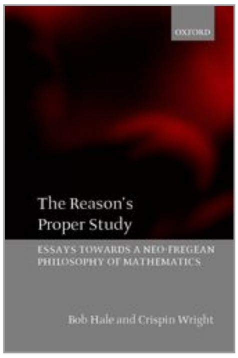


The Reason's Proper Study: Essays towards a Neo-Fregean Philosophy of Mathematics
Bob Hale and Crispin Wright

Print publication date: 2001
Print ISBN-13: 9780198236399
Published to Oxford Scholarship Online: November 2003
DOI: 10.1093/0198236395.001.0001

# To Bury Caesar . . .

Bob Hale (Contributor Webpage)
Crispin Wright (Contributor Webpage)

DOI:10.1093/0198236395.003.0015

## Abstract and Keywords

In this paper, Bob Hale and Crispin Wright discuss the Caesar Problem, which raises the worry that abstraction principles––like the Direction Equivalence and Hume's Principle ––are subject to an indeterminacy of reference which leaves them incapable of excluding the possibility that Caesar is the Number 3. They review––and deem implausible––an attempt by Richard Heck to finesse the Caesar Problem by recasting Hume's Principle in a two-sorted language. Drawing on considerations by Michael Dummett (*Frege: Philosophy of Mathematics*, 1991) and Gideon Rosen (”Refutation of Nominalism(?)”, 1993), they then address the question whether a coherent nominalist response to Fregean abstraction––which is supposed to involve a distinctively realist ontology––can be provided. A section is devoted to a discussion of two criticisms––launched by respectively Dummett and Potter and Sullivan––of Wright's solution to the Caesar Problem in *Frege's Conception of Numbers as Objects* (1983), based on an inclusion principle for sortal concepts. While Hale and Wright disagree with Dummett, they find that a consideration close to the Potter-Sullivan line of thought shows that Wright's earlier proposal stands in need of modification––and provide a modified solution centred around the notions of sortal concept, category, and criterion of identity.

*Keywords:* Caesar Problem, category, criterion of identity, Dummett, Heck, Potter, Rosen, sortal, sortal inclusion principle, Sullivan

Contents



## 1. Frege's Presentation of the Problem in Grundlagen

Frege first raises what has come to be known as the Julius Caesar Problem in *Grundlagen* § 56, in criticizing the proposal, made in the immediately preceding section, to define ‘the number *n* belongs to the concept F' inductively, by stipulating that:

> the number 0 belongs to the concept F iff $\forall x \neg Fx$
> the number 1 belongs to the concept F iff $\exists xFx \& \forall x \forall y (Fx \& Fy \rightarrow x = y)$
> the number $n + 1$ belongs to the concept F iff $\exists x(Fx$ & the number *n* belongs to the concept $Fy \& y \neq x)$

It is, he says, the last clause which is most likely to cause misgivings, because

> strictly speaking we do not know the sense of the expression 'the number *n* belongs to the concept G' any more than we do that of the expression 'the number ($n$ + 1) belongs to the concept F'. We can, of course, by using the last two definitions together, say what is meant by
>
> 'the number 1 + 1 belongs to the concept F'
>
> and then, using this, give the sense of the expression
>
> 'the number 1 + 1 + 1 belongs to the concept F'
>
> **(p.336)** and so on; but we can never—to take a crude example—decide by means of our definitions whether any concept has the number JULIUS CAESAR belonging to it, or whether that same familiar conqueror of Gaul is a number or not.[1]

He then advances an additional objection:

> Moreover we cannot by the aid of our suggested definitions prove that, if the number *a* belongs to the concept F and the number *b* belongs to the same concept, then necessarily $a = b$. Thus we should be unable to justify the expression '*the* number which belongs to the concept F', and therefore should find it impossible to prove a numerical identity, since we should be quite unable to achieve a determinate number.

The additional objection is intimately related to, and helps us to see the point of, the Julius Caesar objection (of which it is, in effect, a corollary). Were we able to transform the hypothesis that the number *a* belongs to the concept F and the number *b* belongs to the concept F into the conjunction of identities: the number belonging to the concept F = *a* and the number belonging to the concept F = *b*, then the desired conclusion would follow immediately by Leibniz's Law. Frege's point accordingly has to be precisely that we are *not* justified in so transforming the hypothesis. And since we would be justified in doing so, had his definitions established a use for expressions of the form ‘the number that belongs to the concept F’ as Fregean *proper names* (singular terms), the point of the first, Julius Caesar objection has to be that they fail to do that. This is confirmed by the remark with which Frege concludes his criticism of the proposed inductive definitions:

> It is only an illusion that we have defined 0 and 1; in reality we have only fixed the sense of the phrases
>
> ‘the number 0 belongs to’
> ‘the number 1 belongs to’;
>
> but we have no authority to pick out the 0 and 1 here as self-subsistent objects that can be recognized as the same again.

Frege does not, as one might perhaps expect him to do, immediately scrap the idea of defining number contextually. The moral he at first draws is rather that this first attempt at a contextual definition fails as a result of fastening on to the wrong sort of (sentential) context. Thus in *Grundlagen* § 62, after reaffirming his Context Principle (‘it is only in the context of a proposition that words have any meaning’), he asserts that—since number words are to be understood as standing for self-subsistent objects (as proper names)—the crucial sentential contexts whose sense must be fixed are identity statements linking terms for numbers. What must be done is to

> define the sense of the proposition
>
> ‘the number which belongs to the concept F is the same as that which belongs to the concept G’;
>
> **(p.337)** that is to say, we must reproduce the content of this proposition in other terms, avoiding the use of the expression
>
> ‘the number which belongs to the concept F’.

> In doing this, we shall be giving a general criterion for the identity of numbers. When we have thus acquired a means of arriving at a determinate number and of recognizing it again as the same, we can assign it a number word as its proper name.[2]

Frege next—without at this point explicitly formulating it—proceeds to consider the suggestion that numerical terms may be contextually defined by means of what is now commonly called *Hume's Principle*, determining that the number of Fs is the same as the number of Gs just in case there is a one-to-one correlation between the Fs and Gs; formally[3]

$$\forall F \forall G\{NxFx = NxGx \leftrightarrow \exists R[\forall x \forall y \forall z ((Rxy \ \& \ Rxz \rightarrow y = z) \\ \& (Rxz \ \& \ Ryz \rightarrow x = y)) \& \forall x(Fx \rightarrow \exists y(Gy \ \& \ Rxy)) \\ \& \forall y(Gy \rightarrow \exists x(Fx \ \& \ Rxy))]\}.$$

He immediately raises a doubt—to the effect that it is a mistake to define identity specially for the case of numbers—but dismisses it with the reply that he is doing no such thing: he is, rather, taking the general concept of identity as already understood, and using that prior understanding ‘as a means for arriving at that which is to be regarded as being identical’. The doubt, therefore, betrays a crucial misunderstanding of how his explanation is intended to work. The remainder of his discussion is largely conducted at one remove, in terms of what he takes to be a relevantly analogous contextual definition of terms for directions by means of the *Direction Equivalence*:

> The direction of line *a* = the direction of line *b* iff lines *a* and *b* are parallel.

Against this, he raises two further doubts or objections. The second doubt—which, he argues, may be easily laid to rest—is that we may, by introducing direction terms in this way, come into conflict with the well-known laws of identity.[4] But his third doubt, he thinks, cannot; he writes

> In the proposition
>
> > ‘the direction of *a* is identical with the direction of *b*’

> **(p.338)** the direction of *a* plays the part of an object, and our definition affords us a means of recognizing this object as the same again, in case it should happen to crop up in some other guise, say as the direction of *b*. But this means does not provide for all cases. It will not, for instance, decide for us whether England is the same as the direction of the Earth's axis—if I may be forgiven an example which looks nonsensical. Naturally no one is going to confuse England with the direction of the Earth's axis; but that is no thanks to our definition of direction. That says nothing as to whether the proposition
>
> 'the direction of *a* is identical with *q*'
>
> should be affirmed or denied, except for the one case where *q* is given in the form of 'the direction of *b*'. What we lack is the concept of direction; for if we had that, then we could lay it down that if *q* is not a direction, our proposition is to be denied, while if it is a direction, our original definition will decide whether it is to be denied or affirmed.[5]

This is plainly a re-run of the Caesar Problem. Frege rapidly convinces himself that it cannot be solved without abandoning the attempt at a contextual explanation altogether. In *Grundlagen* § 68 he does just that, switching to an explicit definition of numbers in terms of extensions of concepts (roughly, sets). The disastrous consequences which attended the course on which he thereby embarked are too well known to require detailed elaboration. If cardinal numbers are to be identified, in Frege's way, with extensions, an axiom governing extensions is required. Frege's own proposal, set forth in *Grundgesetze*, i, was his Basic Law V:

(V)

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

which leads, as is notorious, to a version of Russell's antinomy. It is true, of course, that we know how to frame better axioms for a theory of sets (e.g. the usual ZFC axioms) which are widely believed to be consistent. But that is no help to Frege. The character of the standard set-theoretic axioms makes it utterly implausible to claim that they express general logical laws; and in any case, Frege's explicit definition of the cardinal numbers cannot be consistently adopted in the context of standard set theory.[6]

**(p.339)** In a footnote to his explicit definition (*Grundlagen*: 80, fn.1), Frege says that he assumes it known what the extension of a concept is. He does not say, but must have thought that we know enough about extensions to know that Julius Caesar is not one—or else it would be a mystery how he could have thought that the switch to the explicit definition was any progress. But if, prescinding from its inconsistency, we view Basic Law V as intended to be explanatory of extension terms in much the way that Hume's Principle was supposedly explanatory of numerical terms, then it is clear that there is actually no progress at all—Basic Law V is merely a further, albeit very general principle of the same species[7] as Hume's Principle; if the latter really is incompetent to determine the truth-value of mixed identities linking person-denoting terms with terms for numbers, the former is equally incompetent to determine those of identities linking person-denoting terms with terms for extensions. But it would seem quite unclear what other resources might be supposed—by Frege or anyone else—to have determined that Caesar is no extension.[8] Frege's shift **(p.340)** to explicit definition of numbers as certain extensions thus actually provided no discernible way past the Caesar Problem at all.

## 2. Is a Solution to the Caesar Problem Needed?

Almost anyone's initial reaction to these sections in *Grundlagen* will be one of puzzled impatience. For the mixed identity statements which Frege treated as calling for a redirection of his effort feature neither in pure number theory nor in its applications. It is true that we do have what *feels* like a knowledgeable reaction to such statements in a wide class of cases: it is the normal view, for instance, that Caesar is—of course—*not* a number. But even if that *is* a knowledgeable reaction, it is not immediately clear why a satisfactory foundation for pure and applied arithmetic has to be able to make sense of it. And it is anyway debatable whether it is in fact knowledgeable—whether it should count as any serious deficiency in someone's grasp of Number (or Person, for that matter) if they did indeed profess themselves baffled by the question, or regarded it as indeterminate. Why should an account be reckoned as incomplete which fails to resolve such *outré* questions? What importance do they have for the understanding of number-theory?

Even if a significant kind of deficiency is attested to by the problem, there is a further prima-facie cause for impatience with Frege's handling of the issue. After giving his ill-omened explicit definition of numerical terms by reference to extensions, the only use Frege makes of that definition—in the sketch that follows of the derivation, on that basis, of the fundamental laws of arithmetic—is in proving Hume's Principle. Thereafter, it is Hume's Principle, not the explicit definition nor any appeal to extensions, which plays the key role in his sketch of how the fundamental laws, including the crucial theorem that every finite number is succeeded by another, may be derived. This suggests (a result since confirmed, and dubbed *Frege's Theorem* by George Boolos)[9] that adjoining Hume's Principle as the single additional axiom to a standard formulation of second-order logic yields a theory —reasonably taken to be consistent[10]—in which the Dedekind–Peano axioms for arithmetic can be obtained as theorems. This is at least a finding of considerable *mathematical* interest. But if Hume's Principle could rightly be viewed as some sort of a priori conceptual truth—for instance, as being, though not a definition, at least analytic of the concept of (cardinal) number —then Frege's Theorem would, it seems, open the way to potentially very important philosophical insights into the **(p. 341)** epistemological status of number theory, indeed perhaps to the vindication of a suitably qualified version of logicism. In other words, it may be felt that even if the Caesar Problem is indeed an indication that Hume's Principle is incomplete as an explanation of numerical terms, it could still be a correct *partial* explanation whose epistemological status—assuming it transmissible in second-order logic—might thus rub off on the Dedekind-Peano axioms. Frege's seemingly precipitate and fastidious reaction to his objection caused to him to miss the chance to explore this crucial epistemological possibility.

However, there is reason to regard such impatient reactions to Frege's discussion as misconceived. There are at least two broad reasons why he could justifiably have felt impelled to take the problem as seriously as he did. The first is internal to his general orientation in the philosophy of mathematics: specifically, it flows from his platonism—his conception of numbers as a type of self-subsistent object [*selbstandiger Gegenstand*], one species, as it were, within an all inclusive domain of objects. If such a platonist conception is to be legitimate, there surely have to be facts of the matter about *which* objects the numbers are. The concept of number must possess 'sharp limits to its application'[11]—it cannot just be indeterminate whether the putative referents of the terms which Hume's Principle enables us to introduce coincide with any previously understood or independently intelligible kind of objects. Hence if there really are such referents, a satisfactory conception of the objects in question must somehow implicitly settle whether they are or are not trees, tigers, persons, or countries.[12] So if what is supposedly a sufficient explanation, or characterization, of the cardinal numbers leaves such matters indeterminate, then either its claim to sufficiency as an explanation must be discarded or Frege's platonism must be. Frege, perfectly consistently, chose to discard the former.[13]

The second reason is closely connected with what in *The Varieties of Reference* Gareth Evans called the *Generality Constraint*.[14] Readers familiar **(p.342)** with his discussion will recall that Evans applied this principle in the first instance to *thought*, rather than linguistic understanding: the contention was that a thinker can grasp any particular thought only if possessed of a grasp of every other thought in a relevant *local holism*. For instance, a thinker who understands the thought that *a* is F, must likewise understand each thought of the form, *b* is F, whose subject concept he understands and which presents (if anything, then) something which is apt to be F; and likewise he must grasp any thought of the form, *a* is G, where G is any concept under which items of the kind to which *a* belongs are apt to fall. The principle is plausible enough—though, taken at the level of thoughts, it raises philosophical issues which are unprovoked by the corresponding principle governing sentence-understanding. It is the latter principle which is relevant in the context of the Caesar Problem—and *it* is apt to seem merely platitudinous. It is a platitude that to understand any sentence is to understand the meanings of its constituents and the significance of the way in which they are put together. Since the latter items of understanding are essentially general—grasp of the meaning of any semantically significant component of a sentence just consisting in grasp what it contributes to the meaning of any sentence in which it can significantly occur—it follows that to understand any sentence is to understand that range of significant sentences which can be derived from it by permutation of understood constituents.[15] So—*modulo* a qualification to be entered below —someone who understands a subsentential expression, '*f*', must thereby grasp the meaning of any significant sentence, '—*f*—', whose matrix '—. . . — ' they already understand. It is merely a special case of this that one who understands a range of terms '$a_1$', '$a_2$', . . . , '$a_n$' should understand each significant predication on them of any (n-adic) predicate already understood by him; or that one who understands such a predicate should understand any significant sentence in which it is applied to terms, '$a_1$', '$a_2$', . . . , '$a_n$', which he understands.

The bearing on the Caesar Problem is evident. Suppose the stipulation of Hume's Principle does succeed in fixing the meaning of sentences of the form, 'Nx:Fx = Nx:Gx', and with it their ingredient singular terms. Then, by the Generality Constraint, it must somehow simultaneously explain any sentence which can be derived from such sentences by permutation of understood constituents—for instance, by permuting 'Caesar' with a numerical term.[16] But that, Frege is saying, is just what it does not appear to do. So it would follow that it never did properly explain even the purely numerical identity contexts which it explicitly concerns.

**(p.343)** If that is right, then there is not—in the present case, given the kind of explanation Hume's Principle seeks to provide—the kind of gap between complete and partial explanation which the above complaint—that Frege missed the chance explicitly to derive Frege's Theorem and explore its epistemological significance—presupposes. The claim that Hume's Principle is analytic of the concept of number, or an a priori conceptual truth, could foreseeably be justified only by showing either that it is a consequence of some independently given explanation of that concept or by laying it down as itself (part of) such an explanation. The complainant is canvassing the latter possibility: that, even if (whether or not objectionably) silent on the relevant mixed identities, Hume's Principle can still serve as an at least *partial* explanation. But 'partial' explanation here presumably means: *complete* explanation of *some*, but only some, of the contexts that need explaining. The claim is, in other words, that the explanation puts us in position *fully* to understand such statements as 'The number of planets = the number of numbers less than 10', even if it leaves us inadequately prepared for such statements as 'The first man to reach the South Pole = the number of planets', and even though we satisfy whatever conditions may reasonably be reckoned necessary for fully understanding such terms as 'the first man to reach the South Pole'. It is thus presupposed that we may fully understand two identity statements 'a = b' and 'c = d' and yet be quite unable to grasp the sense of a third identity statement 'a = c', for all that it is composed entirely out of familiar ingredients in an entirely familiar way. But that is just what is excluded by the Generality Constraint. If the Generality Constraint is in full force, there is simply no option of maintaining that Hume's Principle provides at least—though only—a partial explanation of the concept of number (and associated canonical numerical terms).

Now for the—obvious—qualification promised three paragraphs ago. It is that while the Generality Constraint for linguistic understanding is surely a platitude at some level, it may still be a substantial—and controversial—question just how extensive is the appropriate *degree* of Generality in any particular case. It is a pervasive feature of significant predicates (open sentences) in natural languages that they have a *range of categorial significance*—a range of things for which it makes *literal* sense to predicate them and their contraries—which is normally narrower than the range of singular terms to which they may be applied without grammatical incongruity. ('Venus is west of the Earth', 'It is five o'clock on the sun', 'Europe is more cylindrical than it was'.) If that is right, then the proper demand imposed by the Generality Constraint on the understanding of a singular term is only that one understand any sentence resulting from combining that term with a predicate whose range of categorial significance includes the referent (if any) of the term.[17] Thus a possible line of resistance to the complaint that the Direction Equivalence, **(p.344)** for example, fails to comply with the Generality Constraint as applied to terms of the form 'Dx', since it fails to explain the truth conditions of perfectly well-formed contexts which result from their insertion into the already understood sentence frame, ' . . . = q', would be to say that it merely seeks more Generality than is necessary: that a satisfactory explanation of the truth-conditions of identity contexts featuring direction-terms need accomplish no more than to treat those in which the other term falls within the relevant range of categorial appropriateness—which 'the Equator', or 'England', say, do not.

But why not? What *determines* the relevant range of categorial appropriateness? In order to move on to a definite treatment of the Julius Caesar Problem in general, it seems one would need as a minimum to argue something like this: that when 'p' is a term introduced by the stipulation of a Fregean abstraction, no predicate of the form, ' . . . = q', where 'q' is a singular term which can be grasped in innocence of any such abstractive stipulation, includes the referent, if any, of 'p' in its range of significance. This would only be a minimum because, even if the point were made good, the issue would still remain of the Caesar Problem posed by mixed identity statements concerning different types of Fregean abstract. But putting that to one side, the point remains that the (single) *atomic predicate* in question here is ' . . . = . . . ', and the thought dies hard that identity is categorially appropriate simply to any *objects*—numbers, polar explorers, emperors, directions, countries, the Equator, whatever. How can it be a kind of category mistake to raise the issue of the identity of Caesar and the number zero? The point is not whether numbers and persons belong to different ontological categories. Presumably—at least, one would like to see it made out that—they do. But in that case, that very point should suffice for their *distinctness*. What kind of consideration could tend to show, to the contrary, that the statement of their distinctness was *itself* a violation of some kind of categorial boundary? We return to these issues in the next section.

