population—to impart a piece of information which in conjunction—but only in conjunction—with an understanding of some intended range of quantification, will enable the recipient to grasp the truth-conditions of a particular sentence of that kind. There is thus no merit in the suggestion that the only kind of understanding of the right-hand sides available to a novice is one which is unsuitable to facilitate the understanding of the left-hand sides which the Fregean will intend $N^=$ to bestow. What $N^=$ is supposed to bestow, in the first instance, is only, for each particular F and G, a knowledge of how things have to be in a given population if 'Nx:Fx = Nx:Gx' is to be true of it. Why may not that condition be grasped in full readiness, as it were, for populations which include numbers by someone who doesn't yet know what numbers are?

Call this the *range-unspecific* grasp of (an instance of) $N^=$. I have suggested that there is no evident obstacle to the capacity of $N^=$ to impart at least this. And now the view of the Fregean will be, in addition, that there is no obstacle to the intelligent fusion, as it were, of this range-unspecific understanding of $N^=$ with the kind of unrestrictedly general understanding of the objectual quantifiers on (instances of) its right-hand side emphasized above. It is all the same to the range-unspecific understanding of $N^=$ what the intended range of its objectual quantifiers is supposed to be. And it is all the same to the unrestrictedly general understanding of a quantified statement what actual, perhaps undreamed of types of object may turn out to fall within its scope. How is a problem of circularity generated by the union of these understandings? Why can they not be yoked together to produce exactly the state of information from which a recognition of the infinity of the series of natural numbers may follow, just as Frege believed?

A thought to consolidate that question. Quite early in *Grundlagen*, directly after the famous passage about Euclidean geometry, Frege writes:

> [W]e have only to try denying any one of [the fundamental propositions of the science of number] and complete confusion ensues. Even to think at all seems no longer possible. The basis of arithmetic lies deeper, it seems, than that of any of the empirical sciences, and even than of geometry. The truths of arithmetic govern all that is numerable. This is the widest domain of all; for to it belongs not only the actual, not only the intuitable, but everything thinkable. Should not the laws of number, then, be connected very intimately with the laws of thought?[33]

Frege's idea is that this absolute generality of number theory—its applicability wherever logic is applicable[34]—ought to predispose us to the idea that arithmetic is analytic in just the way that—he takes it—logic is. And Dummett for his part emphasizes the point in several places and seems disposed to regard it as one of Frege's abiding insights.[35] But whether or not the point suffices to establish a sense in which arithmetical thought is already logical, even before it comes to any logicist constructions, there seems no denying the *datum*: number theory does have this generality—*and must, therefore, be applicable in particular to its own proper objects*: the finite cardinal numbers.

The corollary is obvious. Since numerical concepts and operations are to be applicable to objects of any kind at all, any satisfactory explanation of them must encompass this generality and hence—if it has recourse to objectual quantification in the statements that make it up—must include the numbers themselves within its range. But that is to say that any stated explanation of the notion of cardinal number in general—assuming, as is surely inevitable, that it involves quantification at all—must either involve *impredicative* quantification or fail to encompass its target notion in full generality. The feature of $N^=$ which Dummett views as hopelessly compromising Fregean aspirations is thus a feature which any account of identity for cardinal numbers in the full generality which Frege, with Dummett's approval, attributes to that notion, has to have.

[33] *Grundlagen*, §14.

[34] The generality is not quite absolute, of course, but belongs with that of predicate logic. Schemes of thought are conceivable—for instance, as articulated in a tensed feature-placing language of the kind envisaged in chs. 6 and 7 of P. F. Strawson's *Individuals*, London, Methuen, 1959—which would naturally be subject to ordinary propositional logic but in which the idea of cardinal number would have no application.

[35] See *Frege: Philosophy of Mathematics*, 43–6, 224, and 308.

In sum: if—as Dummett thinks—Frege is right to ascribe this generality to our arithmetical notions, then it is a fact of our intellectual lives that the concept of (sameness of) cardinal number has a reflexivity which it will take recourse to impredicative quantification to capture. That an attempted account of the notion uses such quantifiers is therefore, not cause for complaint, but a necessary condition of its completeness. And the impression that vicious circularity is thereby imported must accordingly depend on some misconception or other.

