LOGIC, SEMANTICS, ONTOLOGY

by

Richard Gustave Heck, Jnr.

B.S., Mathematics, Duke University
(1985)
B.Phil., Philosophy, Oxford University
(1987)

SUBMITTED TO THE DEPARTMENT OF
LINGUISTICS AND PHILOSOPHY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June 1991

Signature of Author ____________
Department of Linguistics and Philosophy
May 1991

Certified by ____________
George Boolos
Professor, Linguistics and Philosophy
Thesis Supervisor

Accepted by ____________
Wayne O'Neil, Chairman
Department of Linguistics and Philosophy

Table of Contents

LOGIC, SEMANTICS, ONTOLOGY

by

RICHARD GUSTAVE HECK, JNR.

Submitted to the Department of Linguistics and Philosophy on 9 May 1991 in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Philosophy

ABSTRACT

Logic, Semantics, Ontology consists of three papers concerned with ontological issues. The first, "That There Might Be Vague Objects", is a critical study of Gareth Evans's essay, "Can There Be Vague Objects". The author argues that the formal argument presented in Evans's paper is valid and that a contradiction can indeed be derived from the statement that it is indeterminate whether **a** is **b**. However, the deduction theorem fails in the required logic: Hence, one can not derive the validity of the statement that it is determinate whether **a** is **b**.

One who holds the view that there are vague objects is committed to the legitimacy of those principles to which appeal is required in the proof. Hence, the view that there are vague objects is committed to the claim that no statement of the form "It is indeterminate whether **a** is **b**" can be true, but also to denying the validity of its negation. Possible motivations for such a position are sketched and its tenability is defended.

The second paper, "Whether Structure May Be Misleading", is a critical study of Crispin Wright's *Frege's Conception of Numbers as Objects*, in which Wright defends Platonism, the view that there are Abstract Objects of various sorts. The author argues that Wright's view is too promiscuous, that Wright's view appears to commit us to the existence of far too many sorts of objects. The causes of this ontological extravagance are isolated, and the author suggests ways to avoid it. In the process, however, the author also argues that certain classical Reductionist arguments fail and that their failure is closely connected with the strongest motivations for Platonism.

The final paper, "Trans-sortal Identification", continues this discussion. It contains arguments against Reductionism and a suggestion of a form of Platonism new to the literature. The author argues that names of abstract objects are not eliminable and that, therefore, the fundamental question in this area is not whether we have in our language expressions which are truly *names* and which purport to refer to Abstract Objects; rather, the question is to what such names refer. The connection between this problem and Frege's Julius Caesar problem is duly noted.

The author argues for a particular view about what determines the kind of object to which names in a given class refer. From this principle it follows that names of abstract objects *may* refer to objects of different sorts than do names of concrete objects and, indeed, than do names of abstract objects of other sorts. Integral to this claim is the claim that there is a closely related principle which states conditions necessary if names of abstract objects of a given sort are to refer at all: That is, if abstract objects of a given sort are to exist. It is this claim, the claim that there is an important, non-philosophical question whether there are abstract objects of a given kind, which distinguishes the view from those previously discussed.

BIOGRAPHY

Richard Gustave Heck, Jnr.

Richard G. Heck attended R.J. Reynolds Senior High School, in Winston-Salem, North Carolina, from August 1980 until December 1981. In the Spring of 1981, he placed first in the North Carolina State Mathematics Contest, winning a scholarship to Duke University, at which he enrolled in January 1982. He studied Mathematics and Philosophy, concentrating, in Mathematics, on Algebra and Geometry. In Philosophy, he wrote a thesis, under Professor David Sanford, entitled *On the Interpretation of Wittgenstein's Later Philosophy*. He received the degree of Bachelor of Science, with honors, in May 1985 and graduated *summa cum laude*.

Mr. Heck was awarded a Marshall Scholarship, which enabled him to matriculate at New College, Oxford University, in October 1985. He studied under Michael Dummett, Wykeham Professor of Logic, and wrote a thesis, under Professor Dummett's direction, entitled *Rule-following and the Justification of Deduction*. He received the graduate degree of Bachelor of Philosophy in July 1987.

Fearing for the future of Philosophy in Britain under then Prime Minister Margaret Thatcher, Mr. Heck enrolled the following September in the graduate program in Philosophy at the Massachusetts Institute of Technology.