One final point. We have observed the corollary of the Generality Constraint that—unless some appropriate restriction on the relevant degree of Generality can be motivated—sense has been assigned to expressions of the form, ‘Nx:Fx’, only if truth-conditions have been established for each statement of the form, Φ(Nx:Fx)' where ‘Φ’ is an understood predicate. The complaint is then that this is not achieved by Hume's Principle for any such statements where ‘Φ’ is of the form, ‘ . . . = a’, and ‘a’ is some independently understood term. But if this complaint is fair, then Hume's Principle equally fails to explain predicates of the form, ‘ . . . = Nx:Fx’. For (the satisfaction-conditions of) such predicates would have been explained just if the truth-conditions had been explained of each statement of the form, ‘a = Nx:Fx’, where ‘a’ is any independently understood term. Accordingly, from the point of view of the Generality Constraint, the failure of Hume's Principle to explain the truth-conditions of mixed identity contexts implies both a failure to fix the sense of numerical terms and a failure to fix the **(p.345)** satisfaction conditions of the relevant numerical predicates. So much is obvious enough. But several consequences of the latter failure are worth special note. One is that we cannot justify the use of the predicate ‘ . . . is a number’—explained as ‘For some F, . . . = Nx:Fx’—since it involves higher-order quantification into predicates we have failed to explain. Another is that it makes only doubtful sense to quantify over its instances—the numbers themselves. But there is a third consequence—a quite particular reason why failure to provide sense for predicates of the form ‘x = Nx:Fx’ would be a crippling technical defect. For among the predicates which would then be lost to us are some which must be assigned a sense if anything like Frege's proof of the infinity of the number series is to go through. The key idea of that proof, it will be recalled, is that the number of finite numbers less than or equal to any given finite number always follows directly after that number in the series (i.e. always exceeds it by one). This requires us to make sense of numerical terms which embed (simpler) numerical terms as, to take the simplest case, ‘Ny: y = Nx:x ≠

x’ does. And that we cannot do, unless we make sense of the predicate ‘x = Nx:x ≠ x’. In general, it is essential to Frege's construction of arithmetic that the variables in open sentences of that kind—open sentences embedding clauses in which free variables flank the identity sign alongside numerical singular terms—be available for binding both by the ordinary first-order quantifiers and by the numerical operator. But neither operation can be intelligible, it seems, unless the satisfaction-condition of the open sentence is; and that cannot be intelligible—so the Generality Constraint would seem to require—unless we have solved the Caesar Problem.[18]

## 3. A Recent Attempt to Finesse the Problem

Nevertheless, the idea that the Caesar Problem must be confronted and solved by any version of logicism based upon Hume's Principle has recently been challenged by Richard Heck.[19] In Heck's view, what is really at issue, when Frege revives the Julius Caesar Problem at *Grundlagen* § 66, is whether, if our understanding of names of numbers went no further than could be explained by Hume's Principle, that would be a sufficient basis for knowledge of the basic laws of arithmetic, and, more especially, for knowledge of the infinity of **(p.346)** the series of natural numbers.[20] It is thus a version of the final problem mentioned in the preceding section to which he accords centre stage. In doing so, Heck does not mean to deny that the Caesar Problem can properly be seen as raising more general philosophical difficulties—on the contrary, he explicitly concedes that it gives rise to

> two challenges. . . . to show how, on the basis of the understanding of names of numbers captured by Hume's Principle, one can come to understand questions of ‘trans-sortal identification’ and, in particular, to know that numbers are of a different sort from people and other such objects [i.e. from ‘basic objects’, in the terminology Heck subsequently adopts; and] . . . to explain how, on the same basis, one can arrive at an understanding of such open sentences as ‘x is the number of Gs’.[21]

But analogues of these difficulties plausibly afflict any attempt to explain a concept—such as *direction* or *shape*—by means of a Fregean abstraction principle, whereas the most important difficulty raised by Caesar is, Heck thinks, quite specific to *number*, viz. that of providing a cogent proof, consistently with the explanatory limitations exposed by Caesar, of the infinity of the series of natural numbers. And in his concluding remarks, Heck leaves no doubt that he believes that his answer to *this* difficulty removes entirely the apparent need to meet the others.[22]

Heck's principal thesis is that there is a 'way of founding our knowledge of arithmetic on (an analogue of) Hume's Principle which, to some extent, *sidesteps* the Caesar objection.[23]

How is the 'side-step' accomplished? Frege thought that, in order to prove the infinity of the number series, he had to treat numbers as objects of the same type as other less exotic objects such as Julius Caesar; thus the sort of proof he canvasses in *Grundlagen* § § 82–3 requires us to take the domain of the first-level quantifiers in Hume's Principle as comprising numbers themselves, as well as more mundane objects such as Julius Caesar—which is just what it seems we cannot intelligibly do unless sense is given to identity statements linking terms for numbers with, say, terms for persons. But Heck holds that this conception of the matter is not mandatory.

Engagement with the Caesar Problem is an immediate consequence of taking numbers to be entities of the same type (or logical Sort, in Heck's preferred terminology) as objects of other kinds, such as persons. Two expressions are of the same logical type (Sort), Heck explains, only if they are **(p.347)** intersubstitutable *salva significatione*. Thus if there is any sentence of the form ' . . . t . . . ' which has a sense, and a corresponding sentence of the form ' . . . u . . . ' which does not, then the expressions 't' and 'u' cannot be of the same logical Sort.[24] Since '0 = 0' has a sense, '0 = Julius Caesar' must likewise have a sense, if '0' and 'Julius Caesar'—and 0 and Julius Caesar—are to be of the same logical Sort. More generally, if numbers and persons are to be of the same Sort, then mixed identities featuring terms for numbers and terms for persons must always have sense. More generally still, any claim to the effect that entities of sufficiently different kinds are nonetheless of the same Sort will give rise to a version of the Caesar Problem.

As the reader will speedily appreciate, we are being returned by this line of thought to the issues opened earlier concerning the appropriate degree of Generality to be achieved by a satisfactory explanation of numerical singular terms and whether, in particular, it is extensive enough for Caesar to indeed present a problem. Why did Frege himself think that he was obliged to take numbers to be of the same Sort as *basic objects* (where basic objects are ‘such things as people and trees and rocks and rivers’)? The answer, Heck suggests, is that he did so, not because he thought that *all* objects are of the same Sort (although he surely did think exactly that), but for a reason much more directly linked to his logicist project, namely that *numbers can themselves be counted*. Heck writes:

> Once names of numbers have been introduced by means of Hume's Principle, no *further* explanation of such expressions as ‘the number of numbers less than 5’ would appear to be required. One would expect speakers immediately to understand such expressions and to know, for example, that ‘the number of Roman emperors is the same as the number of numbers less than 5’ is true if, and only if, there is a one–one correlation between the Roman emperors and the numbers less than 5. It would appear to follow that the predicates ‘x was a Roman emperor’ and ‘x is a number less than 5’ must be of the same logical Sort. For the names Hume's Principle explains contain only predicates of a single logical Sort, namely, those of the same Sort as ‘x was a Roman emperor’ and other predicates true or false of basic objects. But, if these predicates are of the same Sort, so must *singular terms* to which the predicates can sensibly be attached be of the same logical Sort (for the Sorts of predicates are *determined* by the Sorts of their acceptable arguments).[25]

If that is right, then terms of the kind ‘the number of numbers less than 5’ must be of the same logical Sort as terms for basic objects. Frege could not dispense with such terms without abandoning his claim that ‘numbers can themselves be counted’, and he could not jettison that claim, since—quite **(p.348)** apart form its intuitive correctness—his proof of the infinity of the series of natural numbers depends upon it.

But the foregoing reasoning, as Heck observes, is flawed. In essence, his point is that it is *not* necessary to maintain that numbers must be of the same Sort as basic objects in order to sustain the claim that numbers can themselves be counted. What it *is* necessary to maintain is only that *numbers ascribed to concepts true or false of numbers* are of the same Sort as *numbers ascribed to concepts true or false of basic objects*—but from this the problematic claim that numbers are of the same Sort as basic objects simply fails to follow. Provided that numbers ascribed to concepts true or false of numbers can be held to be of the same Sort as numbers ascribed to concepts true or false of basic objects, nothing stands in the way of proving the infinity of the series of natural numbers, along with the remaining Dedekind–Peano axioms.[26] And since—or so Heck's thought appears to run—this claim can be upheld without claiming that numbers are of the same Sort as basic objects, the Caesar Problem (at least in the form in which Heck assumes it to be intractable) can be side-stepped.

Here, in more detail, is how the proposed solution works. Suppose we are working in a two-sorted language, comprising 'basic' individual variables 'x', 'y', . . . ; 'numeric' individual variables '*x*', '*y*', . . . ; basic and numeric predicate variables 'F', 'G', . . . , and '*F*', '*G*', . . . ; and relation variables of various logical types indicated by subscripting in the obvious way —'$R_{bn}$', '$R_{bb}$', . . . . The language contains not one but two identity predicates: ' . . . $=_b$ . . . ' in which both argument-places must be filled by basic terms, and ' . . . $=_n$ . . . ' in which both must be filled by numeric terms—'mixed' identity statements are disallowed as ill-formed. Now, if Hume's Principle is formulated as:

(Hume's Principle$_{bb}$)

$$\forall F \forall G\{NxFx = NxGx \leftrightarrow \exists R_{bb}[\forall x \forall y \forall z((R_{bb}xy \,\&\, R_{bb}xz \rightarrow y =_b z) \,\&\, (R_{bb}xz \,\&\, R_{bb}yz \rightarrow x =_b y)) \,\&\, \forall x(Fx \rightarrow \exists y(Gy \,\&\, R_{bb}xy)) \,\&\, \forall y(Gy \rightarrow \exists x(Fx \,\&\, R_{bb}xy))]\}$$

then, whilst we shall be able—after defining ‘0’ ‘x precedes z’, ‘x ancestrally precedes z’ and ‘x is a natural number’ in essentially Frege's way-to prove some of Frege's axioms for arithmetic, one—the crucial axiom asserting the existence of successors—will elude us. But there is, Heck explains, a way past this difficulty. For ‘[whilst] we do not, intuitively, consider numbers and **(p.349)** people to be of a single logical Sort, we do intuitively consider all numbers to be of a single Sort’.[27] And we can consistently do so (i.e. consistently take ‘the number of Roman emperors’ and ‘the number of numbers less than 5’, say, as terms of the same Sort, even in the absence of a general answer to the Caesar objection) because the truth-condition of the identity statement linking these terms does not involve that numbers are of the same Sort as people. All that is required is that there be a one–one correlation between the Roman emperors and the numbers less than 5, and one–one correlation—and indeed many other relations—can perfectly intelligibly obtain trans-Sortally. Hence the identity statement can, and does, have a sense, irrespective of whether numbers and people are of the same Sort.

More generally, in the two-sorted theory outlined above, we have not one but *three* formulations of Hume's Principle. Besides (Hume's Principle$_{bb}$), which governs (terms formed from) concepts true or false of basic objects, we have a version governing concepts true or false of *numbers themselves*:

(Hume's Principle$_{nn}$)

$$\forall F \forall G\{NxFx = NxGx \leftrightarrow \exists R_{nn}[\forall x \forall y \forall z((R_{nn}xy \,\&\, R_{nn}xz \rightarrow y =_n z) \,\&\, (R_{nn}xz \,\&\, R_{nn}yz \rightarrow x =_n y)) \,\&\, \forall x(Fx \rightarrow \exists y(Gy \,\&\, R_{nn}xy)) \,\&\, \forall y(Gy \rightarrow \exists x(Fx \,\&\, R_{nn}xy))]\}$$

and a version governing ‘mixed’ identities:

(Hume's Principle$_{bn}$)

$$\forall F \forall G\{NxFx = NxGx \leftrightarrow \exists R_{bn}[\forall x \forall y \forall z((R_{bn}xy \,\&\, R_{bn}xz \rightarrow y =_n z) \,\&\, (R_{bn}xz \,\&\, R_{bn}yz \rightarrow x =_b y)) \,\&\, \forall x(Fx \rightarrow \exists y(Gy \,\&\, R_{bn}xy)) \,\&\, \forall y(Gy \rightarrow \exists x(Fx \,\&\, R_{bn}xy))]\}.$$

Heck calls the two-sorted theory whose non-logical axioms are just (Hume's Principle$_{bb}$), (Hume's Principle$_{nn}$) and (Hume's Principle$_{bn}$) Two-Sorted Fregean Arithmetic (briefly, 2FA). If we now drop all reference to basic objects from 2FA, and omit the then-irrelevant relational suffixes, the result is a single-sorted theory whose sole non-logical axiom is:

$$\forall F \forall G\{NxFx = NxGx \leftrightarrow \exists R[\forall x \forall y \forall z((Rxy \,\&\, Rxz \rightarrow y = z) \,\&\, (Rxz \,\&\, Ryz \rightarrow x = y)) \,\&\, \forall x(Fx \rightarrow \exists y(Gy \,\&\, Rxy)) \,\&\, \forall y(Gy \rightarrow \exists x(Fx \,\&\, Rxy))]\}$$

which is, 'modulo the italics'—as Heck puts it—just Hume's Principle in its simple form. Thus 2FA is revealed as an extension of single-sorted Fregean **(p.350)** Arithmetic, and proves whatever it does. In particular, since infinity is a theorem of FA, it is likewise a theorem of 2FA.

Can Heck's ingenious manœuvre really avoid the Caesar Problem? It would do so, it seems, if—but only if—it really does obviate any need to assign sense (and so truth-conditions) to mixed identity statements such as 'Julius Caesar = the number of Roman emperors'. Heck clearly thinks it achieves this, and does so by showing how the logicist may *avoid allowing* that numbers must be of the same Sort as basic objects. But here we need to be careful. Heck makes a plausible—and indeed, convincing—case that, once the logicist has regimented discourse about numbers and other things in terms of the two-sorted language described, she can proceed with the formal development of arithmetic, in the shape of 2FA, without commitment to or against the claim that numbers are of the same Sort as basic objects—indeed without even so much as concerning herself with the question whether they are or not.[28] It would, however, be a noteworthy step from that result to the further conclusion that the question can simply be denied philosophical house-room—that the logicist can just dig in and declare it out of bounds. Not only does that further conclusion fail to follow; it is hard to see how it can be true. On the face of it, Heck's position must be *either* that our everyday (unregimented) discourse about numbers, etc., is—albeit implicitly and quite untransparently—*already* (at least) two-sorted, *or* that, while it is not, it is open to the Fregean to introduce the two- (or more-) sorted language he describes as a replacement for it. Either way, we cannot avoid confronting something like the question whether numbers are or aren't of the same Sort as people, etc. On the first alternative, the issue would be that of explaining and justifying our alleged (implicitly and untransparently) two-sorted practice. On the second, the question is transformed into: *ought* everyday discourse of numbers and other things be so regimented as to assign numerical expressions to one Sort but, for example, terms for persons, trees, and the like to another (disjoint) Sort, or should they rather be assigned to some single comprehensive Sort?

Heck himself gives little indication how he would have the question, in either form, answered. All he says is: 'No logicist need make this claim [i.e. that numbers are of the same Sort as basic objects]—which is a good thing, since it's very unclear whether it's true.’ His reticence is, in one way, quite unsurprising—after all, he thinks he has found a way to avoid the question altogether; but in another, it *is* surprising. For it is not as if his position is secure either way, whichever answer is returned. If numbers and persons *are* of the same Sort, then—by the very meaning of ‘Sort’—it follows that questions like ‘Is Julius Caesar the number of numbers less than 5?’ must be admitted as significant, and **(p.351)** it should therefore be possible to say what should determine how they should be answered. It would also follow that the range of significant predication possessed by the kind of open sentences embedding numerical terms which the derivation of the infinity of the number series from Hume's Principle must assume to be intelligible, is not after all restricted to numbers, so that the proof which proceeds from Hume's Principle$_{nn}$ would just be an epistemological sham. It thus seems clear that, once the question is allowed to arise at all, Heck must return a negative answer to it: numbers and persons are of different Sorts.

This answer must, however, respect the controls on the question imposed by Heck's own explanation of Sort-identity. It will be the right answer, in other words, if—but only if—terms for persons are not cross-substitutable with numerical terms in all contexts *salva significatione*. How is it to be established that that condition is met? How are we to show, for example, that while ‘The number of Roman emperors = the number of numbers less than 5’ has a sense, ‘Julius Caesar = the number of numbers less than 5’ does not? It would, clearly, be unsatisfactory to rely upon intuitions of significance (whose?)—if only because they are unlikely to speak in unison. Nor is it a matter for stipulation. What is needed is a principled reason for denying that this configuration of individually significant words adds up to an expression which, taken as a whole, expresses something true or false. But on this issue, crucial for his purposes, Heck is silent.

It seems to us utterly unclear how it might be argued that the relevant kind of mixed identity statements might best be regarded as lacking sense. So long, then, as distinctions in Sort are harnessed to that issue, it is unclear in tandem how a two- or multi-sorted approach might possibly be justified. That is not to say that a relevant distinction in Sort, or category, might not be motivated on other grounds of some kind. But as soon as such distinctions are freed from the tie to the putative senselessness of the relevant mixed identity statements, it becomes very difficult to see how any point in Heck's overall approach could remain. For here is the dilemma. If numbers are of the same sort as persons and other ‘basic’ objects, Heck's two-sorted system is based on a philosophical mistake, and the technical distinctions it exploits are baseless. But if numbers are not of the same sort as persons and other ‘basic’ objects, then the knowledge that they are not may now be taken as grounds not for the senselessness of the claim that Caesar is zero, for example, but for its *falsity*.[29] So we should have accomplished a solution to the problem, without the need for Heck's technicalities, by providing a justification for regarding every number and every basic object as of distinct sorts.[30] That will be the gist of our own eventual approach.

## (p.352) 4. Semantic Nominalism and Transworld Abstraction[31]

In *Frege: Philosophy of Mathematics*, Michael Dummett asserts the following indeterministic thesis:

> The stipulation that the direction of a line *a* is to be the same as that of a line *b* just in case *a* is parallel to *b* does not determine whether the direction of a line is itself a line or something quite different. Even if the requirement were to be made that every direction should itself be a line, the stipulation would in no way determine which line any given direction was to be; it could, in fact, be any line whatever. A convenient choice would be to take some point as the origin O, and identify the direction of any line *a* with that line through O that was parallel to *a*; even so, any point could serve for this purpose as the origin.[32]

Dummett is, of course, making no special claim about lines and directions: it is an intended implication of his remarks that *no* (first-order) abstraction principle can enforce construal of the reference of the terms it introduces as lying outwith the domain of the relation on which the abstraction is made: that we are entirely free to take directions to be (particular) lines, geometrical shapes to be (particular) closed figures in a plane, lengths to be (particular) length-bearing objects, and so on.[33]

The indeterminacy Dummett here alleges is more specific than that raised by the Caesar Problem in its most general form but it is hard to see what could prevent it, if justified, from ramifying into the latter. If any particular line exemplifying a particular direction may legitimately, under some suitable selection of an origin, O, be taken to *be* that direction, why not go further and treat a bijection, M, between the lines through O and *any* appropriately ordered field of objects—including Roman Emperors, if you wish—as entitling us to identify the direction of a line, *a*, with the M-correlate of that line through O which is parallel to *a*?

However the special interest of Dummett's claim resides in its bearing upon the feasibility of a coherent nominalist response to Fregean abstraction. Suppose, to fix ideas, that the stipulation of the Direction Equivalence

(i) D*a* = D*b* iff lines *a* and *b* are parallel

**(p.353)** is augmented, for each of a range of predicates, $F_k$, for which parallelism is a congruence,[34] by a battery of principles of the form,

(ii) $\Phi_k$ (D*a*) iff $F_k$ *a*,

determining the satisfaction conditions of a range of new predicates of directions by reference to those of counterpart predicates on lines; and suppose that we further stipulate[35] for each such pair, $\Phi_k$ and $F_k$, that

(iii) (∃ x) ($\Phi_k$x) iff (∃ x) $F_k$x.