V. UNDERSTANDING $N^=$

We have reviewed two unpersuasive lines of argument for the claim that the impredicativity of $N^=$ bars it from playing the stipulative-explanatory role that the neo-Fregean would assign to it, and suggested that the very generality of the concept of cardinal number must be in tension with the existence of any fundamental difficulty in this area of the kind Dummett supposes. Even if all this is correct, however, there remains on the neo-Fregean a clear obligation to offer something more positive: an account of how $N^=$ actually can work as an explanation—how someone innocent of the concept of cardinal number can use the principle to advance not merely to a general grasp of the condition any range of objects and pair of concepts, F and G, must meet if 'Nx:Fx = Nx:Gx' is to be true of them, but to an understanding of the specific uses of the cardinality operator which the neo-Fregean reconstruction of arithmetic demands.

As stressed earlier, this obligation actually arises independently of a concern about impredicativity. For as noted, $N^=$ is not a strict definition in the way the corresponding principle for identity of directions is (or can be). The latter can participate in a regimen of stipulations whose collective effect is to provide for the contextual eliminability of all singular reference to and quantification over directions.[36] $N^=$, by contrast, provides no general means for the contextual definition of occurrences of the cardinality operator: contexts formed by the deletion of one term from a statement of numerical identity and its replacement by a bound variable remain recalcitrant. For all that has so far been said, therefore, it is open to an awkward customer to claim that such occurrences of the operator are wholly unaccounted for. What is true as far as impredicativity is concerned is that *if* $N^=$ can indeed provide all that is necessary for an understanding of such uses, and in general for a construal of a Fregean reconstruction of sentences of number theory along the lines, for example, outlined in *Frege's Conception*, then—prescinding from the Caesar problem, which is another matter—

[36] See *Frege's Conception*, 29 f.

there can be no well-motivated doubt about its effectiveness as an explanation. And if its impredicativity puts in jeopardy neither its consistency nor its explanatory power, what is there left for a worry about impredicativity to be about?

In order to deliver what is needed—to show that $N^=$ provides the basis for an understanding of the kinds of use of the cardinality operator which Frege's reconstruction of arithmetic demands but which that principle does not eliminatively define—one would naturally think in terms of constructing a comprehensive ranking of such uses by degree of complexity and of trying to show how a grasp of the constructions prior to any particular stage in the ranking provides for the construal of those at that stage. I shall turn to that thought shortly. First, though, it will prove expedient to focus on one central, though as we shall see atypical component in the general task: that of explaining how $N^=$ provides the resources to construe Frege's definitions of zero and its successors.[37]

How should one start? The question is under what conditions someone can understand the right-hand side of (an instance of) $N^=$ who doesn't yet know what numbers are. No problem, to be sure, if numbers are excluded from the range of the individual quantifiers. But what we need, rather, is a base class of substitutions for its predicate variables for which, even when its individual quantifiers are construed impredicatively, $N^=$ can function as an explanation of the associated contexts of numerical identity without any presupposition of a prior understanding of arithmetical concepts. The vital initial reflection is therefore that there is indeed such a class of cases. In a large range of examples, an ability to verify whether F and G are equinumerous will not depend on knowing what numbers are or whether any of them are F or G (even if some of them are). Examples will include concepts like 'man born in Chipping Norton between the wars', 'Julio-Claudian emperor', 'apple on the table', or 'book on the shelf'. Even if the number 3 is in fact an apple (the one on top of the pile), someone who doesn't know that is in no way disadvantaged when it comes to determining whether there are exactly as many apples on the table as books on the shelf. Even if Julius Caesar is a number (the number zero), someone who doesn't know that is in no way disadvantaged when it comes to determining whether there were as many Julio-Claudian emperors as Tudor monarchs.

Say that F and G are *unproblematic* just in case the knowledge whether or not either or both are instantiated by some number is not needed by one who is to determine whether they can be one–one correlated. Then, when F and G are unproblematic, $N^=$ gives a perfectly clear, in no way circular,

[37] The argument to follow generalizes and completes the considerations advanced in *Frege's Conception*, §xvii, 132–45.

account of the truth-conditions of 'Nx:Fx = Nx:Gx', even when the objectual variables in the *explanans* are intended to include numbers in their range.

Now introduce a faultlessly rational subject—I'll follow Gareth Evans's example and call him Hero[38]—who has mastered an appropriate higher-order logic, in which the *definiens* of $N^{=}$ can be formulated, together with a range of such unproblematic concrete sortal concepts and a battery of singular terms standing for instances of them but who so far has no specifically arithmetical vocabulary. Then our task is to show how, on receiving $N^{=}$, and thereby coming to understand 'Nx:Fx = Nx:Gx' when F and G are drawn from his unproblematic stock, Hero somehow has all he needs in order to advance to an understanding of Frege's definitions of zero and its successors.