Mr. Heck's essay "On the Consistency of Second-order Contextual Definitions" is forthcoming in *Nous*, and he is currently editing a volume of essays in honor of Michael Dummett, to be published in 1993 by Oxford University Press.

Mr. Heck will be teaching at Harvard University beginning in July 1991.

ACKNOWLEDGEMENTS

I should like, first of all, to thank George Boolos, whose unending support and encouragement have been of inestimable value to me these last four years. He took me under his wing shortly after I arrived at MIT, a lost and lonely soul, and George has been not only an excellent advisor but a good friend.

I should also like to thank my two other advisors, Bob Stalnaker and Jim Higginbotham. Their criticism and support has made the papers contained within far better than they would otherwise have been.

Tom Kuhn has also been a valued friend and colleague. As different as our approaches to and areas of interest in Philosophy are, we found that there were more than enough common points of interest. I shall recall our discussions with awe and warmth.

Finally, I must thank my best friend and partner, Alice Graham-Brown. Throughout all the ups and downs, discoveries and disasters, she has been by me. She has learned how to remind me of past ups and downs, of papers which disintegrated and were revived in far better shape, without making me feel foolish when in one of the troughs: One lovely gift among many she has and shares.

for my parents,
Richard Gustave Heck, Sr. and Helen Mary O'Connor Heck

"The older I get, the smarter my father gets"
--my father

Logic,
Semantics,
Ontology

Richard Gustave Heck, Jnr.

Paper I

THAT THERE MIGHT BE VAGUE OBJECTS (SO FAR AS CONCERNS LOGIC)

0. Opening

Some years ago, Gareth Evans presented an argument which, he claimed, shows that there can be no vague objects.[1] Evans's paper has been the subject of much discussion. Little agreement, however, has been reached even on the nature of Evans's argument: There is little agreement regarding what is in dispute (what a 'vague object' is), what sorts of arguments are relevant, how Evans's argument addresses the problem, or what objections to Evans's argument, in particular, would be relevant.

I shall attempt here to resolve some of these difficulties. First, we shall look at what principles are required if Evans's formal argument is to succeed; we shall then consider objections to them. The most important of these concerns the formulation of Leibniz's Law or the principle of the Indiscernibility of Identicals. I shall argue that, though the standard version of this principle begs the question against one who maintains that there are vague objects, there is a version of the principle which does not.

---

[1] Gareth Evans, "Can There be Vague Objects?", *Analysis* XXXVIII (1978), p. 208.

I shall thus be arguing that Evans's formal argument is valid, but I shall reject his claim that that argument shows that there can be no vague objects. This, the ultimate conclusion of Evans's argument, depends upon a quite specific interpretation of the claim that there are vague objects. I shall argue that there is a weaker, independently plausible interpretation of the view, against which Evans has no argument. I shall develop this alternative view only to a *very limited extent*, only so far as is required to support the claim that it is, first, rightly described as committed to the existence of vague objects and, secondly, not so implausible a view as to be utterly uninteresting.

I shall not even attempt to decide whether, indeed, there are vague objects. My claim is only that logic alone does not show that there are not.

1. Evans's Formal Argument

Before beginning that discussion, however, it is worth reminding or informing the reader of the formal component of Evans's argument. Where '∇' is an operator to be read "It is indeterminate whether..." and 'λx' is a predicate-abstraction operator, the argument is, in short:

$\nabla(a=b)$
$\lambda x[\nabla(a=x)](b)$
$\neg\nabla(a=a)$
$\neg\lambda x[\nabla(a=x)](a)$
$\neg(a=b)$

Or, informally: Suppose that it is indeterminate whether **b** is **a**. Then **b** has the property that it is indeterminate whether it is **a**. But **a** itself does not have this property: For it is perfectly determinate whether **a** is **a**. Hence, there is a property, namely, 'being indeterminately **a**', which **b** has but which **a** does not have. Therefore, **b** can not be **a**.

Evans remarks that this conclusion contradicts the assumption with which we began, that it is indeterminate whether **a** is b. It is not immediately clear why this is so. We shall return to this question.