Finally, suppose—of course Dummett himself wasn't thinking of them this way—that the lines referred to and quantified over on the right-hand sides of these stipulation are conceived as concrete entities of some kind—actual thin black straight-edged strips magnetically anchored on some plane surface, for instance, or laser-beams.

The collective effect of these stipulations will be to make it possible to play a simple ‘language-game’ of directions in which—at least apparently—they are treated as objects, subject to simple predications and to quantification. The neo-Fregean thought about the situation will be that, contrary to any misplaced nominalist scruples, it provides a very simple illustration of how the use of a range of abstract singular terms can be legitimated and disciplined and of how, by getting to know the truth of contexts falling within the scope of the right-hand sides of the stipulations, we can verify simple truths about (one kind of) *abstracta* and, *a fortiori*, gain an assurance that the relevant abstract terms refer.[36]

How is a sceptic about all *abstracta*—a nominalist—to respond to this line of thought?[37] There are a number of well-known, more or less traditional responses. Least interesting is the flat allegation that the Fregean's ‘explanatory stipulations’ are simply illegitimate, that one cannot make an unintelligible claim intelligible, or an unjustifiable one justifiable, merely by stipulating that it is to have the same truth-conditions as an unproblematic one. This is uninteresting because dogmatic: it ignores altogether the neo-Fregean's claim to have allayed the nominalist's concern precisely by giving sense to the problematic contexts in a way that makes them metaphysically and epistemologically tractable.

**(p.354)** A second response allows that the various equivalences may legitimately be stipulated—but only subject to the proviso that no significant syntax is discerned in the left-hand sides, which must be taken as mere notational variants of the nominalistically acceptable right-hand sides. This too is unattractive. The problem with it is that if the stipulations are sufficiently carefully crafted, the expressions occurring within the left-hand sides will behave in all respects just as if they had the syntax and content that the Fregean proposes to assign them. In the language-game for directions, for instance, everything will seem to function just as it ought in a regular first-order language with identity: the '=' symbol will behave as if it expressed an equivalence relation, congruent with respect to each predicate defined on the (apparent) singular terms standing for directions, which will in turn behave just as they should with respect to the (apparent) existential quantifier, and so on. So the nominalist will owe an explanation of why he refuses to take this surface as indicative of the currency of genuine notions of identity, quantification, predication, and reference. The strongest argument for so refusing would be if those notions are unintelligible in the context of purported discourse of abstract objects—but just to re-invoke that claim as an axiom is, of course, is to *presuppose*, not argue, that the Fregean's stipulations fail precisely to bestow the requisite intelligibility. So the nominalist stance again threatens to emerge as *ad hoc* and unmotivated—unless independent fault can be found with the Fregean view of the stipulations.

A third, sometime very popular nominalist move is the more generous reductionist tendency which, applied in the present setting, would have it that while the equivalences are indeed legitimately stipulated and while there is indeed significant semantic structure on their left-hand sides, the very equivalence between their left- and right-hand sides itself somehow serves to neutralize the apparent distinctive commitments of the contexts introduced—that we may conclude from the very obtaining of the equivalences that to talk of directions and their properties, for example, is to talk of nothing 'over and above' lines and their properties, and any appearance of additional ontological commitment is an illusion. But, despite its sometime popularity, the obligations of this proposal are actually extremely awkward. All that distinguishes it from the second proposal is the concession of significant semantic structure to the left-hand sides. But what is this structure if not that discerned by the Fregean—viz. the semantic structure of quantification over, identification and differentiation among, predication on and singular reference to directions? If *that* type of structure is conceded, then it seems the intended nominalism is forfeit—for then there are true, understood statements in which we refer, to quantify over, etc., abstract objects. But to deny structure altogether is to fall back on the second response. So it seems the nominalist must find a third way—a reconstrual of the surface syntax of the left-hand sides which avoids discerning any reference to or quantification over *abstracta* there, but does construe their meanings as compositional, and **(p.355)** does sustain the equivalences. It is very doubtful if any such account is possible.[38]

There is a fourth nominalist option—though perhaps best construed as a variant of the first—provided by Hartry Field. This is the 'rejectionist' response discussed in Wright (1990).[39] Field rejects the unconditional stipulation of the type of equivalences typified by the examples above, but does allow that the syntax of the left-hand sides may properly be taken to be what it appears, and—here is the departure from the first view—that the biconditionals *can* be stipulated to hold, *subject to the proviso that the general ontological presumption of their left-hand sides is satisfied*. Thus we may not legitimately stipulate the Direction Equivalence *simpliciter* but it is, in Field's view, all right to stipulate it *provided there are directions at all*—more specifically, we may stipulate that provided lines *a* and *b* respectively have directions, then those directions will be identical just if *a* and *b* are parallel. The Fregean's stipulations thus give way to a range of conditional stipulations in which (effectively) the original equivalences appear as consequents while the supposition of the requisite abstract ontology occurs as antecedent.

This proposal is well worth separate discussion.[40] Here it must suffice to say that like the first line of response proper, Field's proposal risks the charge of mere dogmatism. The Fregean is offering an account of how reference to abstracta can be made intelligible and how we can know a wide range of statements about them to be true. This account proceeds via stipulation of what are presented as meaning-conferring equivalences. It is therefore no good for the nominalist just to insist that these equivalences must be rejected or further conditionalized on the grounds that we somehow already know that reference to, or knowledge of *abstracta* is problematical. Rather it must be argued on neutral terms that the Fregean stipulations fail in the explanatory task assigned to them. For if they succeed, there is simply no motivation for nominalism in the relevant class of cases.

So far a good judge might have the Fregean a little ahead on points. The whole dialectic is transformed, however, if Dummett is right that (first-order) Fregean abstraction need never be taken to introduce a concept of a *new* kind of thing—and hence cannot be regarded as introducing a concept of a **(p.356)** distinctive kind of thing at all. If that is right, then there is a kind of nominalist—let's call him a *semantic nominalist*[41]—who can acknowledge the success of the Fregean's explanations in the kind of case illustrated: he can wholeheartedly accept the equivalences, without further conditions, allowing that the semantic structure of the left-hand sides is just what it seems; this while he legitimately—according to Dummett—construes the ontology of the left-hand sides as involving no excursion outside nominalist boundaries.[42] Moreover once talk of directions was 'rehabilitated' in this way, there would be no reason why directions and relations upon them may not in turn throw up the material for the right-hand sides of further abstractive stipulations, and so on indefinitely. So far, then, from supplying the resources for legitimating a modest form of platonism, Fregean abstraction promises to emerge—if Dummett's indeterministic thesis is right—as a powerful tool for the rehabilitation of what appears to be talk of *abstracta* in terms of a wholly concrete ontology.

It seems to us, however, that although his claim is technically correct—that is: first-order abstractions, as standardly formulated, must allow of models in which the domain is confined to that of their abstractive relation—the latitude of allowable *interpretation* which Dummett asserts in the above quotation is actually illusory. Specifically, we shall argue that if talk of directions is introduced by laying down the Direction Equivalence *as intuitively intended*, then there is no freedom to go on to stipulate that such talk is actually talk of lines, let alone is it then arbitrary which lines we may accordingly conceive ourselves as thereby talking about. More generally, any first-order abstraction on a concrete domain, provided it is allowed to have what seem to be entirely appropriate *modal and counterfactual implications*, will preclude construal of the terms it introduces as referring within that domain. More generally still, no abstraction of any order—if granted such modal and counterfactual implications—allows the construal of its novel terms as referring to *concrete objects* (on one natural understanding of 'concrete'). The bearing of this point on the Caesar Problem is obvious; we'll return to that in due course.

Consider, then, the following *Simple Argument* against Dummett's specific suggestion about directions and lines, remembering that the lines in question are, for our purposes, concrete, contingently existing objects:

> (i) The direction of a given line exists in all circumstances where that line exists—(so much is entailed by the Direction Equivalence).
> (ii) But any given line might exist even although all the other lines had been destroyed (removed, erased, or switched off).
> So (iii) The only line which is eligible to be the direction of a given line, *a*, is that line itself—for no other is guaranteed to so much as exist in all **(p.357)** circumstances in which *a* has the direction it has. And this holds for any line.
> Hence (iv) if lines are directions, then whenever $a \neq b$, $Da \neq Db$, *even if a and b are parallel*.
> But (v) that contradicts the Direction Equivalence.
> So (vi) Lines are not directions. QED.

Clearly an analogue of this argument can be run in all cases where the abstractive domain consists of destructible—more generally, contingently and mutually-independently—existing objects.
The nominalist will immediately protest that (i) is a wholly tendentious statement of the implications of the Direction Equivalence. What that principle gives us is only that in all circumstances where *a* exists, it will, since self-parallel, have *a* (self-identical!) direction. It does not give us that the object which will then be its direction will be *the same as* the direction it actually has. Compare

> The precise place and time of birth of A is the same as the precise place and time of birth of B iff A = B.

That is a necessary truth, but of course it does not follow that the precise place and time of birth of A is the same in all worlds in which A exists. The intended understanding of (i) in the Simple Argument could be sustained, the nominalist continues, only if direction-terms were *rigid*; that is, only if talk of the direction of *a* in hypothetical contexts would always involve reference to its actual direction. But that is manifestly not so: objects may change the properties on which the identity of an associated Fregean abstract depends—people can change in height and weight, for instance, and *physical* lines may be such that they can change their orientation on a plane.

It is questionable, however, whether the Simple Argument needs to rely on anything as strong as rigidity, properly so regarded. What it essentially requires is only that direction terms have a kind of *quasi-rigidity*. A complex term, $F(\alpha)$, is quasi-rigid with respect to a specific range of properties just in case it has the same referent in all worlds in which α exists and is unchanged in respect of those properties. A direction term, D *a*, is, plausibly, quasi-rigid with respect to the *orientational properties* of the line *a* on the relevant plane: those properties of a line which, modulo the corresponding properties of any other line, determine whether or not the two are parallel. The intuition is strongly that in speaking of the direction *a* would have in a hypothetical scenario in which *a* holds its actual position (relative to some relevant framework) while other lines do not, we should be speaking of the same thing that we actually refer to by categorical uses of 'D *a*'. If that is right, then it would seem lines (i) and (ii) may be replaced by

(i)* The actual direction of a given line exists in all possible circumstances where that line exists and is unchanged in its orientational characteristics (so much is entailed by the Direction Equivalence).

**(p.358)**

(ii) But any given line might exist, unchanged in its orientational characteristics, even although all the other lines had been destroyed (removed, erased, or switched off).

—and now the rest of the argument can run as before.

What are we to make of this? To presuppose that the Direction Equivalence may be appealed to in reasoning about an arbitrary hypothetical scenario of lines and their properties is merely to assume that the principle is a necessary truth. *That* much is presumably all right: the envisaged dialectical situation is one in which the principle is laid down as a putative implicit definition of the concept of direction so—granted the legitimacy of so doing, which the present stripe of nominalist does not contest—we may assume that it is accepted as necessary by all hands. But that only yields that the Direction Equivalence holds *of* any world. What the refashioned argument is additionally assuming is that the scope of the principle encompasses not merely relations of parallelism within a world, but relations of, as it were, *trans-world parallelism*. Such relations include, for example, relations between actual lines as they actually are and hypothetical lines, relations between hypothetical lines in (the same or distinct) hypothetical scenarios, and—the relevant case—relations between actual lines as they actually are and actual lines as they would be in some hypothetical scenario. The specific presupposition of the argument is the legitimacy of claims of the form: if *a*-as-it-is is parallel to *a*-as-it-would-be in scenario such-and-such, then *a* would have the same direction in that scenario as it actually has. How is a claim of that kind supposed to follow from the normal (first- or second-order) statement of an abstraction?

It had better be admitted immediately that the Direction Equivalence, as normally stated, does indeed *not* entail its trans-world counterpart. How could it? An explicit expression of the trans-world principle will require, at the least, means to express the subjunctive conditional and clearly cannot be achieved using just the resources of the first-order language-game of directions described. (So, as granted earlier, the *letter* of Dummett's claim in the above-quoted passage is not challenged by the Simple Argument.) But the fact remains that it makes undoubted intuitive sense to speak of an actual line's being parallel to a line—itself or another—in hypothetical circumstances, and there can surely be no serious objection to understanding the Direction Equivalence as covering suppositions of that general ilk. So much seems well within the compass of the understanding of the concept of direction that we actually—without dwelling on its overt first-order character—elicit from it. It seems entirely intuitive, for instance, both that if *a* is actually parallel to *b*, then *a* as it actually is is parallel to *b* as it would be in any circumstances in which it had undergone no change in its orientational characteristics, and that it should follow that *a*'s actual direction is the same as that which *b* would then have.

**(p.359)** So too for Hume's Principle: it seems entirely intuitive and obvious that if there had been as many knives as there actually are forks, then *the number of knives there would then have been* is the same as *the actual number of forks*. If someone alleges that such a claim receives no mandate from our ordinary understanding of sameness of number, they are wrong. If they allege that there is such a mandate but that it is independent of Hume's Principle, the surely correct reply is that Hume's Principle has an intuitive content which exceeds that of its canonical second-order formulation. Getting a satisfactory formalism in which that larger content can be rigorously conveyed—and a satisfactory overarching modal semantics—may be a very challenging business. But a first stab in English at a compendium of the relevant principles would run:

> (∀F)(∀G)(∀C)(∀ C′) The number of Fs there actually are/would be in C, is the same as the number of Gs there actually are/would be in C′, if and only if the Fs there actually are/would be in C are/would be one—one correspondent with the Gs there actually are/would be in C′

where the C-variables range over worlds or scenarios (however such things are best understood), including the actual world, and each legitimate instantiation will select one of ‘are’ and ‘would be’ at both of the two relevant places in the left-hand side and then make selections on the right-hand side as logically/grammatically dictated by that selection.

If we allow ourselves the jargon of possible worlds for a moment—without necessarily making any metaphysical investment in it—then in general, a *fully modalized* abstraction will encompass both *intra-world* and *inter-world* cases of its abstractive relation, each subdividing in turn into two subtypes: intra-world cases, that is, will include instances where the world in question is actual and instances where it is hypothetical (merely possible); and inter-world cases will include instances where one of the worlds in question is actual and the other hypothetical and instances where both worlds are hypothetical. That is, letting ‘A’ denote the actual world and ‘$H_n$’ any hypothetical world, a fully modalized abstraction will allow its abstractive relation, ≈ to operate over four types of field:

(i) $A \approx A$ (ii) $H_n \approx H_n$ (iii) $A \approx H_n$ (iv) $H_n \approx H_m$.

But there is more. Whenever hypothetical worlds are involved, and the abstraction is first-order, the terms of the relation in a particular case may be either *actual objects* surviving, changed or otherwise, in that world or *hypothetical objects*. Let '$a_i$' and '$h_i$' schematize names of actual and hypothetical objects respectively, and let '$a_{iA}$' be a name of the referent of '$a_i$' as it actually is, '$a_iH_n$' a name of that same object as it would be in $H_n$, '$h_iH_n$' a name of the hypothetical object $h_i$ as it would be in $H_n$. Then we have these **(p. 360)** twelve non-trivial[43] sub-cases:

| (i) | (ii) | (iii) | (iv) |
|---|---|---|---|
| (a) $a_{iA} \approx a_{jA}$ | (a) $a_iH_n \approx a_jH_n$ | (a) $a_{iA} \approx a_iH_n$ | (a) $a_iH_m \approx a_iH_n$ |
| | (b) $a_iH_n \approx h_iH_n$ | (b) $a_{iA} \approx h_iH_n$ | (b) $a_iH_m \approx h_iH_n$ |
| | (c) $h_iH_n \approx h_jH_n$ | (c) $a_{iA} \approx a_jH_n$ | (c) $a_iH_m \approx a_jH_n$ |
| | | | (d) $h_iH_m \approx h_iH_n$ |
| | | | (e) $h_iH_m \approx h_jH_n$ |

Think once again of the direction example as involving fine physical black edges, magnetically attached to a fixed plane surface. The fact is that we, as it seems, effortlessly understand both what it would be for there to be other hypothetical such lines in play besides a particular initial stock, and what it would be for these hypothetical lines to have particular orientations, which then might or might not be changed in further sufficiently specified hypothetical scenarios, along with those of some of our actual lines, and so on. In this setting, each of (ii)(a)–(iv)(e) corresponds, in the best case, to a perfectly intelligible subjunctive thought; essentially, all that is required is that the hypothetical scenarios be sufficiently well specified to ensure determinacy about what lines exist in them and determinacy about what their orientational properties then are.

The suggestion, then, is that we are capable of a perfectly legitimate, fully modalized understanding of any abstraction principle, and that so understood, there is no possibility of systematically interpreting the reference of the terms it introduces as being wholly within any actual concrete population of objects in the field of its abstractive relation.

How should the semantic nominalist respond? Scepticism about modality, worlds, and the counterfactual conditional and an insistence that abstractions be strictly confined to their usual first- or higher-order formulations would be one, rather disappointing reaction. But there are no indications in Dummett's writings that he, at least, would want to pursue that—would want simply to deny that there is any legitimate understanding of abstractions which will sustain premisses of the kind illustrated by (i)* and (ii)* in the revised Simple Argument.

There is however a different possible—rather desperate—line of defence which we need to run down before we can consolidate our conclusion. Allow that the Direction Equivalence, properly formulated in full modal generality, **(p. 361)** does indeed entail that, in any scenario, C, where *a* exists and is orientationally unchanged,

> The direction of *a*-as-it-actually-is = the direction of *a*-as-it-would-be-under-circumstances-C.[44]

That has to be a problematical conclusion for the nominalist—that is, it bars the identification of *a*'s direction with any object that would not exist under suitably selected such C-circumstances—only if we have no choice but to construe the reference of the term on the right-hand side as being within *the range of objects that would exist under C-circumstances*. What, the nominalist may now enquire, enforces that reading? Why may we not take 'The direction of *a*-as-it-would-be-under-circumstances-C' as having reference within the *actual* world, and within an envisaged C-world only insofar as its actual referent—putatively, some particular nominated actual line—exists in that world? In other words: a determined nominalist may respond to the evident possibility of scenarios in which *a* exists with all its actual orientational characteristics, but any particular other line nominated as *a*'s direction does not, by insisting that directions are one and all *actual-world lines* nonetheless. When we speak of the directions that actual, or hypothetical, lines would have in hypothetical circumstances, the items to which we refer—if we do—are always *actual* objects.

This ploy would block the transition from the intuition expressed above like this:

> that in speaking of the direction *a* would have in a hypothetical scenario in which *a* holds its actual position (relative to some relevant framework) while other lines do not, we should be speaking of the same thing that we actually refer to by categorical uses of 'D *a*'.

to

> (i)* The actual direction of a given line exists in all possible circumstances where that line exists and is unchanged in its orientational characteristics.

Simply: the intuition can be accepted without (i)* if the reference of ‘the direction of *a*-as-it-would-be-under-circumstances-C’ can always be construed as to an actual line. In order for the thesis:

> *The direction a actually has* is the same as *the direction a would have in any circumstances where its orientational characteristics are unchanged*

to be true, it suffices that the second italicized term have the same actual referent as the first; it is not required that that object do anything but *actually* exist.

**(p.362)** The move amounts to the insistence that abstraction operators always take wide-scope with respect to counterfactual and modal constructions—that a direction-term occurring in a counterfactual conditional, for instance, never refers to anything except an *actual* direction. This contrasts with the ‘narrow scope’ reading that is generally possible of definite descriptions: a use of ‘The man who would’ (under certain conditions) ‘be King’ may be understood to refer, if to anything, then to an actual man who would be King, and otherwise to lack reference; but it may also—and would often—be understood to refer to, as it were, a hypothetical man—whoever, whether an actual existent or not, would be King in the circumstances in question. The nominalist is in effect now proscribing *ad hoc* any analogue of the latter reading for terms introduced by Fregean abstraction. That may seem mean-spirited. But the question is: what in an abstraction principle requires a more flexible reading of the abstracted singular terms?