There are two aspects to this accomplishment. Hero has to understand the numerical predicates, like 'y = Nx:x ≠ x', which Frege's definitions exploit, and to understand the singular terms formed from them, like 'Ny:[y = Nx:x ≠ x]', given as *definientia* of the standard numerals. Both, however, will arguably be encompassed by an understanding of an appropriate range of identity statements. For reflect: since it ought to be reckoned sufficient for an understanding of any predicate that someone should understand any statement which results from completing it with any term that he understands (or comes to understand), it should suffice for an understanding of, for instance, the above numerical predicate that, for each term, 't', that he understands or comes to understand, Hero understands the corresponding 't = Nx:x ≠ x'. Likewise it ought to suffice for an understanding of the associated term, 'Ny:[y = Nx:x ≠ x]' that, where 't' is any term that he understands (or comes to understand), Hero should understand the corresponding identity statement, 't = Ny:[y = Nx:x ≠ x]'.

Let us take it that an understanding of any such identity-statement will be ensured by an appropriately generated knowledge of its truth-conditions. Then what needs to be shown is how Hero can determine the truth-conditions of each of the series of increasingly complex identity-statements which are associated as indicated with an understanding of the numerical predicates which Frege's definitions of the numerals successively utilize and of the numerical terms which successively result. More specifically, designate each predicate, 'Ax', for which 'Nx:Ax' is Frege's *definiens* for the number, n, *Frege's predicate for n*; and designate 'Nx:Ax', likewise, as *Frege's term for n*. Then for each 't' which he understands or comes to understand, Hero needs to be able to determine the truth-conditions of each of the infinite series of statements,

[38] Evans so names his imaginary subject of experience in 'Things Without the Mind', in Zak van Straaten (ed.), *Philosophical Subjects: Essays Presented to P. F. Strawson*, Oxford, Clarendon Press, 1980, 76–116.

$$t = Nx{:}x \neq x,$$

whereby he understands Frege's term for 0 and predicate for 1;

$$t = Ny{:}\,[y = Nx{:}x \neq x],$$

whereby he understands Frege's term for 1;

$$t = Nx{:}x \neq x \ \vee\ t = Ny{:}\,[y = Nx{:}x \neq x],$$

whereby he understands Frege's predicate for 2;

$$t = Nz{:}\,[z = Nx{:}x \neq x \ \vee\ z = Ny{:}\,[y = Nx{:}x \neq x]],$$

whereby he understands Frege's term for 2;

$$t = Nx{:}x \neq x \ \vee\ t = Ny{:}\,[y = Nx{:}x \neq x] \ \vee$$
$$t = Nz{:}\,[z = Nx{:}x \neq x \ \vee\ z = Ny{:}\,[y = Nx{:}x \neq x]],$$

whereby he understands Frege's predicate for 3;

$$t = Nw{:}\,[w = Nx{:}x \neq x \ \vee\ w = Ny{:}\,[y = Nx{:}x \neq x] \ \vee$$
$$w = Nz{:}\,[z = Nx{:}x \neq x \ \vee\ z = Ny{:}\,[y = Nx{:}x \neq x]]],$$

whereby he understands Frege's term for 3;
. . . and so on.

Note now that every oddth member of that series from the third onwards is a disjunction which Hero will understand if he understands its constituent disjuncts—since he has (higher-order) logic. And each constituent disjunct of such a disjunction occurs in its own right earlier in the series. So the task reduces to that of showing how Hero is in position to know the truth-conditions of each identity-statement of the form,

$$t = n_f,$$

where 'n_f' is the Fregean *definiens* of the numeral for n, and 't' is either an expression of the form 'Nx:Fx' which he already understands (or comes to understand) or else a concrete singular term from his initial repertoire.

Take the latter case first; suppose that 't' is one of the initial concrete singular terms, or indeed any other term which Hero comes to understand as denoting, if anything, then a concrete entity of some kind. Then he will be able to know the truth-conditions—and indeed the truth-value—of '$t = n_f$' provided, and so far as I can see, only provided he has a general solution to the Caesar problem: some principled account of the distinction between numbers and things of other kinds which allows him to determine that no concrete object can be a number.