Evans's argument plainly relies upon a number of different principles. First, it relies upon the principle of the Indiscernibility of Identicals. For the moment, we may assume that Evans would maintain the validity of the schema:

(LL) $a=b \ \& \ \lambda x(Fx)(a) \rightarrow \lambda x(Fx)(b)$

This schema is, of course, equivalent to the following one:

(II) $\lambda x(Fx)(a) \ \& \ \neg\lambda x(Fx)(b) \rightarrow \neg(a=b)$

Such a principle justifies the transition from "$\neg(\lambda x)[\nabla(a=x)](a)$" and "$(\lambda x)[\nabla(a=x)](b)$" to "$\neg(a=b)$".[2]

Secondly, the application of this principle rests upon the claim that the predicate "$\nabla(a=\xi)$" expresses a 'property' of objects; that is, Evans is relying upon the claim that the operator '∇' does not induce an opaque context, so that the step of predicate-abstraction--from "$\nabla(a=b)$" to

---

[2] One might well wonder if Evans needs to appeal to such a strong principle. We shall not be ready to consider such a question until later: See section 5.

"$\lambda x[\nabla(a=x)](b)$"--is not, in general, invalidated by the presence of the operator '∇'. Thirdly, the argument depends upon the assumption that the *specific* inferences, from "$\nabla(a=b)$" to "$\lambda x[\nabla(a=x)](b)$" and from "$\neg\nabla(a=a)$" to "$\neg\lambda x[\nabla(a=x)](a)$", are valid. Fourthly, Evans relies upon the claim that reflexive identities are not of indeterminate truth-value. Where 'Δ' is an operator to be read "It is determinate whether...", we may record the principle as:

(R) $\Delta(a=a)$

Fifthly, if it is determinate whether A, it is not indeterminate whether A:[3]

(C′) $\Delta A \rightarrow \neg\nabla A$

From (R) and (C′), we pass to Evans's third premise.[4]

Evans would seem also to accept the claim that a sentence is determinate if, and only if, it is not indeterminate. We record this as the schema:

(C) $\Delta A \leftrightarrow \neg\nabla A$

These are the only assumptions appeal to which is required for the formal argument in Evans's paper.

---

[3] Throughout, 'A' is a syntactic variable for an arbitrary (open or closed) formula.

[4] Evans does not so derive it in his paper, but simply asserts that "$\neg\nabla(a=a)$" is true. The question is, however, what justifies the claim, and it would seem that only (C′), or some stronger principle, together with (R) can do so--unless, of course, we simply assume it as an axiom. I intend to concentrate attention upon 'Δ', rather than '∇', however, so I record (R) and (C′) as axiom-schemata.

I do not intend to question any of these assumptions in its own right, with the exception of the formulation of the principle of Indiscernibility. As I shall shortly argue, Evans's appeal to some version of this principle is justifiable. However, we shall discuss, later, whether his appeal to the principle, in this form, is legitimate, or whether only some weaker formulation of the principle can be justified.

Evans also remarks that '∇' and 'Δ' are "duals". We should thus probably ascribe the following principle to him:

(D) $\Delta A \leftrightarrow \neg\nabla\neg A$

In any event, the following is surely valid:

(Eq) $\Delta A \leftrightarrow \Delta(\neg A)$

For, if it is determinate whether A, surely it is also determinate whether not-A, and it does not matter which two of (C), (D), and (Eq) one takes as valid, since, as is easily shown, each of the three is derivable from the other two.

Now, a great deal of confusion has been caused by a slip which Evans made in his paper.[8] The slip is the result of an equivocation between the operator 'Δ' and a related but distinct operator '$\Box$', which is to be read "It is defin-

[8] A reference to Lewis's report of Evans's retraction of this slip can be found in Francis Jeffry Pelletier, "Another Argument Against Vague Objects", *Journal of Philosophy* LXXXVI, 9 (1989), pp. 481-92. See the footnote on p. 482. Lewis has not, to the best of my knowledge, published such a report himself.

itely true that..." (or simply, "Definitely:..."). The principle

(T$\Box$) $\Box A \rightarrow A$

is a natural one; we may take

(D$\Box$) $\Diamond A \leftrightarrow \neg\Box\neg A$

as the definition of a dual operator. Principles analogous to (C) and (Eq), however, are plainly invalid: If it is definitely true that A, not only does it not follow that it is definitely true that not-A, it follows that it is not definitely true that not-A.