Once the full range of modal readings of the right-hand side of an abstraction is allowed however, this ploy becomes unsustainable. Suppose the nominalist scheme of stipulation works as envisaged by Dummett, with the direction of each line, *a*, being identified with a particular exemplar selected from the class of lines parallel to *a*. The following should be a consequence of the fully modalized Direction Equivalence:

> (1) If *a* were to exist in circumstances C but were to be orientationally altered in such a way that no actual line has the orientational characteristics *a* would then have,

then the direction of *a* as it would be under C would not be that of any actual line.[45]

So: to what actual line does the occurrence of ‘the direction of *a* as it would be under C’ effect a reference? Since *a* as it would be under C is parallel to no actual line, and each direction has been stipulated, on the Dummettian scheme, to be some actual line parallel to all lines of which it is the direction, there is no actual candidate. But a nominalist who is trying the desperate ploy cannot say the natural thing: that the quoted expression has no actual reference—that it functions like ‘the present King of France’ in:

> If France had recently restored the monarchy, the present King of France would not have been a descendant of Louis XIV.

That option is foreclosed by the ‘mean-spirited’ insistence that all direction terms take wide scope with respect to modal and counterfactual constructions. So the nominalist is hoist with his own petard.

**(p.363)** The dilemma, then, is this. If the nominalist insists on the wide-scope convention, he cannot construe consequences of the Direction Equivalence like (1). But if he abandons the convention, he must allow the permissibility of a narrow scope reading of ‘the direction *a* would have in circumstances C’ in *this* consequence of the Direction Equivalence:

> (2) If *a* were still to exist and were to be orientationally unchanged in circumstances C, then the direction *a* would have in circumstances C would be its actual direction.[46]

And read with narrow scope, ‘the direction *a* would have in circumstances C’ must be reckoned to invoke an object that *would exist under C*, just as the narrow scope reading of ‘the King of France’ in the example above calls for an object that would exist had the French monarchy recently been restored. The difference, of course, is that, according to (2), the object called for is also an *actual* object—the actual direction of *a*, which *must therefore to be reckoned among the objects that would exist under circumstances C*. And that will suffice to set the Simple Argument off.

This is all quite a complex matter. But we see no way to avoid the conclusion that, if abstractions may indeed be taken in fully modalized form, then the Simple Argument succeeds. For the only apparent way of withstanding it—insistence on the wide-scope convention—is not consistent with acknowledging every legitimate consequence of such an abstraction. Thus no semantic nominalist can accept *any* fully modalized first-order abstraction, for no such principle will be consistent with construing the terms it introduces as referring to contingently and mutually-independently existing objects. Effectively, then, their reference cannot be to *concrete* objects at all, at least as ‘concrete’ is normally (vaguely) understood. Any objects whose concept is introduced by a fully modalized first-order abstraction are not to be identified with physical contingents.

How might any of this extend to higher order abstraction? Of course, it would be no surprise if an actual nominalist wanted nothing to do with higher-order languages in the first place. But if such languages are tolerated, whether somehow nominalistically construed or not, there will be a possible analogue of Dummett's claim in the first-order case. The analogous claim will be that no higher-order abstraction can demand construal of the singular terms it introduces as referring outside the domain of the objects on which the concepts which provide the terms of the abstractive relation are defined. How does this fare?

Very badly. To fix ideas, suppose we have a second-order abstraction:

$$\mathcal{F}(\mathrm{F}) = \mathcal{F}(G) \leftrightarrow \mathrm{F} \sim \mathrm{G}.$$

**(p.364)** The analogue of the Dummettian-indeterminist claim will be that we can always construct an interpretation which has each abstract term, $\mathcal{F}$(F), so formed, referring to some object in the range of predication of F and G. This claim confronts one very serious snag immediately. Higher-order abstractions often call for more abstracts—since they generate more equivalence classes of concepts—than there are objects in the range of the concepts which make up their abstractive domain. (Hume's Principle, for instance, always calls for $n + 1$ abstracts on a domain of $n$ objects—that, model-theoretically, is why it calls for an infinite population of numbers.)[47]

But that is not the whole of the matter. Even if we go past that point and restrict attention to cases where no more abstracts are demanded than there are original objects, there are still going to be problems of precisely the kind explored by the Simple Argument. What, to begin with, is going to be offered by way of a construal of ‘$\mathcal{F}(x \neq x)$’—the term for the $\mathcal{F}$-abstract of the empty concept? Reflect that the property of a concept which corresponds to the orientational characteristics of a line—the property on which its bearing any second-order definable equivalence relation to another concept will depend

—is of course its extension. So appropriate higher-order analogues of premisses (i*) and (ii)* in the Simple Argument would be something like

> (i)** *F*(F) exists (and is the same thing) in all possible circumstances where F is unchanged in its extension–(so much is entailed by the *F*-Equivalence).

And

> (ii)** Any given concrete object assigned as the referent of '*F*(F)' will fail to exist in some circumstances where F has its actual extension.

Can we make both those claims good for F = x ≠ x? Well, (ii)** seems evidently correct. But before we can secure (i)**, we have to reckon with the analogue of the nominalist's earlier 'desperate' ploy—the attempt to insist on a wide-scope construal of the abstracted terms in modal and counterfactual contexts—which if allowed would block the transition from the intuitive thought

> that in speaking of the *F*-abstract F would have in any hypothetical scenario in which F holds its actual extension[48] we should be speaking of the same thing that we actually refer to by categorical uses of '*F* (F)' to (i)**.

**(p.365)** Now, the desperate ploy misfired in the first-order case before because thoughts of the type

> If *a* were to exist in circumstances C but were to be relevantly altered in such a way that no actual object has the relevant characteristics *a* would then have, then the [abstract of] *a* as it would be under C would not be that of any actual object

are both intuitively compelling and demand a narrow-scope construal of the abstract term occurring in the consequent. But now we hit the snag that no analogue of that for the higher-order case—specifically, it would be something along the lines:

> If F were to have an extension under circumstances C which no concept actually has, then the *F*-abstract of F as it would be under C would not be that which any concept actually has—

is sustainable if concepts are extensionally individuated or if—the present case—F necessarily has its actual extension. So, unless some intensional conception of concepts, or of whatever else is taken to comprise the range of the higher-order variables, can be enforced, it is not immediately clear how to block the 'desperate ploy' in the present, higher-order setting.

Maybe hypothetical objects can come to the rescue. Suppose C is a scenario in which there would be some new objects, some of which would be F. Consider the notion of the (extensional) concept 'F' would express in C if the language were intensionally unchanged—call that the C-concept for 'F'. That seems perfectly intelligible. But no actual (extensional) concept is referred to by 'the C-concept for "F" '. So the term demands a narrow scope reading in any true hypothetical context. Hence, so will 'the *F*-abstract of the C-Concept for "F" ' in say

> If under circumstances C the C-concept for 'F' would have an extension which no concept actually has, then the *F*-abstract of the C-concept for 'F' would not be one which any concept actually has.

And now, failing the production of any good reason why such a narrow scope understanding may not be extended to the occurrence of 'the *F*-abstract of non-self-identity' in

> No matter what circumstances C might be and no matter what objects did or did not exist in them, the *F*-abstract of non-self-identity would be the same as the *F*-abstract it actually has

the analogy with the Simple Argument at first-order is restored.

It thus appears that, in order to maintain any analogue of Dummett's indeterministic thesis with respect to higher-order abstractions, two restrictions will be necessary: only abstractions will be acceptable whose abstractive **(p.366)** relations generate no more equivalence classes of concepts than there are objects in the original domain; and any *necessarily empty* concept will be excluded from the range of the abstractive operator. Both those are restrictions which one can imagine a certain kind of nominalist temperament welcoming in any case. However, there are still further subtleties for the indeterminist thesis to negotiate. Clearly, no assignment to the abstracted terms in which '*F*(F)' receives a referent which *would not exist under certain conditions in which F retained its actual extension* can avoid confrontation with an adaptation of the argument just outlined. So for each F, any eligible referent for '*F*(F)' which is to participate in an interpretation which avoids the difficulty must be an *actual instance* of F. And that constraint will demand that the relation, ~, partitions the concepts into *exactly as many* equivalence classes as there are objects, since each unit concept—concept of a single object—must now be assigned to a different such equivalence class. Thus ~ cannot be 'deflationary' either—must not generate *fewer* equivalence classes than there are objects in the original domain. There are of course such relations.[49] But the position is that even with the empty concept banished, an extension of the Simple Argument still promises to block any interpretation of the terms introduced by a higher-order abstraction within the original objectual domain save in this one very special kind of —non-inflationary and non-deflationary—case.

If all this is correct, then it is, needless to say, a helpful contribution towards a solution to the Caesar Problem. Indeed we may conclude that—because one—one correspondence does not fit the narrow condition just described (even after proscription of the empty concept)[50]—there is actually no *Caesar* Problem for Hume's Principle, granted the manifest existential contingency of the conqueror of Gaul. But it is not a complete response. The Simple Argument says nothing about the relations between lines and directions, for instance, when lines are *not* construed as destructible objects; and indeed, whether first- or higher-order abstraction is involved, the whole basic strategy of argument we have been looking at in this section must fail to engage the issue of sortal distinction and overlap in any case where—at least for the neo-Fregean—*necessary* existents are involved. A purchase on the issues arising in such cases, if indeed possible at all, will require a different approach.

## (p.367) 5. On the Sortal Exclusion Principle of Frege's Conception of Numbers as Objects

When Frege insists that numbers are ‘self-subsistent’ objects, he does not of course mean that they are objects in the way in which doorstops or paperweights are objects. To be an object of either of these kinds is to be an object of *some other* kind first and then to play a certain additional role. Any kind of thing—of a certain size and weight—can be a doorstop or paperweight. But it cannot *just* be a doorstop, or paperweight: there must be some other basic sort of thing which it is, to qualify as which it does not have to *do* anything and being which is essential to it. Such underlying basic sorts are what are expressed by *sortal concepts*, properly so-styled. That the concepts introduced by Fregean abstraction principles rank, when all goes well, as sortal is the cardinal thesis of Frege's platonism as we understand it.

*Sortal concept* was a key notion in the approach to the Caesar Problem essayed in *Frege's Conception of Numbers as Objects*.[51] Sortal concepts are usually thought of as somehow intimately tied up with identity contexts. But both *doorstop* and *paperweight* are associated with significant identity contexts; the doorstop I had by my back door in 1994, for example, may or may not be the same as the one which my stepdaughter now has in her flat in Edinburgh. However, what marks those concepts off from sortals proper is that whether x and y are the same doorstop, or paperweight, depends first on whether they are both doorstops, etc., and then on whether they are the same F for some *other* concept F—as it might be: *brick*, or *small bone-china toad*, or *horseshoe*—which has this feature: that knowing what it is for x to be the same F as y does not in turn depend on knowing that x and y are Gs, for some G distinct from F, and knowing what it is for them to be the same G. Sortal concepts, strictly so regarded, are not merely those which make for intelligible contexts of identity—concepts for which ‘x is the same F as’ makes sense—but which make for identity contexts which are, in the sense just outlined, *epistemically autonomous*. And when Frege insists that numbers are objects, what he is best taken to mean is that Number is a (non-empty) sortal concept in this more technical sense.

The reflection that Number, and other concepts introduced by Fregean abstraction, are meant to be sortal concepts in the sense just gestured at places the Caesar Problem in a certain perspective. Grasping a sortal concept will, in the light of the above, involve having an appropriately autonomous understanding of relevant contexts of identity. But it must, of course, also involve understanding an *application-condition*: understanding what it takes for something to be (an) F in the first place. In some cases, it seems as though someone could grasp the latter, or nearly do so, while ignorant of the former. **(p.368)** For instance, someone could be adept at distinguishing Cyrillic letters from other signs, and from other things in general, or books from objects of other sorts, while innocent of, or confused about, the type–token distinction and so not clear what it would be for x to be the same (type, respectively token) letter or book as y. What seems to be going wrong with Hume's Principle, or the Direction Equivalence, conceived as purported explanations of sortal concepts of number and direction, is a kind of converse failing: such principles purport to explain a criterion of identity under the relevant abstract sortal F, but altogether neglect the more basic issue of the application-condition: of what it takes to be an F (and thereby do not really explain identity under F either).[52]

The leading idea of the approach of *Frege's Conception* was a direct response to this suggestion. That approach maintained that, contrary to appearances, abstraction principles *do*, at least implicitly, impose constraints upon the range of application of the concepts they purport to explain. Here is the basic motivational thought. Evidently, if any numbers *were* people, then not only would it be possible[53] to frame some true *mixed* identity statements—such as ‘Julius Caesar = 0’—linking terms ostensibly designating people with terms ostensibly designating numbers; it would also be possible to formulate pairs of (unmixed) identities ‘a = b’, ‘c = d’—in which ‘a’, ‘b’ recognizably purport reference to persons, and ‘c’, ‘d’ recognizably purport reference to numbers—whose truth-values should be determined by essentially the same factors. The idea might be expressed by saying that if any numbers are people, then—in a perfectly ordinary sense of ‘fact’—some facts about the identity or distinctness of people just are facts about the identity or distinctness of numbers, and what determines the one determines the other. Tigers are animals because facts about the identity and distinctness of tigers are all facts about the identity and distinctness of certain animals: whatever distinguishes a particular pair of tigers, or determines that they are identical, thereby *eo ipso* does the same for a particular pair of animals. In general, for arbitrary sortal concepts F and G: if any Fs are Gs, then there must be some ostensible F-identity questions and some ostensible G-identity questions which turn on exactly the same considerations. Some Fs are Gs only if certain F- and G-identity statements coincide in what determines them as true or false.

This was the admittedly very vague motivation underlying the treatment of the Caesar Problem offered in *Frege's Conception*, where it was proposed that a solution for the specific case of numbers might be had by appeal to the principle

> $N^d$ Gx is a sortal concept under which numbers fall (if? and) only if there are, or could be, singular terms ‘a’ and ‘b’ purporting to denote instances of Gx such that the **(p.369)** truth-conditions of ‘a = b’ could adequately be explained as those of some statement to the effect that a 1-1 correlation obtains between a pair of concepts.[54]

The leading thought was that whilst the truth-value of an identity statement linking terms purporting reference to a person or persons—such as ‘The murderer = Three-fingered Jake’—might, as a matter of fact and in a certain informational context, be determined by verifying the existence of a 1–1 correlation between a pair of concepts (e.g. if either Jake or the Butler did it and there are exactly as many fingerprints on the axe as fingers on Jake's left hand), it is clear enough that the condition for the truth of a statement of the former kind cannot adequately be explained by means of a statement of the latter kind. It would then follow, via $N^d$, that no numbers are persons. In more general terms, the idea was that if (any) numbers fall under a further, distinct sortal concept G, then this should be reflected in facts about the truth-conditions of certain identity statements overtly concerning certain Gs, viz. those Gs which some numbers are; that is, some identity statements connecting actual or possible G-terms ought to coincide in their truth-conditions with corresponding statements of numerical identity, and hence have truth-conditions expressible in terms of the condition which Hume's Principle lays down as necessary and sufficient for numbers to be the same (i.e. 1-1 correlation among suitable concepts). $N^d$ thus gives expression to the idea that a *criterion of application* for the sortal concept *number* is extractable from its associated *criterion of identity*.[55]

**(p.370)** The proposal is not to be confused with the attempt, effectively criticized by Frege himself in *Grundlagen* § 67, to wriggle out of the Caesar Problem by laying down that an expression refers to a number if (and only if) it—or a definitional equivalent—is introduced by means of (an instance of) Hume's Principle. That no numbers are people, and hence that Caesar is not the number of planets, is a consequence—not of Hume's Principle directly but—of $N^d$, together with the additional premiss that no identity statements linking person-denoting terms have their truth-conditions adequately explained as those of any statement of 1-1 correspondence among concepts. The *Frege's Conception* proposal thus bears on the kind of *sense* a range of terms possess, as reflected in what it takes satisfactorily to explain the contents of identity statements configuring them; and the way such terms are actually introduced is relevant only insofar as it bears on that issue about their sense. Further, $N^d$ itself is not conceived as a stipulation, but as a consequence of Hume's Principle together with a quite general principle concerning sortal inclusion (**SI**):

> No objects falling under a given sortal concept Fx can be reputable candidates for identity with objects falling under another sortal concept Gx, unless . . .: to every identity statement ‘a = b’, where a and b purport to have reference among the relevant instances of Fx, corresponds at least one statement ‘A = B’, where A and B purport to denote instances of Gx, which has the same truth-conditions.[56]

This latter principle is, again, not advanced as a stipulation. The root idea was—as sketched above—that if instances of one sortal concept are *in fact* instances of another, then the truth-values of certain identity statements linking terms for instances of the latter sortal must be constituted by, or owed to, the very same facts as determine the truth-values of appropriate identity statements linking terms for instances of the former. The general principle seeks to encapsulate that idea. It amounts to a substantial metaphysical claim about when the extension of one sortal may properly be regarded as included in—or as overlapping with—that of another.[57]

**(p.371)** Whether it is correct or not is very discussible, however. In the years since the publication of *Frege's Conception*, a series of objections have emerged. Some suggest that the principle $N^d$ is too generous—that it doesn't in the end even banish Caesar from the domain of numbers; others suggest that it, or the more general principle from which it flows, is too strong, proscribing the identification of Members of Parliament, for instance, with flesh-and-blood humans and imposing unwanted a priori resolutions of a large class of issues of sortal taxonomy. Here is Michael Dummett making both kinds of objection (without distinguishing them) in a single paragraph: the proposed solution to the Caesar Problem is vitiated, he suggests, because

> The determination of the truth-value of a statement of identity between numbers may well turn on the criterion of identity for human beings; the number of Dr Jekyll's cousins coincides with the number of Mr Hyde's because Dr Jekyll and Mr Hyde are one and the same person . . .

Thus as far as $N^d$ is concerned, Dummett is suggesting, Dr Jekyll could still be a number. On the other hand

> The oddity of [the *Frege's Conception*] position is that the legitimacy of a stipulation depends upon whether it follows or precedes another. The eccentricity of one ellipse coincides with that of another just in case they are similar. This follows from a definition of eccentricity as the ratio between the distance between the foci and the length of the major axis; but on [the present proposal] that definition would be illegitimate if we had first stipulated the condition for the eccentricity of two ellipses to be the same: the geometrical criterion would preclude the identification of the eccentricity with a real number, with which is associated a quite different criterion of identity.[58]

So—at least on one order of stipulation—eccentricities cannot be real numbers.

Is the *Frege's Conception* proposal too generous? Let's begin with the Jekyll and Hyde example. It would be a mistake to set it aside as another of the murderer/Three-Fingered Jake kind, that is: a case where issues of numerical and personal identity are merely contingently coincident. For, of course, it *follows from* the identity of Jekyll and Hyde that the number of their cousins is the same, just as it follows, for any function, $F$, including persons among its admissible arguments, that $F(\text{Jekyll}) = F(\text{Hyde})$. Nevertheless the actual example still misfires rather badly, for the converse entailment fails. It does not follow from a's and b's having the same number of cousins that a = b. So this is not a case of identity of truth-conditions (since however precisely that notion is best understood in order to give the *Frege's Conception* proposal a run for its money, mutual entailment will presumably be a necessary condition).

Still, Dummett's example does serve to point up a *potential* difficulty: if the function expressed by 'the number of x's such that Rx . . . ' on a domain **(p.372)** of what were intuitively non-numbers were necessarily one–one—as for 'Rxy' = 'x is a cousin of y' it is not—then the entailment *would* go in both directions. And in that case, statements of identity concerning the elements of that domain would all share their truth-conditions with—at least in the sense of being necessarily equivalent to—certain numerical identity statements; absent some further refinement of the notion of sameness of truth-conditions, $N^d$ would accordingly be impotent to distinguish the elements of that domain from the numbers in question.