The Caesar problem is often viewed as the ineluctable nemesis of any form of mathematical Platonism. Certainly it needs an extended treatment, not to be ventured here. But I should at least explain my own continuing feelings of optimism about it. Is it really plausible that Hero is in no position to make *anything* of the question whether any particular object of a kind previously familiar to him coincides with an object denoted by one of the numerical terms which $N^=$ introduces? There is one thought which surely should be salient to him. Admitted that nothing directly follows from the actual statement of $N^=$ about the external identity, so to speak, of the objects of which it speaks, a constraint *is* exerted by the fact that the principle is offered as the explanation of a *kind*: a determination of the sort of thing that numbers, *qua* numbers, are. For one who receives $N^=$ in this way, it will be understood to be of the *essence* of cardinal numbers that facts about identity and distinctness among them are constituted in facts about one–one correspondence among concepts.[39] So nothing can be a number whose every distinction from other numbers is not constituted in such facts. Since this is not true of persons, or stones, or concrete objects generally—since identity and distinctness among such things is grounded in facts of quite different kinds—Hero has good grounds for the thought—with which of course, as Frege noted, we sympathize[40] (and what source of information do we have that Hero lacks?)—that there is no co-reference between any of his initial stock of terms and any of the numerical singular terms introducible via $N^=$.

This is, to stress, merely to make a suggestion about how the general distinction which, seemingly, it comes quite naturally to us to recognize, between the abstracts associated with principles like $N^=$ and many ordinary kinds of object, is grounded in the concepts concerned. It is not a suggestion which will cover all cases; in particular, it will not cover the relations between Fregean abstracts and *classes* generally since it *is*, plausibly, of the essence of equivalence classes under a given relation that identities and distinctions among them are grounded in facts about the obtaining of that relation. Here, though, it seems to me, it is plausible to regard the indeterminacy as real: Frege's identification of cardinal numbers with equivalence classes of concepts under one–one correspondence, for instance, strikes one merely

[39] This aspect of abstraction principles, conceived as forms of conceptual innovation, goes unrecognized in Dummett's discussion in *Frege: Philosophy of Mathematics*; see e.g. the criticism offered at 161–2 of the *Frege's Conception* suggestion about the Caesar problem (which more elaborate formulation was essentially an attempt to develop the idea presently entertained) and Dummett's suggestion at 126 of the identifiability of directions and lines-through-a-fixed-point. For discussion of Dummett's misinterpretation of the point of the *Frege's Conception* proposal, and Hale's own related proposals, see Hale's 'Dummett's Critique of Wright's Attempt to Resuscitate Frege', 122–47.

[40] In *Grundlagen* §66 when the Caesar problem is first brought up in connection with the abstraction for directions, Frege makes a remark which, transposed to the case of $N^=$, would run as follows: 'Naturally no one is going to confuse [Caesar] with [the number 0]; but that is no thanks to our definition of [number].' A sceptic about the solubility of the Caesar problem had better feel that, to the contrary, such a confusion is precisely what is invited; otherwise, there has to be something which discourages it and which has somehow been implanted in our understanding, so that the Caesar problem is less that of making a distinction than that of formulating the means whereby it is already drawn.

as an available move—one which nothing intuitive either mandates or forbids. If that is right, then its neutrality on such examples is actually to the credit of the mooted proposal: if there *is* indeterminacy in some instances, a correct account of what determines the determinate ones will not say anything about those.

The other class of cases to consider are identity-statements of the form, '$t = n_f$', where 'n_f' is the Fregean *definiens* of the numeral for n, and 't' is an expression of the form 'Nx:Fx' which Hero already understands. We may take it that it will suffice for a knowledge of the truth-conditions of these statements that Hero should know, for each particular 'n_f', *how many* F's there have to be if '$Nx{:}Fx = n_f$' is to be true; and that this knowledge will certainly be captured by knowledge, for each 'n_f', of a truth of the form:

$$Nx{:}Fx = n_f \leftrightarrow (\exists_n x)\,Fx,$$

where '$(\exists_n x) \ldots x$' expresses the numerical quantifier, 'There are exactly n . . .' as standardly recursively defined.[41] That definition can, of course, be given in fully explicit form in higher-order logic. It therefore suffices for the immediate purpose to reflect that, with the numerical quantifiers so defined, each instance of the above schema may be derived from $N^=$ in higher-order logic and is therefore available to Hero.[42]

Our result, then, is that $N^=$ provides the resources for a construal of any identity-statement linking one of Frege's numerals with an antecedently understood term, 't', in two cases: where 't' is understood to stand for a type of object which cannot be a number if the essentialist proposal about the Caesar problem outlined above is right; and where 't' is (an abbreviation of something) of the form, 'Nx:Fx'. These cases are not exhaustive; for instance, they exclude terms for classes. Nevertheless the result is enough, I argued, for the conclusion that the principle explains the meaning of Frege's numerals.