If we do not keep these operators separate, we are going to have some problems. At one point in his paper, Evans appeals to the principle:

(T) $\Delta A \rightarrow A$

As was said, the analogue, (T$\Box$), of this principle is valid for the operator "Definitely". But given the interpretation of 'Δ', as "It is determinate whether...", (T) is plainly invalid: If it is determinate whether A, it does not follow that A is true; A may be either determinately true or determinately false. One will have no great difficulty deriving, from (Eq) and (T), that "$\neg\Delta A$" is a valid schema.[♠] In the

---

[♠] For $\Delta A \rightarrow \Delta\neg A$, by (Eq), and $\Delta\neg A \rightarrow \neg A$, by (T); hence, $\Delta A \rightarrow \neg A$, so since, by (T), $\Delta A \rightarrow A$, it follows that $\Delta A \rightarrow A \,\&\, \neg A$, and so $\neg\Delta A$.

Pelletier's attempt to derive a contradiction from the conditional "$\nabla a{=}b \rightarrow \neg a{=}b$" is invalidated precisely by an appeal to (T). Pelletier in fact notes Evans's retraction of this slip, but he seems to miss the point. (In fairness, Pelletier refers to writers who hold that (T) is valid; I do not know if it is (T) or (T$\Box$) which they accept.) See pp.

presence of (R), or of any principle asserting that not every sentence is of indeterminate truth-value, contradiction is immediate. The operators '$\Box$' and '$\triangle$' are closely related, however. Given an operator '$\Box$', like "Definitely", for which (T$\Box$) and (D$\Box$) are valid, we can define an operator '$\triangle$', like "Determinately", for which (Eq), (C), and (D) are valid. Viz.:

$$\begin{aligned} \triangle A &\equiv_{df} \Box A \vee \Box\neg A \\ \nabla A &\equiv_{df} \neg\triangle A \\ &\equiv \Diamond A \;\&\; \Diamond\neg A \end{aligned}$$

(Eq) is then obvious; (C) is just the definition of '∇'; and, as mentioned above, (D) follows from (Eq) and (C). Conversely, given our operator '$\triangle$', we can define an operator '$\Box$':

$$\Box A \equiv_{df} A \;\&\; \triangle A$$

(T$\Box$) is then obvious. We may take (D$\Box$) as the definition of the dual, giving:

$$\begin{aligned} \Diamond A &\equiv_{df} \neg\Box\neg A \\ &\equiv \neg A \vee \nabla A \end{aligned}$$

But (C) and (Eq), again, are plainly invalid.

Operators akin to "It is determinate whether..." and "It is definitely true that..." are thus interdefinable. Our reading of '$\triangle$' as "It is determinate whether..." may now be further explained: To say that it is determinate whether A is to say that either A is definitely true or it is definitely false. Since operators such as "Definitely" are rather

---

483-4.

more often discussed in this connection, perhaps this reading is more helpful than the official interpretation with which we began.

2. What Evans Argued

For the purposes of our discussion here, I shall assume, as earlier, that Evans would hold the principles (C), (D), and (Eq), as well as the unobjectionable (R), to be valid.

Evans also assumes, *for the purposes of argument*, that the operator 'Δ' does not induce an opaque context. Evans is not arguing that no identity-statement is vague; he is arguing that there can be no vague objects. Now, I think that we should know well enough what a vague object was meant to be if we understood what Evans would need to prove to show that there can be no vague objects. As a first approximation, we may take the following: To say that there are vague objects is to say that the vagueness of a statement about such an object *may* be a consequence, not of how the object is *described*, but of the nature of the object itself. That is, whether certain statements of the form "Fa" are of determinate truth-value must depend, in respect of the term "a", only upon to what "a" refers; it can not depend, as Evans here puts it, upon how the bearer of "a" is 'described' or, in Fregean terminology, upon what sense the name "a" bears.[7]

To say that whether "Fa" is of determinate truth-value *may*, in certain cases, depend only upon to what "a" refers is to acknowledge that the explanation why "Fa" is not of determinate truth-value may be just that the expression "a" is vague. Evans's claim is that a sentence of the form "a=b" may be of indeterminate truth-value *only* if one of the terms "a" and "b" is vague, *only* if it is indeterminate to what the terms refer. Conversely, Evans's opponent holds that there may be (or are) identity-statements "a=b" which are of indeterminate truth-value, whose truth-value is indeterminate *not* because it is indeterminate to what "a" and "b"

---

[7] These remarks are in agreement with those of David Lewis, "Vague Identity: Evans Misunderstood", *Analysis* XLVIII, 3 (1988), pp. 128-30. Lewis's remarks are in a rather different terminology: But the point is that, for one who maintains that there are vague objects, the step from "∇Fa" to "$\lambda x(Fx)(a)$" must be valid.