Consider, for instance, this case. Let the domain to consist of the Roman Emperors from Julius Caesar to Constantine, and let Rxy be the relation: x is an imperial predecessor of y. And suppose that no two Emperors may constitutionally reign at the same time, so that it belongs to the concept of being Emperor at a certain time that no one else is. Then where a and b are Roman Emperors, it will follow from the identity of Nx:Rxa and Nx:Rxb that a and b are the same; and from the distinctness of those numbers that a and b are distinct. The truth-conditions of the two types of statement will thus coincide and, for all that $N^d$ has to say on the matter, Julius will be up for identification with the number of his predecessors, that is: with the number Zero, after all.

The immediate reply should be that the impression of mutual entailment here is an illusion, contrived merely by the stipulated background assumptions: that it is no part of the sense of any of their names, or otherwise of the essence of the Emperors just cited that they were ever Emperors. So the distinction between Tiberius and Galba, say, is quite consistent with the numbers of their imperial predecessors being the same, viz. 0.

A similar observation deals with a recent example of William Stirton.[59] Stirton proposes that, in a tight sense of sameness of truth-conditions which we shall canvass shortly, and for arbitrary objects, a and b

$$a = b$$

and

$$Nx : x = a = Nx : (x = a \vee x = b)$$

coincide in their truth-conditions and hence that no matter what sort of objects, F, we are concerned with, the sortal overlap condition attempted by $N^d$ will be mutually satisfied by F and Number. Again, the immediate reply should be that there is no mutual entailment in the first place. For take the case where a and b are distinct, contingent existents. Then in any world where a exists but b does not, the numerical identity will be true but the proposition that a = b will fail.

**(p.373)** Still, what if a and b are *not* contingent existents? In that case it will be necessarily true that a = b holds just when Nx:x = a = Nx:(x = a ∨ x = b) does. But a proponent of the *Frege's Conception* approach still has a decisive rejoinder. Even if mere mutual entailment constitutes the relevant sense of sameness of truth-conditions, there is a stronger constraint in the offing than those actually formulated in $N^d$ and its generalization, **SI**. Those principles merely demand that when number overlaps another sortal, or when any two sortals overlap, some identity statements under the one have the same truth conditions—however that is to be interpreted—as some identity statements under the others. But there is a clearly motivated stronger constraint—if the overall approach is motivated at all. What is required is not merely that there be some identity statements under the respective sortals with the same truth-conditions but, more, that the objectual identifications and distinctions respectively effected by the truths among those two sets of statements be *isomorphic*. The explicit proposal was merely that if (some) Fs are to be Gs, then there must be a range of F-identities, $f_i = f_j$ $f_i = f_k$, $f_j = f_k$, etc., with the same truth-conditions as certain G-identities, $g_i = g_j$ $g_i = g_k$, $g_j = g_k$, etc. But of course the underlying idea is that these respective ranges of identity statements should be dealing with *the very same objects* under two sets of guises. So the mapping, under sameness of truth-conditions, between the two ranges of statements should be such as to allow for a coherent scheme of cross-identification of the referents of the two sets of terms. And that condition the Stirton example signally fails to meet.

Consider for instance the true imperial distinctions:

Julius ≠ Tiberius

Tiberius ≠ Domitian
Julius ≠ Domitian

Under the Stirton translation—and prescinding now from the point about possible non-existence—these are correlated with

Nx: x = Julius ≠ Nx: (x = Julius ∨ x = Tiberius)
Nx: x = Tiberius ≠ Nx: (x = Tiberius ∨ x = Domitian)
Nx: x = Julius ≠ Nx: (x = Julius ∨ x = Domitian)

But now, which of these numbers is Julius? If we pick Nx: x = Julius, then Julius is also Nx: x = Tiberius—since those are both the number 1. So which number is Tiberius? Not Nx: x = Tiberius, since that number is Julius. Could he be Nx: x = Tiberius ∨ x = Domitian? If he is, he is the number 2. But in that case, since Nx: (x = Julius ∨ x = Domitian) is also the number 2, all three numerical distinctions become *de re* equivalent to the first imperial distinction, and no reference to Domitian is effected within them. So the isomorphism constraint is violated—and necessarily so, of course, since the **(p.374)** Stirton scheme only ever gives at most two objects to work with whereas we have three Emperors to locate in the realm of numbers.

It would, however, be an oversight to suppose that **SI** plus the isomorphism constraint copes with all cases. Consider the mutual entailment of

$$a = b$$

and

$$\{x : x = a\} = \{x : x = b\}.$$

Since the unit sets corresponding to a range of individuals are always exactly as many in number as those individuals, **SI** plus the isomorphism constraint will pose no obstacle to the identification of each object with a set, viz. its unit set—at least so long as the identity of truth-conditions to which it appeals demands no more than mutual entailment. Whether or not those constraints effectively give us a solution to the Caesar Problem for numbers, therefore, it would seem that the general problem goes deeper.

It may be supposed that what the example shows is that **SI** and its kin do indeed require formulation in terms of a stronger notion of sameness of truth-conditions than mere mutual entailment: a notion whereby the truth-conditions of ‘a = b’ and ‘{x: x = a} = {x: x = b}’ may be regarded as different even though necessarily satisfied or unsatisfied together. What might that notion be? An interestingly tight notion of sameness of truth-conditions is presupposed by Frege's imagery in *Grundlagen* § 64 where, as he suggests, the transition from right to left across an (instance of an) abstraction principle involves ‘carving up’ *one and the same content* in a new way. We may, he tells us, conceive the very judgement that a pair of lines are parallel *as* an identity statement, and thereby ‘obtain the concept of direction’. Clearly what is required if such an image is to be apt is more than mere necessary coincidence in truth-value—if only because it would be implausible to claim that one of any arbitrary pair of necessarily true statements effects a ‘recarving’ of the content of the other. On the other hand, precisely because a prior understanding of the right-hand side of an abstraction is intended to furnish a route to forming the concepts involved in a grasp of the left-hand side, it cannot be required that the two statements coincide in *sense* (or one who understood the right-hand side would already possess the concepts involved in understanding the left-hand side). So what is the right way to think about the matter?

One way to circumscribe a suitably intermediate notion of truth-condition, serviceable for Frege's purposes, is to stipulate that two statements share their truth-condition just in case a thinker who understands both of them is thereby in a position to tell, without determining the truth-value of either statement separately,[60] that they must be alike in truth-value. As it stands, **(p.375)** this is vague in a crucial respect—it allows that the thinker may employ (deductive) reasoning to reach the conclusion that the two statements must be alike in truth-value, but says nothing about what kinds of deductive inference are permitted. If no restriction is imposed—if, say, the thinker is allowed to reason via any classically valid entailments—then pairs of statements of the forms ‘P’ and ‘(P & Q) ∨ (P & ¬ Q)’ (and more generally, pairs in which one member is, so to speak, a logical complication of the other, importing arbitrary redundant constituents) will qualify as having the same truth-condition. One way to avoid this, presumably undesirable, consequence would be to restrict the thinker to reasoning via *compact* entailments, where a compact entailment is one in which the premiss includes no ‘passengers’—constituents which do not contribute towards the holding of the entailment. More precisely, A compactly entails B iff (i) A entails B and (ii) for every non-logical constituent E of A, there is some uniform substitution E′/E such that A[E′/E] does not entail B.

It is clear enough that if sameness of truth-condition is explained in this way,[61] the left- and right-hand sides of Fregean abstractions will coincide in truth-conditions while differing in sense—in the conceptual repertoire exercised by an understanding of them. So the notion does appear a promising candidate to underwrite Frege's purposes in *Grundlagen* § 64.[62]. But as the reader will immediately see, this notion of sameness of truth-conditions is quite useless for the present purpose since both legs of the mutual entailment between

$$a = b$$

and

$$\{x : x = a\} = \{x : x = b\}$$

are compact! If not everything is its unit class, more is demanded, it appears to underpin our knowledge of the fact than what can be elicited from **SI**, even with sameness of truth-conditions interpreted in terms of (ancestral) compact equivalence, even with the supplementary constraint of isomorphism in attendance.

However, it need be no very serious matter should **SI** indeed be too weak to exclude certain unwanted sortal overlap hypotheses provided it, or its specialization, $N^d$—supplemented by other considerations like those reviewed in Section 4 or those concerning isomorphism sketched above—provide sufficient strength to solve the immediate problem: that of determining **(p.376)** sharp, or sharp enough, limits to the extension of the concept of Number introduced by Hume's Principle. But the complementary line of objection is quite another matter. If the principles invoked are intuitively too *strong*—if they serve to outlaw what are evident cases of sortal overlap or inclusion—then naturally we can place no confidence in any boundaries they may be argued to impose on Number.

That the *Frege's Conception* proposal is indeed too strong was the gist of Dummett's second example, about ‘eccentricities’ and real numbers. Consider, where e and e′ are any ellipses and $E$ is an operator forming terms standing for eccentricities, the possible Fregean abstraction principle:

(Eccentricity Equivalence)

$$E(e) = E(e') \leftrightarrow \text{Similar}(e, e').$$

Dummett's point, recall, was that if this ‘geometrical’ criterion of ellipse identity were adopted, then—according to the *Frege's Conception* proposal—we would not, in view of the difference between the geometrical criterion and the criterion of identity for real numbers generally, be at liberty to identify the eccentricity of an ellipse, e, with the ratio between the distance between the foci and the length of e's major axis. Yet, he implies, that natural stipulation ought surely to remain an option. On the other hand, were we *first* to stipulate that *E*(e) just *is* the relevant real number, the aptness of the geometrical criterion would follow. Hence the allegedly bizarre upshot that ‘the legitimacy of a stipulation’—that identifying eccentricities with the appropriate real numbers —‘depends upon whether it follows or precedes another’—that of the Eccentricity Equivalence.

In appraising this criticism it is as well to distinguish the allegation about stipulations pre-empting each other or not according to their order of introduction from the claim that stipulation of the Eccentricity Equivalence forecloses on the identification of eccentricities with reals. The former is open to the simple counter that if we did first stipulate that eccentricities are the appropriate real numbers, then it would not thereafter, contrary to Dummett's implicit suggestion, be open to us to adopt the Eccentricity Equivalence as a *stipulation*; rather that biconditional would then be a conceptually substantial *theorem* depending on a background context of geometry and real number theory in which would be no option but to accept it. So the real issue raised by the example is only whether if the stipulation, properly so regarded, of the Eccentricity Equivalence would indeed, by the lights of the *Frege's Conception* proposal, pre-empt identification of eccentricities with real numbers, that would be any serious objection to that proposal.

The answer is that it would not. If we lay down the Eccentricity Equivalence as a Fregean abstraction, we in effect explain the truth-conditions of statements of eccentricity-identity as a special case of those of statements of **(p.377)** identity of shape for plane figures. But shapes are not real numbers. If they were, then different shapes would be identified. For instance, let t and t′ be isosceles triangles. We might identify the *sharpness* of t with the ratio of the length of t's height to its base. But sharpnesses too, if introduced by an appropriate abstraction on similarity, would be a kind of shape. If we then went on, having so introduced them, to identify them with the appropriate real numbers, we would thereby be forced to *cross-identify* them with appropriate eccentricities, and thereby absurdly to identify the shapes of figures as dissimilar as ellipses and triangles.

In other words: you have to decide what concept of eccentricity, etc., you want. Dummett is quite right that if it is fixed by the abstraction on similarity, then we are not at liberty to identify eccentricities with ratios. But that, rather than an anomalous consequence of the proposed way with the Caesar Problem, is actually entirely as it should be. The shapes of ellipses, and isosceles triangles, are indeed not real numbers. Sure, they are open to real *measurement*: they allow of *representation* by a real number, but that is not the same thing —as the simple antinomy just outlined shows. Of course we are free to go on to *refashion the concept* of eccentricity so as to allow the identification of its instances with reals. In that sense the latter stipulation is *not* precluded by the former. This may be theoretically convenient in all sorts of ways. But then the Eccentricity Equivalence no longer gives the criterion of eccentricity-identity but rather, as stated, will be a conceptually substantial consequence of it.

In sum, Dummett is wrong to imply that the identification of eccentricities with reals leaves us free to stipulate the Eccentricity Equivalence and, while right to imply that the latter stipulation would—according to the spirit of the *Frege's Conception* proposal—preclude conceiving of eccentricities as reals, wrong again to suggest that there is any anomaly in that. On the contrary, we already think of shapes—and lengths, and volumes, and weights—as *different from* but *measurable by* reals. For certain theoretical purposes we may represent the former by the latter—and thereby, in whatever is the attendant sense, ‘identify’ them. But that is not a serious ontological sense of ‘identify’—if it were, we'd have the simple antinomy.[63]

**(p.378)** The suspicion that the *Frege's Conception* approach relies upon an unacceptably strong requirement for sortal overlap has been pursued independently by Peter Sullivan and Michael Potter,[64] who offer the following counter-example. Suppose we lay down the following principle, by (purported) analogy to the Direction Equivalence, etc.:

(Members of Parliament)

The MP of x = the MP of y iff x and y are co-constituents.

*Members of Parliament*, their suggestion is, is an unobjectionable principle. But the criterion if identity for MPs so introduced is quite different from that (complex) criterion—whatever it is—which governs personal identity. No statement of personal identity could have its truth-condition adequately explained in terms of the obtaining of co-constituency relations among people. So it should follow, via an appropriate analogue of $N^d$—at least if we take *Members of Parliament* as giving a criterion for MPs to be identical—that MPs aren't people!

**(p.379)** The counterpart of the response just given to Dummett would be to accept that conclusion and argue that it is benign. But it may seem that there is no option of that move this time; for MPs *are* people—the two concepts do overlap, period. So a principle which requires that they don't is simply unsound. However, this thought is less compelling than may at first appear. If the conclusion that MPs aren't people impresses as plainly unacceptable, that is because we are tacitly drawing on some *independent* conception of MPs, according to which they are flesh and blood individuals who play a certain political role. We are, therefore, not treating the principle *Members of Parliament* as serving to introduce the concept MP in the manner in which Fregean abstractions such as the Direction Equivalence and Hume's Principle are intended to introduce the concepts of direction and number.

To elaborate. The objection misfires because it fails to take seriously what is involved in introducing a sortal concept S by means of a Fregean abstraction

$$s(a) = s(b) \text{ iff } Eq(a, b).$$

The crucial point is that the content of any identity statement linking S-denoting terms s(a) and s(b) is to be the *same* as (no more and no less than) that of a statement to the effect that a and b stand in the specified equivalence relation Eq. So no more (and, of course, no less) is to be required for the existence of s(a) and s(b) than that there exist Eq-related entities a and b. It follows that when the abstraction is first order and, more specifically, proceeds in terms of an equivalence relation on *concrete* objects, 's(a) = s(b)' must demand the existence of no concrete objects beyond a and b themselves—or more precisely, none beyond those which are required for the truth of 'Eq(a, b)'. But if *Members of Parliament* is taken as providing a criterion of identity for Members of Parliament, conceived of as flesh and blood individuals, the truth of its right-hand side will *not* be conceptually sufficient for that of its left—if only, for instance, because a parliamentary constituency may obviously be temporarily unrepresented as a result of death or resignation. So construed, the truth of (instances of) *Members of Parliament's* left-hand side would—save in the special case where x or y happens to *be* the relevant Member—call for the existence of concrete objects beyond what are required for x and y to be co-constituents. But then *Members of Parliament* would not, for the reason given, be serving to introduce a concept by Fregean abstraction—any concept so introduced is bound to be one under which only abstract objects fall.[65] Thus if, in particular, we introduce a concept, Member of Parliament, by the suggested stipulation, we have no right to complain that, by the lights of **SI**, Jones's MP cannot be Smith, or anyone else—for MPs so introduced are *abstracta*, while people, presumably, **(p.380)** are not. Putting the point the other way round, if we want a concept of MP—our usual one, maybe—that leaves room for Jones's MP to have a spouse, an overdraft, and a flat in Kensington, that constrains us precisely *not* to introduce it in this way.

In sum, the objection founders on a dilemma. If *Members of Parliament* is intended as a genuine abstraction, purportedly introductory of a sortal concept and fully comparable in all relevant respects to the Direction Equivalence, *et al.*, then it is no paradox to deny that MPs are people and no objection to the *Frege's Conception* approach if it carries that consequence. On the other hand, if the principle is offered as analytic of the concept, Member of Parliament, as ordinarily understood, then, first it is arguably not even a truth as it stands—it needs an existential proviso to cover the case of temporarily unrepresented constituencies, etc.; and second, it in any case remains that the concept it governs is not *sortal* but—in the sense illustrated at the beginning of this section—*functional*: to be an MP is not to be a certain basic sort of object but is rather to occupy a certain kind of functional role (and hence to be some other kind of object first and foremost). But if *Members of Parliament* does not introduce a sortal concept (in the relevant sense), then it is not properly subjected to principles constraining overlap and inclusion among such concepts.

These reflections suffice to neutralize the purported counter-example. But they are insufficient to dispose of a more general worry which lies close by, and which may charitably be seen as underlying Sullivan and Potter's concern. The condition which **SI** lays down as necessary for a pair of sortals C and D to overlap is, in effect, that there be some range of terms purporting reference to instances of D such that where d and d′ are any two of these terms, there is some true identity statement, ‘c = c′’ whose terms purport reference to Cs and which shares its truth-condition with ‘d = d′’. But, on anything like the conception of sameness of truth-condition with which we are working, the questions whether or not ‘c = c′’ shares its truth-condition with ‘d = d′’ will seemingly always be determinable a priori, on the basis merely of an understanding of the two statements and (maybe compact) reasoning. The worry, accordingly, is that anything broadly in keeping with the *Frege's Conception* approach—anything that might pass for a refinement of it—is bound to make for *unwanted a priori resolution* of questions of sortal inclusion or overlap—unwanted, because even if some such questions ought to admit of armchair decision, there are plenty of others which ought to be resolvable only a posteriori, through empirical, scientific investigation. It is, for instance, not a priori that whales are not fish, that chimpanzees are not persons, or that the Stepford Wives—if that may pass as a sortal—are not humans.[66]

**(p.381)** Let us look at this worry in more detail. To begin with, there is room for manœuvre on two points. **SI** imposes a necessary condition on sortal overlap which it had better be possible to determine a priori is not met *in certain cases*—for instance, in the case of persons and numbers. It has to do that much or it cannot assist a purely philosophical solution of the Caesar Problem. But why should it be thought to be implied that *whenever* two sortals do not overlap, that fact can be recognized a priori? To get to that strong claim, two supplementary premisses are required: (i) that two sortal concepts do not overlap if and *only if* the **SI** condition is not met—if and *only if* no ‘c = c′’ and ‘d = d′’ correspond as required; and (ii) that whenever no appropriate identity-statements so correspond, it can be recognized a priori that that is so. Since any doubt that the *Frege*'s *Conception* approach can deliver a *sufficient* condition for sortal overlap/inclusion—and we closed our discussion of that on a note of doubt—is a doubt about (i), it would seem that a critic who wanted to press the present concern had first better help us over the former.

But suppose in any case that (i) is made good. What is the ground for thinking that (ii) holds—that if a pair of sortals, C and D, give rise to no identity statements, ‘c = c′’ and ‘d = d′’, specified as above, whose truth-conditions are the same, that fact will always be recognizable a priori? Certainly, it is in accordance with the type of notion of sameness of truth-conditions with which we are working that whether or not a *particular* ‘c = c′’ shares its truth-condition with a *particular* ‘d = d′’ will be determinable a priori just on the basis merely of a knowledge of their respective truth-conditions and reasoning. But it cannot directly follow that if *no* such identity-statements share their truth-conditions, that too will be determinable a priori; trivially, that a relation—here, sameness of truth-conditions—is a priori decidable in any particular case does not entail that the claim that no two items stand in that relation is knowable a priori whenever true.