There are two respects in which this finding falls short of the generality that must ultimately be accomplished. First, nothing explicit has so far been said about how Hero is to construe the results of binding the free variable in a numerical predicate not by the cardinality operator but by the ordinary first-order quantifiers. Rectifying this omission raises no difficulty. If a predicate is agreed to be determinate in content, then the contents of its quantifications may presumably be conceived as determined in line with whatever semantics for the quantifiers, classical or constructivist, is found agreeable. But a predicate ought to be regarded as determinate in content whose every completion by a term whose meaning is understood is likewise determinate in content. So it must suffice to provide meanings for the quan-

[41] For details, see the Appendix. (To the best of my knowledge, the relevant clauses were first formulated in *Grundlagen* §55.)

[42] A proof is sketched in the Appendix.

tifications of numerical predicates that each of their instantiations to a term which is previously understood, or which comes to be understood, is determinate in content. Ultimately, therefore, provided we have under semantic control all the terms which Hero's basic vocabulary allows us to form, quantification will look after itself.

The second omission is another matter. Hero's initial vocabulary comprises higher-order logical vocabulary, an indefinite stock of concrete singular terms, a range of unproblematic predicates, and the cardinality operator as governed by $N^{=}$. Many other numerical predicates, both pure and mixed,[43] besides Frege's predicates for the natural numbers will, of course, be formulable in terms of these materials. And each such predicate will be attachable to the cardinality operator to form a term whose contribution to the meaning of identity-statements containing it $N^{=}$ will somehow have to be harnessed to explain. But obviously it cannot in general be expected that, for each such predicate, Bx, it will be possible for Hero to prove a result of the form,

$$Nx{:}Fx = Nx{:}Bx \leftrightarrow (\exists_n x)Fx.$$

For instance, since Hero's base language is adequate for the formulation of number theory, we know we can construct a predicate, θx, within it meaning: is a member of a pair of twin primes. Since there is no surety that anyone can know how many twin primes there are, we are in no position to expect that Hero will be able to recover the truth-conditions of contexts of the form,

$$Nx{:}Fx = Nx{:}\theta x,$$

by the kind of method which works for contexts involving Frege's terms for the natural numbers.

The *general* form, then, of the means whereby $N^{=}$ fixes the meaning of numerical singular terms cannot be that outlined for the Fregean numerals, which are something of a special case. And the general account the Fregean needs, of how Hero is empowered to construe *any* numerical predicate, cannot be based on the principle that it will suffice to be able to associate a numerical quantifier with it in the fashion illustrated for Frege's predicates for the natural numbers. So how to proceed?

Well, let us review the prospects for an induction on the degree of complexity of expressions in Hero's language, as earlier anticipated. Let a *numerical term* be any term of the form, 'Nx:Ax', where 'Ax' is a predicate expressible in Hero's language. If 'Nx:Ax' and 'Nx:Bx' are a pair of such terms, let 'Nx:Ax = Nx:Bx' be their *corresponding identity-statement*. Let S

[43] These notions to be understood in the obvious way: pure numerical predicates are any formulable using just the N-operator and the resources of Hero's logic, while mixed numerical predicates deploy additional non-logical constituents.

be any such identity-statement formulable in Hero's language. Define the *rank* of S as follows:

(i) S is of rank 0 if 'Ax' and 'Bx' are each unproblematic;
(ii) S is of rank n + 1 if there are numerical terms occurring within 'Ax' and/or 'Bx' whose corresponding identity-statement is of rank n, and for no such terms is the corresponding identity-statement of rank higher than n.

Now extend the idea of rank first to numerical terms, and then to statements in general, in the obvious way:

(i) A term 'Nx:Ax' is of rank n just in case 'Nx:Ax = Nx:Ax' is of rank n.
(ii) An arbitrary sentence, S, of Hero's language is of rank n just in case some numerical term occurring within S is of rank n and no numerical term occurring within S is of rank higher than n.

Hero understands all numerical identity-statements of rank 0. Suppose he also understands all numerical identity-statements of rank n: Does it follow that he understands all of rank n+1?

Clearly that will follow if an understanding of all numerical identity-statements of rank n ensures an understanding of all statements (in Hero's language) *of any kind* of rank n. For an understanding of all statements of any kind of rank n will ensure the comprehensibility of all *right-hand sides* of $N^=$ of rank n; and that in turn will suffice for an understanding of the truth-conditions of every numerical identity-statement of rank n+1. But given that the syntax and non-numerical constituents of any statement of rank n are familiar to Hero, the point ought to carry on grounds of mere compositionality that he will understand any statement of rank n provided he understands all *numerical terms* of rank n. So our question—the crux of the whole issue—is precisely whether an understanding of all *numerical identity-statements* of rank n ensures an understanding of all *numerical terms* of rank n. If so, Hero will be able to construe all numerical identity-statements of whatever rank and indeed all arithmetical contexts, pure and applied, whose formulation his language allows.