In Lewis's terminology, only one who holds that a vague name does not "rigidly [denote] a vague object" can balk at the transition from "$\nabla a=b$" to $\lambda x(a=x)(b)$". I avoid such terminology, being rather unhappy about the use of the notion of rigid designation here: I suppose that the notion of necessity, the accessibility relation, with respect to which such names are meant to be rigid, is that relevant to a semantics for '∇'. That now seems no better an explanation than that '∇' should be transparent. Moreover, it does not seem quite right, since, as shall be shown below, one who maintains that there are vague objects may accept a modal semantics, based upon S4, which validates "$a=b \rightarrow \Box a=b$" but not "$\neg a=b \rightarrow \Box\neg a=b$". Hence, a given term need *not* refer to the same object 'in every world'.

refer, but because *the objects to which* they refer are indeterminate.[6]

A note on terminology is now required. I shall speak throughout of "transparent" operators. In my usage, an operator is transparent if it poses no barrier to predicate-abstraction: An operator 'Ω' is transparent if and only if "$\lambda x[\Omega(Fx)](a)$" follows from "$\Omega[\lambda x(Fx)(a)]$", and *vice versa*, so long as "a" is an expression of the appropriate sort. I shall simply call such expressions *names*, for our purposes (since it is common for the relevant class of expressions to exclude descriptions, as in the case of necessity).[7] Hence, the validity of "$(\forall x)(\forall y)(x=y \ \& \ \Box Fx) \rightarrow \Box Fy$" is a consequence, not merely of the transparency of '$\Box$', but of the above-mentioned principle (LL). If (LL) is valid, then '$\Box$' (or 'Δ') will not only be transparent but will be *extensional*, in the sense that "$(\forall x)(\forall y)(x=y \ \& \ \Box Fx) \rightarrow \Box Fy$" is

---

[6] Evans's opponent thus holds that, e.g., 'being definitely red' (or 'being definitely identical to a') is a legitimate 'property' of an object, since--so long as the *names* do not suffer some indeterminacy--whether "$\Box$Red(a)" (or "$\Box$a=b") is true must depend, in respect of "a", only upon to what it refers. Likewise, 'being such that it is determinate whether it is a' is a legitimate property of an object, *or so must one who maintains that there are vague objects hold*.

[7] The notion of a 'name' to which I am appealing here will have to be specified in more detail for any particular operator. In the case we are discussing, the relevant 'names' are those in which there is no essential indeterminacy of reference, as argued in the last section. The argument here does not depend upon any particular way of specifying these names.

As mentioned above, Lewis suggests extending the notion of rigidity to this case.

valid. But there might be good reason to question the validity of (LL), without questioning whether 'is definitely red' is a predicate satisfied by (vague) objects; without, that is, questioning the transparency of '□'.[10]

It is worth emphasizing that the argument just given depends upon the assumption, if such it can be called, that the ontological or metaphysical view that there are vague objects has a semantic component: Namely, that the fact that we refer to such objects has some explanatory force, that it explains certain features of our use of (apparent) names of such objects. The view in question is[11] that there are vague objects and (if it adds anything) that it is the vagueness of such objects which is responsible for the vagueness of certain statements, including identity-statements, which we make about those objects.

Now, it is tempting to conclude, from that argument, that, if there are vague objects, any operator which means something like "It is vague whether..." must be transparent. But one who maintains that there are vague objects need not hold that *every* such operator is transparent. She, like everyone else, can make a place for epistemic or otherwise intensional operators of this sort. What is essential to her case is that there may be such operators; that there might

---

[10] The distinction between transparency and extensionality will only become important when we discuss objections to the principle (LL) itself.

[11] As Lewis too notes.

be, or that we could introduce, an operator, which could plausibly be read "It is vague whether..." or "It is indeterminate whether...", which was transparent.

Conversely, one who denies that there are vague objects need not deny that operators like "It is vague whether..." may be transparent. What is in question is not just whether such an operator may be transparent, but whether, if such an operator is transparent, there are any identity-statements which are, in the sense of that operator, vague. That is: One might have the view that there is a transparent operator "It is vague whether..."; that some statements are (in that sense) vague and others are not; but that "For all x and y, it is not vague whether x=y" is valid.[12] On such a view, there would be no vague objects, for any identity-statement of indeterminate truth-value should be so because it was indeterminate to what the expressions "a" and "b" referred, not because to what they refer is indeterminate.