But presumably an objector can jump that gap. Our entitlement, if we have it, to the belief that the statement that Caesar = the author of *De Bello Gallico* shares its truth-condition with no statement of numerical identity relies on nothing special to the particular example. It is not as if we might know a priori that

$$\text{Caesar} = \text{the author of } \textit{De Bello Gallico}$$

differs in truth-condition from, say,

$$\text{The number of knives in the drawer} = \text{the number of forks on the table},$$

but in such a way as to leave it open whether the personal identity statement is alike in truth-condition with *some other* statement of numerical identity! If that is the best we could do, **SI** would be incompetent to deliver any purchase on the issue of overlap between person and number. **SI** can deliver a negative verdict on that issue only if the distinction in truth-conditions between the **(p. 382)** two ranges of identity statement is both *systematic* and *available to general reflection*. The crucial question might therefore seem to be this: does it make sense to suppose that a systematic distinction in the truth-conditions of two relevant ranges of identity statement might—depending on the choice of the sortal concepts concerned—sometimes be available to general a priori reflection and sometimes demand additional a posteriori information before it can be recognized?

That might *appear* to be the crucial question; and it might seem very hard to see how an affirmative answer could be justified, at least so long as the relevant notion of sameness of truth-conditions is no weaker than mutual entailment. But there is an ambiguity. When it is asked whether it is knowable a priori that there is a systematic difference between two ranges of truth-conditions, the scope of 'a priori' may or may not include an appreciation of *what those truth-conditions are*. What may plausibly be taken to be an a priori matter is a systematic difference between truth-conditions *once known*. But it may or may not be an a priori matter what are the truth-conditions of a given range of identity statements. And the cases where it is plausibly a posteriori whether (any) Cs are Ds are all, we suggest, cases where this is not so—cases where it takes empirical investigation to determine (one or the other of) the criteria of identity involved. If a concept is set up by contextual stipulations of the Fregean sort, it will indeed be given with its very explanation what determines identity under it. But many concepts—including, importantly, (putative) natural kind concepts—are not set up in this way, or in any other way which renders it an a priori matter what should canonically determine that x is the same F as y.

Suppose, for example, that we have come across some smallish, hair-coated quadrupeds, very like groundhogs on casual inspection, but different enough in certain details to make us conjecture they are a new mammalian species. (Suppose we haven't gone in for any dissection, or had opportunities to observe them breeding.) We dub them 'gridhogs'—the name suggests itself to us by their curious tendency to pass the nights huddled close to electricity substations. We certainly don't know *a priori* that they are mammals; that is, that 'is the same gridhog as' may be determined in the same way as 'is the same mammal as'. Our belief that this is so has no better basis than their general appearance. Suppose, in fact, we are quite wrong about this: that gridhogs are little robots which have been secretly introduced into the countryside by the Nature Conservancy Council to monitor wildlife activity. Their night-time habits might have aroused our suspicions—they are in fact replenishing their internal batteries by means of an ingenious device which enables them to draw power from high-voltage sources without direct contact. Their criterion of identity is thus quite different from that appropriate to mammals—to do with the physical integrity of their CPUs and maybe program architecture. Our belief about the truth-conditions of gridhog-identity statements is thus quite mistaken, but there's nothing a priori about our **(p.383)** mistake—we might (epistemically speaking, at least) have been right, just as we are right in taking groundhog identity statements to be decided by the same considerations as decide mammalian identity statements, although there's nothing a priori about that either.

In sum: to allow that, on certain assumptions about what the truth-conditions of C- and D-identity-statements are, it can be determined, without further empirical investigation, whether they are systematically different is quite consistent with holding that it cannot always be determined a priori, just by reflection on the way a (putative) sortal concept is introduced. What the truth-conditions of identity statements under it are. The cases where sortal overlap and inclusion are a posteriori matters are all cases where this is not so: typically, where a concept is fixed by reference to paradigms or a stereotype so that a (defeasible) explanation of criteria of application for it *precedes* knowledge of its criteria of identity.[67] If that is right, then the worry that **SI**—or any sortal overlap principle like it in working with a notion of sameness of truth-conditions at least as strong as mutual entailment—must just on that account lead to unwanted (negative) a priori verdicts is misconceived.

At this point, it might appear that the *Frege's Conception* solution to the Caesar Problem is very much intact. Unfortunately, we have come to think that that is not so: there is a better objection than any so far considered—one which shows that the condition on sortal overlap imposed by **SI** is indeed unsatisfied by certain genuine cases of sortal inclusion.

In which case, of course, and if nothing further is said, the arguable failure of person—or any other particular concept of what ought intuitively not to be numbers—to pass the test vis-à-vis number matters not a jot.

In order for any C to be a D, **SI** requires that at least one suitable statement of C-identity, 'c = c′', shares its truth-condition with some suitable D-identity, 'd = d′'. In responding to the previous objection, we made space for the possibility that someone may have followed the usual explanation of C, or of D, and yet not know—because it is an a posteriori matter—what the truth-conditions of such identity statements are. Still, it remains that if **SI** is right, and if C and D do overlap, then *provided someone knows the criteria of identity for Cs and Ds*, she should be able to tell just by reflection that a suitable pair of what are in fact true statements, 'c = c′' and 'd = d′', must be alike in truth-value. And this requirement seems just too strong.

Why? Well, let C be *horse* and D be *mammal*. And suppose that Jack paid £2,000 at Appleby Fair last September for a horse that is now grazing in that field. Horses are mammals. So **SI** requires that suitable horse-identities share **(p.384)** their truth-conditions with suitable mammal identities. Suppose we select as our D-terms 'the mammal which is grazing in that field' and 'the mammal for which Jack paid £2,000 at Appleby Fair last September'. It seems clear that someone might be fully apprised of the criteria of identity for mammals, know that the criteria of identity for horses are the same as those for mammals, and understand the true identity statement

> (a) The mammal which is grazing in that field = the mammal for which Jack paid £2,000 at Appleby Fair last September

without there being any horse-identity statement of which she can tell a priori that it must be alike in truth-value with (a). For *what* horse-identity might turn the trick? There is no better candidate, surely, than:

> (b) The horse which is grazing in that field = the horse for which Jack paid £2,000 at Appleby Fair last September.

But this—best case—will not do. Even for one fully apprised of the criteria of identity for horses and mammals there can be no telling a priori that (a) and (b) are alike in truth-value, for the simple reason that there can be no a priori ground to exclude the possibility that the beast to which the terms in (a) refer is a prize bull! Indeed, one cannot even tell a priori that the converse conditional holds: that if (b) is true, (a) is—at least not if all one has to go on is an understanding of the two sentences, a knowledge of the respective criteria of application of horse and mammal, and a knowledge of their respective criteria of identity. Certainly, one will then know that the latter are the same. But horses and more generally mammals are kinds of *animal* and their criteria of identity are plausibly just those common to all animals. Since reptiles, for instance, share those, a knowledge of the criteria of application and identity of horse will be a priori quite consistent with horses being reptiles.

It is true, of course, that it has not strictly been shown that *no* horse- and mammal-identities correspond in the required way —we have drawn a blank with (a) but that is only one example. However, that is no comfort. The mammal in our hypothetical scenario is indeed a horse so—if the *Frege's Conception* proposal is broadly sound—it is one of the objects whose identifications and distinctions qua mammal should appropriately correspond to its identifications and distinctions qua horse. If the ways of referring to it involved in (b) and (a) do not generate the required correspondence in truth-conditions, it is incumbent on a proponent of the *Frege's Conception* proposal to nominate other ways of referring to it which fare better in that regard. But it seems wholly obscure how to accomplish that.[68]

**(p.385)** The moral to be drawn, it seems, is that sortal inclusion simply does not in general imply satisfaction of **SI**, or of any similar principle which requires that suitably constructed C- and D-identity statements will share their truth-conditions in any a priori detectable sense. One reason is because it has in general to be a posteriori what (narrower) kind of D any particular D-term denotes. Another is because even when it is given that someone knows what the criteria of application for C and D are, and that those concepts share criteria of identity which are known to her, she need so far have no basis for the conclusion that they overlap rather than are exclusive subconcepts of some more inclusive sortal.

So **SI** and its ilk must be abandoned. Can anything be rescued from the rubble?

## 6. The Proposal Reconceived

The basic idea underlying the *Frege's Conception* proposal was that there are at least some restrictions on the extension of any given sortal concept which are mandated by the type of canonical ground—*criterion*—associated with identity statements concerning its instances: specifically, that because they are canonically decided by reference to inequivalent considerations, statements of numerical and of, say, personal identity must be reckoned to concern different categories of objects. The final lesson of the previous section is that this *grundgedanke* does not allow of a satisfactory exegesis simply in terms of difference between the truth-conditions of relevant identity statements. With hindsight, it is perhaps obvious that this direction was destined for failure—at least so long as coincidence in the truth-conditions of any pair of statements is taken to require that speakers who understand them both can thereby know that the associated biconditional is true. For in the case of a pair of identity statements, such knowledge—that a = b if and only if c = d—is bound in general to be conditioned not just by an understanding of the sort(s) of objects concerned and the associated criteria of identity but also by the modes of presentation effected by the ingredient singular terms. Any criterion for sortal overlap which requires suitable such biconditionals to be knowable a priori, is therefore going to go past the *grundgedanke*—which laid emphasis merely on the criteria of identity concerned—into territory where hostages are offered to whatever a priori relationships do or do not obtain between the senses of terms involved on different sides of the relevant biconditionals: hostages which, as the horse/mammal example showed, are not in general redeemed. A developed version of the *Frege's Conception* proposal which did not offer such hostages would focus on a condition for sortal overlap which required not *coincidence* in the truth-conditions of (certain) identity statements under a pair of sortals but merely *restrictions on the allowable sources of their difference*.

**(p.386)** We continue to have confidence in the *grundgedanke* and in the natural philosophical ontology of which it is an integral part. Fully developing this ontology would involve detailed exploration of a range of associated principles and notions: *sortal concept* itself, *category, criterion of identity, object, property*, and so on. Such a development is well beyond the scope of the present paper, but in this section we will offer some preliminary outlines, or reminders, of certain aspects of these notions which are relevant to the present context, and give just the barest sketch how a solution to a wide class of instances of the Caesar Problem may be facilitated thereby. It will be the task of the final section to add detail and respond to some objections. Perhaps it is too much to hope finally to silence a resourceful opponent who is really determined to hang on the problem—philosophers (rightly, for the most part) hate to lose a problem. But the effect ought to be at least to clarify how unattractive the options are for such an awkward customer.

We begin by returning to the notion of a sortal concept. A contrast between sortal concepts and concepts in general is endorsed, under this or other terminologies, by much recent writing.[69] But the usual—somewhat indefinite—understanding of the notion is much more general than that briskly introduced earlier. According to the usual understanding, a sortal concept is any concept expressible by a common noun or noun phrase—thus calling, in contrast with adjectives and adjectival phrases, for the indefinite article in contexts of predication (‘NN is a . . . ’)—whose sense is such as to subserve a determinate question of number (‘How many . . .'s are there (meeting such-and-such further condition)?’). The latter point underlies the terminology of ‘count nouns’ widely used for expressions for sortals and implicitly demands the connection between sortals and contexts of identity emphasized at the start of Sect. 5.[70] For in order for questions of the form, ‘How many Fs are there?’, asked of some specified range of entities, to have determinately correct answers, it seems necessary and sufficient that it be determinate not merely which things in that range are and are not Fs but also, among the specified Fs, which are the same as or distinct from each other. By contrast, it is the lack of (the right kind of) association with contexts of identity of the concepts expressed by noun phrases such as ‘yellow (thing)’, ‘large (object)’, and so on, which may make for indeterminacy in questions like, ‘How many yellow things are there (in this bowl)?’ The question requires some antecedent understanding of what kind of yellow thing—bananas, maybe—is relevant to its intent. In a case where no such specification is given and no understanding is determined by the context, an expression of the form, ‘the number of Fs’, need have no specific reference.

This is just an corollary of the Frege's famous insight in part III of **(p.387)** *Grundlagen* that statements of number are predications of—i.e. are relative to—a concept:

> While looking at one and the same external phenomenon, I can say with equal truth both ‘It is a copse’ and ‘it is five trees’.[71]

Since both copses and trees may be correctly described as ‘green’, the question, ‘How many green things are there over there?’ takes on the same relativity, and is disambiguated only by an understanding that it is green Fs that are relevant, for some appropriate sortal F.

A very wide class of concepts meet the characterization thus outlined. Each of *person, tree, river, city, number, caterpillar, sapling, tadpole, man with an ice cream, brown cow, river longer than 1,000 miles, tiger with a thorn in his foot, Member of Parliament, kidney*, and *doorstop* is at the service both of assessable claims of the form, ‘There are *n* . . .'s (meeting such-and-such a further condition)’ and assessable claims of the form, ‘a is the same . . . as b’. But each of four importantly different subspecies of ‘count noun’ is illustrated in the list. The most basic case are those we characterized simply as sortals earlier and what we may now term *pure sortals*. For an object to fall under a pure sortal concept is for it to be a thing of a particular generic kind—a person, a tree, a river, a city, or a number, for instance—such that it belongs to the essence of the object to be a thing of that kind. If an object is an instance of a pure sortal concept, then it is so necessarily and thus could not survive ceasing to be so. Pure sortals thus contrast with what Wiggins[72] calls *phase sortals*—caterpillar, sapling, tadpole, and so on—where we precisely allow for one continuing object to survive the transition from one phase to another. They also contrast with *impure sortals*, formed by restriction of a pure sortal by some further inessential characteristic—man with an ice cream, brown cow, river longer than 1,000 miles, tiger with a thorn in his foot. In general, an object may discard, so to speak, an impure sortal without risk to its survival, merely by ceasing to satisfy the relevant restriction. And distinct from all these three groups, finally, are what we earlier called functional concepts—Member of Parliament, paperweight, and doorstop—to instantiate which is to occupy a certain role, which is inessential to the object concerned, but which may restrict the range of (pure) sortals an occupant may exemplify.

Naturally there is much more to say about these proposed distinctions. There are concerns about circularity in the rough characterization of count nouns,[73] issues about the stability and comprehensiveness of—and how best to elaborate on—the contrasts drawn, and questions about borderline cases (is *kidney* functional or sortal? How should artefact-concepts, like *statue* and **(p.388)** *screwdriver*, be classified?). But all we depend on in what follows is the viability of the notion of a pure sortal and its contrast with functional concepts in the crucial respect noted earlier: that where F is a pure sortal, 'x is the same F as y' will be epistemically autonomous—will, in appropriate instances, be assessable as determinately true or false independently of any tacit understanding of what further kind of thing x and y may be.

How in general is the truth of such an identity statement to be determined? Whether or not the Identity of Indiscernibles offers a correct *analysis* of identity, it is clear that the idea of items' sharing all their (bona fide) properties cannot possibly serve as a *working standard* of identity. First, it would hardly ever be possible to determine, property by property, that such was indeed the case. Second, since objects generally can survive *change*, some independent demarcation is needed of which properties an object can shed and which are such that their loss would constitute its demise—would suffice for non-identity of the stages before and after, as it were, in accordance with Leibniz's principle; and that in turn calls for some independent determination of what constitutes the survival of the object, so that we can then say that properties discarded in the process merely reflect changes in the object in question. (That distinction determined, we can then save the letter of Leibniz's account by eternalization of the properties in question—thus for instance I always have the property, so long as I survive, of being 5 feet 6 inches tall on my thirteenth birthday.) There is thus a need for an understanding, supplementary to Leibniz's account, of how to determine identity—a need for, precisely, to use the normal term of art, *criteria of identity*.

Pure sortal concepts, then, are distinctively associated with identity contexts whose understanding demands no further conception of what sort of thing the object(s) in question may be; and such contexts stand in need of criteria if they are to be appraisable. But it is salient that such concepts often import *local* such criteria—local standards, specific to the sort in question, for the appraisal of identity statements or, less verificationistically, local conceptions of that wherein the truth of such statements consists. That is, the kinds of consideration which are treated as *primitively* (canonically) determining (that is, as determining without the support of additional, contingent evidence of their relevance) the identity of, for instance, Mt. Everest with Gaurisankar, the Congo river with the Lualaba, that swan with the one I ringed yesterday, this ship with Theseus' ship, this tree with the cedar which Queen Anne planted, or the identity of persons, numbers, sets, shapes, and so on, vary as a function of the covering sortal—*mountain, river, ship, tree, person, number*, etc.—involved. All pure sortal concepts are in this way associated with their own specific criteria of identity: what counts as establishing—or wherein consists—the truth of an identity statement depends upon what (pure) sort of thing the claim of identity concerns.

It is clear that distinct sortal concepts often share an association with the same criterion of identity, differing only in their criteria of application. Such **(p.389)** for example is the situation of *tiger* and *tiger with a thorn in its paw*. Both in turn share their criterion of identity with *cat* or more generally with *mammal* or more generally still with *animal*. A *category* may usefully and naturally identified with a *maximally extensive sortal*: that is, a sortal concept, C, all of whose sub-sortals share the same criterion of identity and such that any object to which C does not apply must fall under sortal concepts not associated with that criterion of identity.[74] Note that it is a consequence of this characterization that a sortal concept can be subsumed only by one category.

It will already be evident how the framework supplied by these notions—sortal concept, criterion of identity, category—promises to facilitate the intended consequence: that the criterion of identity associated with a sortal concept constrains its extension in a fashion which is helpful with the Caesar Problem. It is an essential feature of an object that it falls under any particular pure sortal which it does fall under. Moreover any (pure) sortal concept is partially individuated by its criterion of identity[75] and so is necessarily subsumed under whatever category does subsume it. So it is essential to an object to belong to whatever category subsumes any pure sortal it falls under. Thus, granting that Number and Person are each pure sortals, numbers essentially belong to whatever category subsumes Number and persons essentially belong to whatever category subsumes Person. Those are clearly different categories since the first will be associated with a criterion of identity which necessarily coincides in application with the criterion stipulated in Hume's Principle while the second will be associated with a criterion of identity which necessarily coincides in application with whatever it takes to determine personal identity. And those are evidently independent criteria: you cannot *necessarily and in general* assess issues of personal identity by reference to considerations of one–one correspondence between concepts. That was the *Frege's Conception grundgedanke*. And now we can see the outline of a philosophical ontology in which it will have its intended significance. It is the outline of a world in which all objects belong to one or another of a smallish range of very general categories, each of these subdividing into its own respective more or less general pure sorts; and in which all objects have an essential nature given by the most specific pure sort to which they belong. Within a category, all distinctions between objects are accountable by reference to the criterion of identity distinctive of it, while across categories, objects are distinguished by just that—the fact that they belong to different categories. It is surely because we already **(p.390)** inchoately think in terms such as these that it strikes is as just obvious that Caesar is no number.

## 7. Caesar Laid to Rest: Which Things Numbers Could not Be

This gives us in effect a *recipe* for a solution to the Caesar Problem. But there are a number of points which require further explanation, and a number of objections to which at least a preliminary response is called for.

First and foremost, more needs to be said about the notion of a criterion of identity: sortal concepts are individuated by their respective associated pairings of criteria of identity and criteria of application; categories were explained as maximally inclusive sortals under a given criterion of identity, and the case that persons, or anything else, are not numbers is entirely dependent upon the alleged difference between their respective criteria of identity. The notion is thus carrying great theoretical responsibility in our discussion. It needs to be made as explicit as possible; and the route from the inequivalence of the respective criteria of identity for numbers and persons to the conclusion that they cannot belong to a common category needs a more contoured description.

When we are dealing with Fregean abstracts, it is natural to think of the criterion of identity in any particular case as the condition encoded on the right-hand side of the relevant abstraction: so the criterion of identity for directions is parallelism of associated lines, that for numbers one–one correspondence of associated concepts, and so on. By contrast, with non-abstracts the criterion of identity, when we can actually state it, will normally be the obtaining of a specific equivalence relation whose domain is the very objects in question; spatio-temporal continuity, for instance, is a relation whose domain is the very things for which it may serve as the criterion of identity; psychological continuity—were it a necessary and sufficient condition for personal identity—is a relation whose domain is persons; and so on.