It might seem that the point does not require very much argument. Surely, if someone understands an identity-statement, he is bound to understand the ingredient terms. But that of course is too quick in the present context. To know the truth-conditions of a numerical identity-statement, as determined via $N^=$, need not be—it might seem—to be empowered to conceive it full-bloodedly *as* an identity-statement. Such a full-blooded conception will demand that one's understanding of the ingredient terms comply with certain platitudes about what it is to understand any singular term—or more specifically, in view of the semantic complexity of numerical terms, any *descriptive* singular term. It cannot be assumed without further ado that an understanding of the right-hand side of an instance of $N^=$, coupled with the instruction that one is to treat the left-hand side in all respects as if it were a statement of identity, somehow ensures that the latter is genuinely so conceived.

So, what are the germane platitudes? They are two, it seems to me. To understand a descriptive term, it is arguably necessary—and will certainly suffice, which is what matters for our purpose—to know, first, what *kind* of thing, if anything, it stands for; and second, what would qualify a thing of that kind to *be* the thing it stands for. Let 'Nx:Ax', then, be any numerical term of rank n. Does Hero have an adequate conception of the kind of thing it may stand for? And does he, if so, have a grasp of the specific condition which a thing of that kind must meet in order to qualify as its referent?

'Nx:Ax' is apt to refer, if to anything, then to a cardinal number. So Hero has the requisite general conception if he knows what cardinal numbers are—which he does if he understands that cardinal numbers are individuated by facts about one–one correspondence among concepts and is apprised of a solution of the Caesar problem to whatever extent is necessary in order to have a conception of the difference between numbers and other kinds of thing. So this issue breaks no new ground: Hero's understanding complies with the first platitude provided the Caesar problem does indeed admit of solution.

What of the second: Does Hero have a grasp of the specific condition which a number should meet in order to qualify as the referent of 'Nx:Ax'? By hypothesis, he understands all numerical identity-statements of rank n, and so understands 'Ax' (since it occurs in the *definientia*, via $N^=$, of a range of numerical identity-statements of that rank which contain 'Nx:Ax' as one related term). But—as an upshot of our lemma concerning the Fregean numerals—Hero also knows, for each finite number, n_f, that for that number to be the referent of 'Nx:Ax', it is necessary and sufficient that there be exactly n A's—a condition, to stress, which he can formulate and appreciate. So he does indeed grasp, for each finite number, what condition it must meet in order to be the referent of 'Nx:Ax'.

True, this condition varies as a function of which number we consider; but that merely betokens that the descriptive content of 'Nx:Ax' incorporates a certain, quite common type of reflexivity—and is easily given a uniform statement once explicit provision is made for that. The condition for a number to be the referent of 'Nx:Ax' is that 'Ax' should instantiate the numerical quantifier associated with *that number*. (Compare: 'the MP with the largest absolute majority in the House of Commons': in order to qualify as the referent, a member must be he/she for whom the number of votes cast in *his/her* constituency minus the total number cast in *that* constituency for any other candidate is greater than the corresponding number for any other MP.) True, also: the understanding described amounts at best to a grasp of what it is for a *finite* number to be the referent of 'Nx:Ax'. But—assuming it clear that more is required—corresponding connections will in any case presumably hold between whatever infinite cardinals Hero is in position to conceive and appropriately specialized higher-order definable numerical quantifiers ('There are countably infinitely many . . .', 'There are uncountably infinitely many . . .', and so on).

Once the lemma concerning the special case of the Fregean numerals is in place, then—and assuming, as always, the means to bury Caesar—really quite simple considerations suffice to supply the generalization which was owed. The *proof-theoretic* unification which $N^{=}$ supplies for the fundamental laws of arithmetic is accordingly matched by an *explanatory* unification: to have understood the principle as an explanation of the concept of cardinal number is to be apprised of a complete explanation of all possible uses of the N-operator within a higher-order language of Hero's kind[44]—which is to have a basis on which all the distinctive concepts of number theory can be defined.