Somebody might be tempted to try to massage this formal point into a fast track to the conclusion that no sortal concept of Fregean abstracts can share its criterion of identity with any other kind of sortal. (The criterion of identity for Fregean abstracts is always an equivalence relation on other kinds of thing; the criterion of identity for non-abstracts is always an equivalence relation on those very objects.) But that would be misconceived: if

$$S(a) = S(b) \leftrightarrow Eq(a, b)$$

is an abstraction, then the condition for the identity of Ss can always be expressed as a relation whose domain is those very things as follows:

$$x = y \text{ iff} (\forall a)(\forall b)(x = S(a) \,\&\, y = S(b) \rightarrow Eq(a, b)).$$

We may accordingly take it that the criterion of identity for things of a certain sort may always be expressed as a specific equivalence relation on things of that very sort.

**(p.391)** However that skirmish now makes for a natural proposal. Consider, for any sortal concepts F and G, the respective open sentences, x $eq_F$ y and x $eq_G$ y, which result from statements of the relevant specific equivalence relations[76] by replacing both related terms with variables. Then we may propose that the concepts F and G share a criterion of identify just when

(*)

$$(\forall x, y)(x eq_F y \leftrightarrow x eq_G y$$

is a (conceptually) necessary truth. It is in this sense that tiger shares its criterion of identity with animal, and philosopher shares its criterion of identity with person. This is the usual case, and it is immediately clear how, if categories are to be maximally inclusive sortal concepts under a single criterion of identity in this sense, then there is no question of number and person being sub-sortals of a single category. For that would require a C such that both (∀ x, y) (x $eq_{Num.}$ y ↔ x $eq_C$ y and (∀ x, y) (x $eq_{Pers.}$ y ↔ x $eq_C$ y) each held good of conceptual necessity, and hence that (∀ x, y) (x $eq_{Num.}$ y ↔ x$eq_{Pers.}$ y) did so. And there is—it seems obvious enough—no such conceptual necessity. Categories, in the case, are such that all their sub-sortals actually share a criterion of identity. So no objects falling under sortals with different criteria of identity can instantiate a common category.

However, the usual case is perhaps not the only possible kind of case. Class (extension) is plausibly a categorial sortal. Consider (again) the vexed issue whether numbers are classes. Frege felt free to stipulate that they are, but not because number shares its criterion of identity with class in the sense just sketched. For it doesn't. What is true is that there is a range of classes the identifications and distinctions among which are isomorphic to those among numbers, being constituted in the very same considerations: Nx:Fx and Nx:Gx will be the same, or distinct, according as F is equinumerous to G; and so will the corresponding equivalence classes, $\{H\colon H\ 1{-}1\ F\}$ and $\{H\colon H\ 1{-}1\ G\}$. The relationship between the associated criteria of identity is thus not that indicated in (*) but rather—as a first approximation[77]—that it holds of conceptual necessity that to any numbers x and y correspond classes w and z such that x and y bear the relation, . . . $eq_{Num.}$ . . . if and only if w and z bear the relation, . . . $eq_{Class.}$ . . . More generally, the relationship is that between concepts F and G for which it holds of conceptual necessity that

(**)

$$(\forall x, y)(Fx \,\&\, Fy \rightarrow (\exists w, z)(Gw \,\&\, Gz \,\&\, x eq_F y \leftrightarrow w\, eq_G z))$$

(**) is necessary but not sufficient for (*). Call the relation it expresses the *subsumption* of ‘ . . . $eq_F$ . . . ’ within ‘ . . . $eq_G$ . . . ’ The line of thought that would regard that relation as sufficient for the sortal inclusion of F within G would **(p.392)** presumably be that if every identification and distinction among the Fs is mirrored by one among the Gs, dictated by the very same considerations, then how can they be distinct existences—what could ground their distinction?[78]

That thought is resistible, for a reason to be noted in a minute. But if it is sustained, then we must recognize that it is no longer necessary in order for a pair of concepts to be included in a single category that they share criteria of identity in the sense marked by (*). So we cannot directly conclude that numbers and persons do not belong to any common category just by that reflection. Rather there is the possibility that ‘ . . . $eq_{Num.}$ . . . ’ and ‘ . . . $eq_{Pers.}$ . . . ’ may alike be subsumed under a more inclusive equivalence relation, just as ‘ . . . and . . . being such that any lines of which they are the respective directions are parallel’ and ‘ . . . and . . . being such that any triangles of which they are the respective shapes are similar’ may be so subsumed under ‘ . . . and . . . being such that any properties of which they are the respective classes are coextensive’.

Might there be a category which stood to person and number as class thus stands to direction and shape? Well, yes, of course: *class* itself! For evidently whenever the identifications and distinctions among a sort of objects are grounded in the obtaining of a certain equivalence relation, they will be perfectly isomorphic to those among the corresponding equivalence classes. But then, first, that seems cause to be wary that subsumption of corresponding equivalence relations is generally good enough to ground categorial inclusion—unless we want everything to be a class; so the concept of category annexed to (**) is actually under some pressure. And second, the thought in any case does nothing to create a difficulty for the would-be solver of the Caesar Problem. For if numbers and persons are both taken to be classes, they will be distinct in any case, since of different types: numbers will be equivalence classes of concepts and persons will be equivalence classes of . . . well, persons. And—prescinding from consideration of the horrible impredicativity implicit in that proposal—no one had suggested that we have no secure grasp of the distinction between concepts and persons—or between concepts and any other spatio-temporal continuants.

In general, if categorical inclusion is determined by *subsumption* rather than *identity* of criteria of identity—and we have just seen reason to doubt it should be—it will be hard to *prove* that there is no single category including both person and number. But trouble for a would-be solver of the Caesar Problem can be expected from this direction only on two conditions: first, that it is possible to construct an equivalence relation, ' . . . $eq_C$ . . . ', subsuming both ' . . . $eq_{Num.}$ . . . ' and ' . . . $eq_{Pers.}$ . . . ' and determining a genuine category, C; and second that the relationship between the instances of C whose **(p.393)** identifications and distinctions mirror those of the numbers and those whose identifications and distinctions mirror those of persons then remains somehow imponderable —that it may not be directly settled by reference to ' . . . $eq_C$ . . . ', as happened when numbers and persons were taken to be classes. We cannot claim to have shown that these conditions cannot be met—we leave the issue as an exercise for a stubborn critic. But absent any reason to think a construction to the contrary is possible, it seems reasonable to take it that number and person are sub-sortals of no single category.

So much then by way of refinement of the notion of a criterion of identity and an indication of how it may subserve the belief that the number 0 and Julius Caesar thus belong to distinct categories. Does that suffice for their distinctness? Not yet: we need in addition that

> (U) No object can belong to more than one category.

Can that be shown?

Not conclusively, we think, but it will be instructive to explore the matter. Let C and D be distinct categories, $c^=$ and $d^=$ their respectively associated criteria of identity, and suppose—with a view to a *reductio*—that x falls under both. Call any identity statement one of whose related terms refers to x an *identification* of x. Then it might seem that since such an identification may now be answerable to assessment by independent criteria, nothing need stand in the way of a clash in such assessments: that x may be identified with y via $c^{=}$, for instance, but distinguished from y under $d^{=}$. Of course, the sortal relativity of identity has been defended by some able philosophers, including Geach and Noonan.[79] But it is a difficult pill to swallow: for if x and y are genuinely discernible under D, the merest common sense seems to allow no option but to regard them as distinct *tout court* so that their 'identity' under C is reduced to the obtaining of a mere equivalence relation, and C loses its status as a genuine sort of object.[80]

In the case that exercises us, however, matters do not take that shape. For the supposition that Julius Caesar is the number 0 carries no implication of the existence of any *single* identification which may be open to assessment by the independent respective criteria of personal and numerical identity. What the supposition entails, rather, is that some identifications involving reference to Caesar will be assessable by criteria of personal identity and others by the criterion of numerical identity—there is no implication that **(p.394)** any one identification will be open to assessment in both ways. It may be replied that, for all that has been shown to the contrary, it might still be the case that we wind up distinguishing emperors—Julius and Tiberius, say—who are identified under certain numerical guises—say '0' and 'the number of primes between 48 and 52'. But that response seems vulnerable to the simple reflection that since Julius and Tiberius are indeed distinct, there is no question of *any* singular term referring to both, so no question of any numerical term which refers to Caesar co-referring with a numerical term referring to Tiberius. There thus seems to be no way of enforcing the charge that a denial of (U) would jeopardize the absoluteness of identity.

The real objection is different. What is clear is that no place can be provided by the ontological framework outlined for *recognition* of dual-category status, even if such a thing could indeed occur. In that framework, all negotiable judgements of identity are essentially *internal* to a category, to be supported or confounded by appeal to whatever criterion of identity is associated with it. There is simply no provision for evidence for or against cross-categorical identities. So one who stubbornly presses the Caesar Problem in the context of that framework has to face up to certain costs. Clearly, it cannot be right to reproach the introduction of a class of expressions via an abstraction principle for failing to resolve a class of issues which there is no resolving in principle for *any* singular terms. When Frege moved past Hume's Principle to explicit definitions of the natural numbers in terms of extensions, it was because he supposed that he could thereby resolve an *objectionable and avoidable* indeterminacy—and a type of indeterminacy which is resolved in other cases. But once it is granted that Hume's Principle, conceived as fixing the content of statements of numerical identity, does at least afford us that number and person are of different categories, to persist with the ‘problem’ is to persist with a worry about a species of indeterminacy which there is no resolving *anywhere*. We have and can have no controls on cross-categorical identifications. The indeterminacy as we now have it would afflict relations among sortals which are explained by means quite other than Fregean abstraction. Might Frege be identical with a statue that existed in Caesar's time? All the obvious distinctions—that Frege only came into being with his birth, that the statue was lifeless, etc.—are merely question-begging when what is at issue is something acknowledgedly undetectable. If the reply that persons and statues belong to different categories in the sense we have been concerned with is not sufficient for a negative answer, then nothing is.

Stubborn pressers of the Caesar Problem certainly include our good friends, Sullivan and Potter. It's worth relating the foregoing to what seems to be their main line of objection, to which we have so far given no attention. They write:

> What did Locke realise about ‘gold’. Effectively that there is an element of blind pointing in our use of such a term, so that our aim outstrips our vision. Our **(p.395)** conception fixes what (if anything) we are pointing at but cannot settle its nature: that is a matter of what's out there. One image of the way [Hume's Principle] is to secure a reference for its terms shares a great deal with this picture.[81]

They go on to add in a note the qualification that it does not share everything; in particularly, the common idea that the guiding conception of the referent of a natural kind term may actually be false to its referent need have no counterpart here; what is relevant is only that

> the guiding conception of a thing will not contain the whole truth about it, not that it will not perhaps be wholly true.[82]

It is important to be clear about the limits of the particular objection. There is no suggestion that Hume's Principle fails to fix a concept, or fails to ensure that referents are assigned to the terms it enables us to introduce which instantiate that concept; we are not here dealing with Dummett-style worries about impredicative specification or Field-style concerns that to lay down Hume's Principle as an explanatory stipulation can only ensure something about the character of any referents of its terms, not that there are indeed such referents. It is being allowed—at least *pro tem*—that Hume's Principle is indeed a necessary truth, exactly as normally formulated, and hence that we know that it suffices for the existence of numbers that certain concepts enter into relations of one–one correspondence and that numbers are things which are essentially identified and distinguished from each other by reference to facts of 1-1 correspondence between concepts.[83] What is being challenged is that anything has been done to ensure that this is the *whole* truth about them. For if it may not be the whole truth, then we have no right to take it that numbers are *distinguished* from other familiar types of thing by their having this characteristic.

However, provided it is granted that Hume's Principle is indeed a necessary truth, it *is*—to stress—thereby being allowed that being so identified and distinguished is an essential characteristic of numbers. (If the proviso is not granted, that takes us back to the Field-style objection again.) So the things which are the numbers are essentially individuated by facts to do with 1-1 correspondence between concepts. Hence if this is not the whole truth about them in some relevant way—i.e. if it is to be consistent with their being persons, or rivers, etc.—they must also be essentially individuated by reference to considerations of some quite other kind. But that is just to say that they will belong to two *different* categories in the sense latterly introduced. If, as we have just argued, cross-categorical identifications are imponderable as a class, then the reply to Sullivan and Potter's principal point is this: while **(p.396)** one would expect that a satisfactory ontology will indeed legitimate the idea that there are *no* cross-categorical identities, it remains that even before we have won through to such an ontology, we can see that the ‘Caesar Problem’ is nothing special to abstractive explanation—that the concern Sullivan and Potter try to make vivid must afflict all explanations of sortal concepts; and that our grasp of mixed identity contexts featuring numerical terms, and the associated open sentences, is as secure as our grasp of mixed identity contexts everywhere. If there is a problem for the neo-Fregean construction of arithmetic hereby, it is a problem for all sortal discourse.

Let's take stock. In effect, the *Frege's Conception* proposal, properly developed, culminates in a dilemma:

> (i) If it is granted that there are no true cross-categorical identifications, then we have the following general sortal inclusion principle: for sortal concepts, F and G,
>
> > (SI#) Some Fs are G only if F and G are each sub-sortals of one and the same category.
>
> No Fs are Fregean abstracts, therefore, unless Fs belong to the category to which those abstracts belong. F's and G's inclusion within a common category requires at least the subsumability of their respective criteria of identity under a single such criterion, and perhaps the identity—in the sense of (*)—of those criteria. So no Fs are Fregean abstracts if the criterion of identity for Fs and the criterion for identity for those abstracts, as given by the appropriate abstraction principle, cannot be so subsumed. The recognition that Caesar is not a number may thus indeed be accomplished by the simple reflection that whatever determines personal identity—and however elusive a philosophical account of it may be—it is surely not to be captured by any generalized criterion of identity of which one–one correspondence between concepts is another instance. (It is true, as noted, that insofar as criteria of identity are given by equivalence relations, they may always be subsumed under the criterion of identity for (the appropriate equivalence) classes; but if we make that move, then—to repeat—Caesar and Zero will be distinguished qua classes in any case.)

(ii) If, on the other hand, we are up against a ‘stubborn presser’ and it is *not* granted that there are no true cross-categorical identifications, then the Caesar Problem becomes just an instance of a problem that afflicts *all* sortal concepts whose instances would belong to different categories, and there is no special cause to see the vulnerability of Hume's Principle to the problem as marking a deficiency in the understanding of numerical terms which it may be used to bestow. And that is good enough. The Caesar Problem was only ever a problem if the indeterminacy involved was special to the terms introduced by Fregean abstraction and did not affect singular terms in general. So if sustaining it demands such generalization, it has effectively gone away.

Notes:

(1) This and the next two quotations are from Frege (1884: § 56).

(2) Frege (1884: § 62).

(3) As usual, we use ‘NxFx’ to mean ‘the number of x's such that Fx’. Frege defines ‘the relation φ correlates the Fs with the Gs’ and ‘the relation φ is one–one’ in essentially this way in Frege (1884: § § 71–2).

(4) Having endorsed Leibniz's definition of identity [‘Eadem sunt, quorum unum potest substitui alteri salva veritiate’], Frege declares that since ‘in universal substitutability all the laws of identity are contained’ all that has to be shown is that provided lines *a* and *b* are parallel, ‘the direction of *b*’ may be substituted everywhere, *salva veritate*, for ‘the direction of *a*’ (and conversely, of course)—that is, that parallelism serves as an unrestricted congruence relation for open sentences into whose argument-places direction-terms may be inserted. The task, he observes, is made easier by the fact that the only sentential contexts involving direction-terms which we are taking ourselves initially to understand are identity contexts, so that other contexts may be dealt with by stipulating, in effect, that any newly introduced predicates of directions shall be defined so as to ensure that parallelism functions as a congruence for them. If we have reservations about Leibniz's definition—perhaps prompted by doubts about the Identity of Indiscernibles and problems about substitutability in non-extensional contexts—we might instead take identity as undefined, but governed by the basic laws that identity is reflexive and that co-designative terms are interchangeable *salva veritate* in all extensional contexts. Either way, the task reduces to showing that parallelism (in the case of directions) and 1-1 correlation (in the case of numbers) function as congruences for all properties expressible by means of open sentences free of (essential occurrences of) non-extensional vocabulary. See Wright (1983: 105–6).

(5) Frege (1884: § 66).

(6) The well-known difficulty is that, (Fregean) concepts being extensionally individuated, the extension with which the Fregean definition would identify NxFx would contain a member corresponding to every set having as many elements as there are objects falling under F, and so would be as big as the universe, and thus ‘too big’ (on pain of contradiction) to be a set. It could, of course, be a proper class in a theory which works with the set–proper class distinction—but that is obviously no help, since it would remain the case that numbers could not be elements (of either sets or proper classes).

(7) i.e., what is often now called a (Fregean) *abstraction principle*. Generally, an abstraction principle has the form:

$$\sigma(\alpha) = \sigma(\beta) \leftrightarrow \mathrm{Eq}(\alpha,\beta),$$

where Eq is an equivalence relation on entities of the type of α, β, and 'σ' is an operator which forms singular terms from expressions for entities of that type. In general, the referents (if any) of these terms may be entities of the same or a new type. In the case of Hume's Principle—and of any other second- or higher-order abstraction, which works with an equivalence relation on concepts of some level—the entities to which the left-hand-side terms purport reference must be of a new type, being objects, not concepts. In the Direction Equivalence—and all other first-order abstractions—the equivalence relation is on objects, so that we do not automatically get a difference in type in this way. But if the Direction Equivalence is construed, as it may be, as working with an equivalence relation (parallelism) on physical lines, we may wish to say that, because direction-terms denote, if anything, abstract objects, the abstraction principle serves to introduce talk of objects of a different kind—even if not of a new type in the foregoing sense.

(8) Frege confronts a special case of the Caesar Problem at *Grundgesetze* § 10, where he observes that his stipulation concerning courses-of-values—that the course-of-values of a function Φ is to be the same as that of a function Ψ just in case they have the same values for the same arguments—does not fully determine the reference of a course-of-values name, because it does not enable him to decide 'whether an object is a course-of-values that is not given as such'. Noting that at this point in the development of his system, the only objects he has recognized which are not given as courses-of-values are the two truth-values, he proposes to remove the indeterminacy by stipulating that each truth-value is to be identified with the extension of a concept under which it and it alone falls, i.e. in effect, with its own unit class. It should be clear that (as Frege himself in effect acknowledges) this affords no general solution to the Caesar Problem for extensions, since before we can safely stipulate that some object not given as a course-of-values is a certain extension, we need an assurance that it is not (behind our back, as it were) some other extension—else our new stipulation might conflict with the original stipulation of identity-conditions for courses-of-values. A solution to the Caesar Problem is thus presupposed, and cannot be provided, by generalizing the kind of stipulation Frege envisages for truth-values. For further discussion, see Wright (1983: 112–13) and, for an analysis in some depth of Frege's move, Moore and Rein (1986).

(9) For the dubbing, see Boolos (1990*a*: 268). See Wright (1983: 158–69) for a detailed sketch of the derivation of the

Peano axioms from Hume's principle (there labelled $N^{=}$); an outline is also given in the Appendix to Boolos (1990*a*).

(10) As noted independently by George Boolos, see his (1987*b*: 6–10), and John Burgess (1984); also by Harold Hodes (1984) —(see remarks at 138 concerning an equivalent of Hume's Principle involving branching quantifiers). For a detailed proof, see the first appendix to Boolos and Heck (1998).

(11) Frege (1884: § 68).

(12) This line of thought assumes, of course, that any facts of the matter about which objects the numbers are have to be determined by (a satisfactory explanation of) the *concept* of number. A more thoroughgoing platonism than Frege's might reject this assumption: perhaps the best explanation of the concept of number simply stops short of determining which objects are numbers, rather as the best explanation of the concept of water stops short of determining—on the standard conception of natural kind concepts—what sort of stuff water actually is. The difference would be that, in the case of numbers, no pursuable empirical issue arises either.

(13) He would not have had to do either had his platonism been of the extreme variety bruited in the preceding note. Such a position: that while the truth of contexts involving numerical terms suffices to ensure their objectual reference, we may just not know, nor be capable of knowing, facts about which objects we thereby secure reference to, might find allies among friends of the so-called epistemic conception of vagueness (most ably defended by Timothy Williamson; see his (1994)). But whatever one's view of that, it needs to be acknowledged that, as the water example shows, it is at least not always reasonable to expect questions about what (fundamental) kind of thing we succeed in talking about by means of certain linguistic devices to be answerable a priori just on the basis of reflection on canonical ways of explaining those ways of talking. We shall have cause to return to this point in the sequel.

(14) See Evans (1982: 100–5).

(15) Of course, qualifications on the principle thus baldly stated will be required by the fact that we are not able to understand expressions of arbitrarily great length (for permutation of understood constituents may make an understood sentence longer), and by possibilities of various forms of cognitive dissonance—somebody might have a disabling phobia of sentences in italic type, for instance. But these considerations do not affect the basic point, nor its application in the present setting.

(16) Or, more accurately, play its part in fitting such sentences with a determinate sense—obviously it will not do this by itself, but only in combination with the contributions of the other constituent expressions.

(17) In a footnote, Evans remarks that a proviso ought to be added to the effect that the predicates in question should be 'categorially appropriate' to the subjects—see Evans (1982: 101, fn.17).

(18) There are other reasons, besides those emphasized in the text, why a solution to the Caesar Problem is needed. An especially important one, in our view, is that the availability of a positive solution is integral to defusing what is arguably the clearest, and potentially most damaging, form of the charge—pressed on several occasions by Michael Dummett—that the impredicativity of Hume's Principle debars it from discharging the explanatory task we assign to it. See Wright (1983: 139–45, 180–84); Sect. 6 of Hale, Essay 8; and the concluding Sect. of Wright, Essay 10.

(19) See Heck (1997*b*). Heck does not claim that there are no serious problems confronting the version of logicism he canvasses. He expresses no view, at least in this paper, about whether the problems he thinks his version of logicism, along with more familiar versions, must face can be solved.

(20) Heck contends that essentially the same, or at least a closely related, difficulty—the apparent impossibility of proving that there are infinitely many objects—is what really underlies Frege's first rehearsal in (1884: § 56) of the Caesar Problem, in his criticism of the inductive definitions proposed in § 55 of that work. No doubt that is a (potentially lethal) objection to the inductive definitions, considered as a basis for arithmetic. But as an interpretation of what Frege was driving

at, specifically in his remark that the definitions do not enable him to prove that if the number *a* belongs to a concept F and the number *b* belongs to the same concept, then $a = b$, it is—in our view—forced and quite implausible. Contrast the much simpler and straightforward explanation offered in Sect. 1 above.

(21) Heck (1997*b*: 278).

(22) Ibid. 305–8.

(23) Ibid. 279.

(24) On this explanation, logical Sorts comprise expressions, rather than the entities for which those expressions stand. But it is entirely natural to speak, as Heck does without explicitly advertising the fact, of (non-linguistic) entities as being of the same Sort (when and) only when corresponding expressions are. We shall follow suit.

(25) Heck (1997*b*: 287).

(26) In case it is not immediately obvious why it does have to be held that numbers ascribed to concepts true or false of numbers have to be of the same Sort as concepts true or false of basic objects, recall that 0 will be defined as $N\text{x:x} \neq \text{x}$, while the proof of the infinity of the natural numbers will make essential reference to numbers of the type $N\text{x:x} \leq n$. The latter had better be of the same Sort as 0, for all that $\xi \neq \xi$ is a concept true or false of, *inter alia*, basic objects, while the latter involve concepts true or false of numbers.

(27) Cf. Heck (1997*b*: 302). Heck supports this contention by appeal to various examples he marshals against the suggestion that Hume's Principle might be formulated in a type-theoretic framework.

(28) Indeed, as Heck observes, [once she has got this far] she need not even take a stand on whether numbers are *objects* of any Sort at all, as opposed to, say, second-level concepts. All that matters, for the formal development of arithmetic, is that all numbers should be of the same Sort—it does not matter what Sort that is.

(29) For some sympathetic thoughts, see O'Leary-Hawthorne and McDonough (1998).

(30) In fact, it is anything but clear that Heck would have found a way of bypassing the Caesar Problem altogether, even if we prescind from the dilemma just explained. For unless he simply assumes—but with what right?—that numbers exhaust the Sort to which they belong, there will still be some mixed identities—perhaps ones linking terms for numbers with terms for sets—which must be assigned sense but on which ($HP_{nn}$) gives no purchase.

(31) This section is indebted to Gideon Rosen (1993), and to many discussions with the author. For those readers familiar with Rosen's essay, the line eventually developed here is best viewed as an alternative response to the ‘problematic’ that Rosen has developed by the start of Sect. IX of that paper, at which point he initiates the path towards his own canvassed ‘Refutation of Nominalism(?)’, which makes essential play with higher-order abstraction.

(32) Dummett (1991*a*: 126).

(33) The restriction to first-order abstractions, that is: abstractions on relations on objects, is needed if, following Frege, the terms of the abstractive relations in higher-order cases are conceived as unsuited to be referred to by any singular term; in that case, there is no question of taking the new singular terms as referring back within the domain of the abstractive relation, since the entities comprising that domain are simply not the kind of thing to which singular reference is possible. So Dummett's claim would not engage Hume's Principle itself, on Frege's view.

(34) An equivalence relation *R* is a congruence for a predicate *F* just in case for any objects *x* and *y*, *x*'s bearing *R* to *y* ensures that either both *x* and *y* are *F* or neither is.

(35) Not unnecessarily: we cannot yet presume that the left-hand sides of the stipulations under (ii) have a significant semantic structure.

(36) This train of thought belongs with one—what we regard as the most fruitful—interpretation of Frege's Context Principle. For further elaboration see Wright (1983: § § iii, v, viii–x, xii, xviii); Hale (1987: chs. 1, 2, 7); and the Introduction to the present volume.

(37) Such scepticism may, of course be more or less extreme, ranging from doubt about the justifiability of ‘postulating’ abstract objects (Field) to the conviction that all purported thought of, or reference to, them is simply unintelligible (sometime Quine and Goodman).

(38) For argument that the trick cannot be pulled by any means involving radical syntactic regimentation, see Heck (2000). Heck makes a convincing case the construal of (seemingly perfectly legitimate) contexts generated by applying quantifiers such as ‘most’, ‘many’, and ‘few’ to the purported ontology of the left-hand sides of Fregean abstractions inevitably calls for a representation function defined on the ontology of the right-hand sides which then regenerates a range of singular terms functioning exactly as the original abstract terms which the reductionist had set herself to eliminate.

Some of the issues raised by the second and third nominalist tendencies recur in the context of Michael Dummett's attempt at an intermediate position—intermediate, that is, between the ‘mere notational variant’ view and that of the Fregean. See Dummett (1973: 495–505) and (1991*a*: 191–9); Wright (1983: § x), and Essay 9; Hale (1987: ch. 7), and Essay 8.

(39) This volume, Essay 6.

(40) And receives it in Essay 5.

(41) Dummett himself, of course, is not such a nominalist.

(42) Cf. Rosen (1993: Sect. VI, 168–9).

(43) That is, cases whose truth-values will depend on details concerning the objects and worlds involved and are not simply settled by ≈'s being an equivalence relation. The comparative simplicity of the ‘first stab’ for numbers above reflects the fact that there are no merely hypothetical numbers—no need for an analogue of the ‘$h_i$’ terms in representing the corresponding range of cases for numbers. We will not spare the reader the pleasure of confirming for herself that the number of non-trivial cases for numbers is six (and of course only one if, as for Frege, concepts are individuated by their actual instances, though that is, naturally, not the way we understand the expression, ‘the number of Fs’, as is attested

to by the ready intelligibility of thoughts of the form, ‘the number of Fs might have been different’).

(44) A proposition under case (iii)(a) above.

(45) In effect the thought is:

$$(\forall a_i \neq a) \neg (a_i A \approx aH_n) \rightarrow \neg (Da_i A = DaH_n)$$

which will be accessible from the fully modalized Direction Equivalence by instantiation to *a* and contraposition on the relevant case of (iii)(c).

(46) Again, the relevant application of the Direction Equivalence is to a modal thought of type (iii)(a).

(47) It is a very nice question whether this point can be enforced on any nominalistically acceptable interpretation of the higher-order quantifiers. The heart of Gideon Rosen's argument in his (1993) is a cautiously affirmative answer—see Sect. IX: 174–80.

(48) A superfluous qualification, of course, if concepts are extensional.

(49) e.g. consider

1. Φ (*F*) & Φ (G) ∨ F one–one corresponds to G,

where Φ (F) holds just if F is empty or applies to everything.

(50) To be sure, once any empty concept is excluded from consideration, one—one correspondence will indeed be neither inflationary nor deflationary on any domain. But it will also, of course, associate all unit concepts with the same abstract whereas any higher-order abstraction which is to avoid the Simple Argument must associate each ‘*F*(F)’ with an object falling under F, as we have seen.

(51) Wright (1983: § § vii, xiii–xv, 41–50, 104–29). The term originates with Locke (1690: Bk I.ii. 15), but its modern usage owes much to Strawson (1959: 168 ff.), and Wiggins (1967).

(52) See above, Sect. 2

(53) Assuming that no numbers—or people—are in principle unnameable.

(54) Wright (1983: 116–17).

(55) Another of our earlier discussions of the idea that a criterion of application for a sortal concept might be derived from its associated criterion of identity is Hale (1987: ch. 8, Sect. 2). The key principle there proposed is:

> (S) Singular terms from a given range stand for instances of a sortal concept *F* iff there is some sortal *G*, whose extension is included in that of *F*, such that where *a* and *b* are any terms from that range, understanding ‘a = b’ involves exercising a grasp of the criterion of identity for *G*s. (1987: 206)

(S), not being specific to *number*, but applying rather to any sortal principle *F*, invites comparison with the more general principle formulated in Wright (1983: 122), and reproduced here as (**SI**) a little later in the text. Just as $N^d$ results from (SI) coupled with the criterion, $N^=$, for identity on numbers, so (S) coupled with $N^=$ yields condition for a sortal *G* to overlap the concept of *number*—but a stronger condition. In effect, the difference is that while $N^d$ requires only that suitably chosen identities ‘a = b’ linking *G*-denoting terms *can* have their truth-conditions explained in terms of one—one correlation among certain concepts, the corresponding instance of (S) has it that the truth-conditions of ‘a = b’ *must* be so explained. The interest of the stronger principle resides principally in the fact that, if sustainable, it would arguably—and in contrast, apparently, with $N^d$—exclude the finite von Neumann ordinals (∅, {∅}, {∅, {∅}}, . . . ) from being numbers, and would—or so Hale (1987) argued—facilitate a head-on response to Benacerraf's (1965) challenge to the Fregean view that numbers are self-subsistent objects. Wright (1983: § xv) advocates a different response to Benacerraf's challenge, arguing that even if, as Benacerraf in effect contends, it is radically underdetermined which sets numbers are, that is no argument against regarding numbers as objects until it is shown that the indeterminacy is peculiar to number, rather than merely one instance of the general indeterminacy argued by Quine to afflict all sortal concepts. These differences are irrelevant to the discussion to follow here.

(56) Wright (1983: 122).

(57) Thus the criticism of the Wright (1983) proposal advanced by Michael Dummett (1991*a*: 160–1) rests upon a misunderstanding. Dummett glosses the reasoning by which he thinks the conclusion that Caesar is no number is intended to be reached as follows:

> the stipulated equivalence [e.g. Hume's Principle, or the Direction Equivalence] lays down the criterion of identity —for directions or for numbers. . . Any identity-statement concerning a direction or a number must therefore be determined as true or false according to that criterion. This can happen only if the statement asserts the identity of a direction with a direction, or of a number with a number: if different criteria of identity are associated with the terms on either side of the identity, there is no way in which either criterion can be applied, and hence such an identity-statement is ruled out as false without further ado. The criterion for the identity of human beings is quite different from that for the identity of numbers; and hence the stipulation specifying the latter criterion of itself determines that the statement ‘The number of planets is Julius Caesar’ is false.

It should be clear from the paragraph to which this note is appended that this is simply not how the proposed solution was intended to work. For a fuller discussion of this point, see Hale Essay 8, Sect. 3.

(58) Dummett (1991*a*: 161–2).

(59) Stirton (2000); Stirton presents the example as an embarrassment to the claims Hale makes for the notion of *compact entailment*—see below—in Hale, Essay 4.

(60) This is required to preclude arbitrary pairs of necessary truths (or necessary falsehoods) from qualifying as having the same truth-condition.

(61) i.e. two sentences share their truth-condition iff anyone who understands both is able to tell, using at most reasoning sanctioned by compact entailments, and without determining their truth-values separately, that they must coincide in truth-value.

(62) Compact entailment was first introduced, for quite different purposes, in Wright (1989). For fuller discussion of its use in this connection, see Hale (1997) reprinted here as Essay 4, along with a new postscript which includes some important clarification and amendments.

(63) It might be supposed that Dummett could just as well have constructed his example around Hume's Principle and Frege's proposed stipulative identification of numbers with

extensions (classes), arguing that the *Frege's Conception* proposal has the unacceptable consequence that the legitimacy of what ought to be a perfectly acceptable stipulation identifying cardinal numbers with extensions is made to depend upon whether it follows or precedes a stipulation of Hume's Principle. In fact, he could not have done so. The parallel breaks at a crucial point. It is true enough that, just as a stipulative identification the eccentricity of an ellipse with the ratio of the distance between its foci to the length of its major axis entails acceptance of the geometrical criterion for identity of eccentricities, so a stipulative identification of the cardinal number belonging to the concept *F* with the extension of the concept: *equal to the concept F* entails acceptance of the criterion given by Hume's Principle for identity of cardinal numbers—in neither case is there room for a *further stipulation*. But it would be a mistake to suppose that, if the concept of *cardinal number* is first fully explained by stipulating Hume's Principle, an argument parallel to that sketched in the text to show eccentricities cannot be real numbers will show that cardinal numbers cannot be extensions. According to that argument, if both eccentricities and sharpnesses are introduced, via the obvious abstraction principles, as kinds of shape, they must be disjoint kinds of shape (since no ellipse is similar to any isosceles triangle), but if eccentricities are identified with real numbers, we are obliged to cross-identify shapes of these disjoint kinds. An immediate problem in constructing a parallel argument is that it is quite unclear what sortal, if any, might play the role discharged by *sharpness* in the original. Whereas laying down the Eccentricity Equivalence introduces eccentricities as a proper sub-sort of the generic sort *shape*, laying down Hume's Principle precisely does not introduce numbers as a sub-sort of any more general sort. This difficulty could be avoided by running the argument in terms of disjoint sub-sorts of number itself—for example, we might take even number and its complement (comprising all finite odd numbers together with all transfinite cardinals) as counterparts of eccentricity and sharpness—provided we could show that identification of numbers of these two sorts with extensions would result in their unwanted cross-identification with each other. However, there is a more fundamental obstacle. In the case of the original argument, the unwanted conclusion that (some) eccentricities are sharpnesses does not straightforwardly follow from the assumption that both eccentricities and sharpnesses are real numbers—so far, the numbers the eccentricities are could be disjoint from the numbers the sharpnesses are. This causes no problem for the present argument, but only because a further assumption is satisfied. We are taking it that if eccentricities are identical with real numbers at all, they are identical with certain ratios—*E*(e) is identical, if with any real number, then with the real that is the ratio of the distance between e's foci to the length of e's major axis; and likewise that, if sharpnesses are identical with reals, they too are identical with certain ratios—S(t) is identical, if with any real, then with the real that is the ratio of t's height to t's base length. Now, under these assumptions, it is not merely the case that every eccentricity is a positive real number, but also that every positive real number is an eccentricity; and likewise, it is not merely the case that every sharpness is a positive real number, but also that every positive real number is a sharpness. But then, of course, it follows that each and every eccentricity is a sharpness, and conversely. There is, however, no reason why identifying the even numbers and their complement with classes must result in their (erroneous) cross-identification with each other. If, for example, they are identified with equivalence classes of concepts under the relation of equinumerosity, then they will be identified with disjoint classes of classes.

(64) Sullivan and Potter (1997: 139 ff.).

(65) Strictly speaking, this last claim goes beyond what is warranted by the considerations immediately preceding it—since, as noted, it is consistent with those considerations that σ(α) should be one of α and β. But the claim is true, as the earlier argument of Sect. 4 shows.

(66) We are indebted here to discussions with Gideon Rosen. The objection as formulated below is prefigured in his (1993).

(67) This need not — and doesn't — mean questions about sortal overlap never admit of resolution a priori. In particular it need not preclude such a resolution of the question whether numbers are ever people. For without settling any question of detail concerning the truth-conditions of e.g. statements of personal identity, it can be clear, and knowable a priori (a) that they will involve considerations of spatio-temporal (or at least temporal) continuity, and (b) that such considerations can have

no bearing upon the satisfaction of the criteria for identity of numbers.

(68) It wouldn't even help to resort to terms which explicitly bring their referents under both relevant concepts—'The mammal which is a horse and which is grazing in that field', e.g. For even: 'The mammal which is a horse and which is grazing in that field = the mammal which is a horse and for which Jack paid £2,000 at Appleby Fair last September' is likewise a priori equivalent neither to (a) nor to (b).

(69) See Strawson (1959: esp. 169 ff.) and (1974: chs. 4, 5); Wiggins (1967: esp. Part 2 and Appendix) and (1980: esp. chs. 2, 3); Dummett (1973: 173 ff.); Lowe (1989, 1997); and Noonan (1976, 1978, and 1980).

(70) For useful discussion, see Rumfitt (forthcoming), also Lowe (1989, 1997: 615).

(71) Frege (1884: § 46).

(72) Wiggins (1967: 7, 29).

(73) Since the filler of the gap marked by 'such-and-such a further condition' in 'There are *n* . . .'s (meeting such-and-such a further condition)' had better not itself be sortal or non-sortals will pass the test.

(74) See Dummett (1973: 76). However, this characterization might be liberalized to allow that the criterion of identity associated with a category may involve some element of generalization of the criteria of identity associated with sub-sortals of it—the way that coextensiveness among the appropriate equivalence classes can be seen as generalizing the equivalence relation used in a particular Fregean abstraction. This would allow *class* to be a category subsuming Fregean abstracts, like numbers and directions, as a species. More of this below.

(75) Notwithstanding the point that the nature of that criterion may in particular cases be an a posteriori matter.

(76) Whose identification, we emphasize once again, may be an a posteriori matter.

(77) A first approximation because of the complications associated with cases where intuitively unrelated identifications each hold—or fail to hold—as a matter of conceptual necessity.

(78) This line of thought would propose as a sufficient condition for the cross-identification of Fs and Gs what was observed, in the context of the earlier Isomorphism constraint, to be a necessary condition.

(79) See Geach (1962, 1967), Noonan (1980). Noonan (1997) provides a useful discussion and further references.

(80) For example, the present British Prime Minister shares his surname with the author of *Animal Farm* and *Nineteen Eighty Four*—and is hence the same *surman* in Peter Geach's sense. But Tony Blair and Eric Blair are distinct people; the common-sense view would be that their 'identity' under the concept surman consists merely in the obtaining of the equivalence relation . . . has the same surname as . . . , and that surmen are no genuine sort of object. For a useful recent survey of the debate over the alleged relativity of identity, and further references, see Noonan (1997).

(81) Sullivan and Potter (1997: 145–6).

(82) Ibid. 146, n. 6.

(83) So the thought is not that numbers may have an additional nature, undisclosed by Hume's Principle, in the same way that MPs have an additional nature undisclosed by the principle *Members of Parliament. That* concern—in effect, that for all we have said, Number might be not a sortal but a *functional* concept—is effectively precluded by the admission that Hume's Principle is a necessary truth.