It follows that nothing can be essentially involved in the epistemology of number theory that is not involved in an understanding, and knowledge of the truth of 'Hume's Principle'. The neo-Fregean view is that such knowledge may be had, for relative cheap, after the fashion of knowledge of *any* bona fide Fregean abstraction—that such principles may properly be viewed as keeping company with definitions and other freely undertaken forms of concept-determination. That comparison should continue to be the focus of the debate about the neo-Fregean project. My point here has been that, whatever else may or may not prove to be wrong with it, it is certainly not spoiled at the outset by the impredicativity in which Dummett sees the doom of Frege's enterprise.

APPENDIX: ON THE PRINCIPLE, N^q

The schematic principle,

$$Nx{:}Fx = n \leftrightarrow (\exists_n x)Fx$$

is christened N^q in Bob Hale's *Abstract Objects*,[45] and is there invoked to scotch the charge of Harold Hodes[46] that we are quite at liberty to interpret the N-operator as denoting *any* one–one function to natural numbers from equivalence classes of concepts under one–one correspondence. Apart from the use to which I have put it

[44] Michael Dummett quite correctly observes (this volume, p. 383) that certain cases are not covered by the foregoing considerations. The remaining details are outlined in my response to him (this volume, pp. 399–402).

[45] Hale, *Abstract Objects*, 223.

[46] Hodes, 'Logicism and the Ontological Commitments of Arithmetic', 134–5.

above, in grounding the interpretation of the Fregean numerals, the principle is important both for its application to Hodes's argument, and because of the role it plays in explaining how Frege's account of cardinal number illuminates the applications of arithmetic. Hale, however, presents the principle as something that we might use to *supplement* $N^=$ in a response to Hodes's challenge. That it is no supplement may be seen as follows:

Stage-setting

Assume the standard recursive definitions of the numerically definite quantifiers:

$$(\exists_0 x)Fx \leftrightarrow (\forall x)-Fx$$

$$(\exists_{n+1} x)Fx \leftrightarrow (\exists x)(Fx \ \& \ (\exists_n y)(Fy \ \& \ y \neq x))$$

and let 'n_f', as before, abbreviate Frege's *definiens* for n. Define 'Pxy'—immediate predecession—as:

$$(\exists F)(\exists w)(Fw \ \& \ y = Nv{:}Fv \ \& \ x = Nz{:}[Fz \ \& \ z \neq w]).$$

Define 'Nat(x)'—x is a natural number—as:

$$x = 0 \vee P^*0x,$$

where 'P^*xy' expresses ancestral predecession.
Let '$(\exists R)(F\ 1\text{–}1_R\ G)$' express that there is a one–one correspondence between F and G.

We shall use three main lemmas from the proof of the Peano axioms from $N^=$ outlined in the concluding section of *Frege's Conception* (numbering as there assigned):

Lemma 51: $(\forall x)(Nat(x) \rightarrow x = Ny{:}[Nat(y) \ \& \ P^*yx]$,
—every natural number is the number of its ancestral predecessors.
Lemma 52: $(\forall x)(Nat(x)\rightarrow -P^*xx)$,
—no natural number ancestrally precedes itself.
Lemma 5121: $(\forall x)(\forall y)(Nat(x) \ \& \ Nat(y) \rightarrow (Pxy \rightarrow (\forall z)(Nat(z) \ \& \ (P^*zx \vee z{=}x) \leftrightarrow (Nat(z) \ \& \ P^*zy)))$
—if one natural number immediately precedes another, then the natural numbers which ancestrally precede the second are precisely the first and those which ancestrally precede the first.

Finally, recall that Frege's 0 is $Nx{:}x \neq x$ and that each successive $(n+1)_f$ is $Nx{:}[x = 0 \vee \ldots \vee x = n_f]$. Each of these objects qualifies as a natural number in the light of the above definition of 'Nat(x)'. *Proof*: 0_f qualifies by stipulation; $(n+1)_f$ qualifies if n_f does—take 'F' in the definition of 'Pxy' as '$[x=0 \vee \ldots \vee x=n_f]$' and 'w' as '$n_f$' to show that $P(n_f, n+1_f)$; then reflect that $Pxy \rightarrow P^*xy$ and that P^*xy is transitive. (*Frege's Conception*, Lemmas 3 and 4 respectively.)

Proof of N^q for Frege's natural numbers

Base: To show

$$Nx{:}Fx = 0_f \leftrightarrow (\exists_0 x)Fx,$$

it suffices to reflect that the lhs holds just if $(\exists R)(Fx\ 1\text{–}1_R\ x \neq x)$, which in turn holds just if $-(\exists x)Fx$.[47]

Hypothesis:
Suppose $Nx{:}\ Fx = n_f \leftrightarrow (\exists_n x)Fx$.
We need to show that it follows that

$$Nx{:}Fx = (n+1)_f \leftrightarrow (\exists_{n+1}x)Fx.$$

PROOF

Left-to-right:
Consider any F such that $Nx{:}\ Fx = (n+1)_f$. By lemma 51 and the reflection that $Nat(n_f)$, $n_f = Nx{:}\,[Nat(x)\ \&\ P^*xn_f]$. So by the Hypothesis, $(\exists_n x)(Nat(x)\ \&\ P^*xn_f)$. But by lemma 52, $-P^*n_f, n_f$. So $(\exists_n x)(Nat(x)\ \&\ (P^*xn_f \vee x{=}n_f)\ \&\ x{\neq}n_f)$. So $(\exists y)(Nat(y)\ \&\ (P^*yn_f \vee y{=}n_f)\ \&\ (\exists_n x)(Nat(x)\ \&\ (P^*xn_f \vee x{=}n_f)\ \&\ x{\neq}y))$. So by the recursion for the quantifiers, $(\exists_{n+1}x)(Nat(x)\ \&\ (P^*xn_f \vee x{=}n_f)$. But by lemma 5121 and since $P(n_f, n{+}1_f)$, we have that $(\forall x)(Nat(x)\ \&\ (P^*xn_f \vee x{=}n_f) \leftrightarrow Nat(x)\ \&\ P^*(x, n{+}1_f))$. So $(\exists_{n+1}x)(Nat(x)\ \&\ P^*(x, n{+}1_f))$.

That establishes the desired result for one concept of which $(n + 1)_f$ is the number. But by $N^=$, any G such that $(n + 1)_f = Nx{:}Gx$ will admit a one–one correspondence with that concept. So a lemma to the following effect will now suffice:

$$(\forall F)(\forall G)(\exists R)(F\ 1\text{–}1_R\ G) \rightarrow ((\exists_{n+1}x)Fx \leftrightarrow (\exists_{n+1}x)Gx)$$

A proof by induction—strictly, at third order—suggests itself:

Base:
It suffices to show $(\forall F)(\forall G)((\exists R)(F\ 1\text{–}1_R\ G) \rightarrow ((\forall x) - Fx \leftrightarrow (\forall x) - Gx)$.

Hypothesis:
Suppose $(\forall F)(\forall G)((\exists R)(F\ 1\text{–}1_R\ G) \rightarrow ((\exists_n x)Fx \leftrightarrow (\exists_n x)Gx)$.
Consider any H such that $(\exists_{n+1}x)Hx$. Then $(\exists x)(Hx\ \&\ (\exists_n y)(Hy\ \&\ y{\neq}x))$. Let a be such that $Ha\ \&\ (\exists_n y)(Hy\ \&\ y{\neq}a)$. Let J be one–one correlated with H by R. Let b be such that Jb & Rab. Then R one–one correlates $Hx\ \&\ x{\neq}a$ with $Jx\ \&\ x{\neq}b$. So, by the Hypothesis, $(\exists_n x)(Jx\ \&\ x{\neq}b)$. So $(\exists x)(Jx\ \&\ (\exists_n x)(Jx\ \&\ x{\neq}b))$. So $(\exists_{n+1}x)Jx$.

Right-to-left:
Consider any F such that $(\exists_{n+1}x)(Fx)$. Then there is some a such that $Fa\ \&\ (\exists_n y)(Fy\ \&\ y{\neq}a)$. So by the Hypothesis $Ny(Fy\ \&\ y{\neq}a) = n_f$. So, by $N^=$, there is an R such that $(Fy\ \&\ y{\neq}a)\ 1\text{–}1_R\ (Nat(x)\ \&\ P^*xn_f)$. Let $R^{\#}$ correlate $(Fy\ \&\ y{\neq}a)$ with $(Nat(x)\ \&\ P^*xn_f)$ in just the fashion of R, and let it also correlate a with n_f. Then $(Fy)\ 1\text{–}1_{R^{\#}}\ (Nat(x)\ \&\ (P^*xn_f \vee x{=}n_f))$. But, as established above, $(\forall x)(Nat(x)\ \&\ (P^*xn_f \vee x{=}n_f) \leftrightarrow Nat(x)\ \&\ P^*(x, n{+}1_f))$. So $Nx{:}Fx = (n{+}1)_f$.

[47] As George Boolos has reminded me, Frege himself observes, at *Grundlagen* §75 and §78, that he is in a position to obtain proofs of N^q for 0_f and 1_f respectively.