To summarize: The view that there are vague objects can not properly be characterized as the view that some identity-statements are vague. Almost everyone (including Evans) believes that some identity-statements are vague: Many of these people (including Evans, again) believe,

---

[12] Similarly, one would hold that, so long as "a" and "b" are (in the appropriate sense) names, "It is not vague whether a=b" is valid.

It is, indeed, not even clear that it is relevant to the dispute whether the operator "It is vague whether...", in English, is transparent. I myself have no settled opinion on this question nor on the question whether it is.

however, that the vagueness of such statements is a product, not of the vagueness of the *objects* themselves, but of our *language*. One who maintains that there are vague objects must, additionally, hold that operators such as "It is determinate whether..." may be transparent and that, in that sense, not every identity-statement is of determinate truth-value *even if it is determinate to what the relevant expressions refer*: For only if such an operator may be transparent can it be said that the truth-value of a sentence containing it (and so the vagueness of a sentence) depends not upon how the objects to which we refer are 'described' but rather upon the nature of the objects themselves; only if such an operator may be transparent does the hypoth sis that we refer to such objects serve any explanatory function.

We may conclude that it is not to respond to Evans, but to concede his point, to claim that "It is indeterminate whether..." can not be, and any similar operator would not be, transparent. For, if so, then the indeterminacy of a given statement depends upon how we refer to the objects to which we refer; that is to say that the vagueness of the statement is a product not of *reality* but of *language*; and that is to say, at best, that the claim that we refer to vague objects can be made only in a theoretical vacuum.

Given this account of what an argument designed to show that there can be no vague objects must accomplish, we may formulate a simple restriction upon such arguments. To show that there can be no vague objects, what one must show is that, if 'Δ' is an operator which can plausibly be construed as "It is determinate whether...", then, if 'Δ' is transparent, no identity-statement is of indeterminate truth-value.[13] But that is to say that the argument must show that there is some *special* problem which arises if we treat 'Δ' as an transparent operator. To contrapose: If "ΔA" is valid *whether or not* 'Δ' is transparent, then we have not been given an argument that there are no vague objects. Rather, we have been given an argument that the principles taken to govern 'Δ' are inappropriate for an operator intended to be read s 'Δ' is intended to be read, namely, as "It is determinate whether...".

3. Whence the Contradiction?

The question before us now is, therefore, whether Evans's argument, the assumptions made thus far being g'anted, establishes his claim. I am therefore granting that the formal argument Evans sets out is one which must be accepted by one who maintains that there are vague objects. The question is whether one who maintains that there are

---

[13] I shall henceforth drop the qualifier "so long as 'a' and 'b' are in the relevant sense names". When I speak of an identity-statement, I shall always mean a statement which asserts the identity of two objects **a** and **b**.

vague objects is thus committed to the truth of a contradiction.

Evans argues, recall, that we can derive, from the assumption "$\nabla a=b$", the conclusion "$\neg a=b$", which, he says, contradicts the assumption. As was said earlier, it is not obvious why this should be so. We may take Evans to have meant that, if "$\neg a=b$" is true, then, since "$a=b$" is false, its truth-value is determinate. Hence, there would seem to be an unrecorded step from "$\neg a=b$" to "$\Delta a=b$". Presumably, Evans intended us to construe the argument in just this way: He writes that "$\neg a=b$" contradicts the assumption "that the identity-statement '$a=b$' is *of indeterminate truth-value*".[14] But to what principle is Evans appealing here? Just what justifies this transition?

The simplest principle to which we might take Evans to be appealing is:

(N) $A \rightarrow \Delta A$

However, if he means to appeal to this principle, then his argument might have avoided questions of identity altogether. Viz.:[15]

$A \rightarrow \Delta A$	(N)	(1)
$\neg A \rightarrow \Delta\neg A$	(N)	(2)
$\Delta A \leftrightarrow \Delta\neg A$	(Eq)	(3)
$\neg A \rightarrow \Delta A$	(2,3)	(4)
$A \vee \neg A \rightarrow \Delta A$	(1,4)	(5)
ΔA	(5)	

[14] Evans. My italics.

[15] Note that I am assuming the validity of classical logic here. That is not to say that I am assuming Bivalence.

No appeal to the transparency of 'Δ' is required: "ΔA" is valid whether 'Δ' is transparent or not.

Thus, if Evans intends to appeal to (N), he has no argument against the existence of vague objects. As I argued in the last section, the possibility of an argument such as that just given shows, not that, if 'Δ' is transparent, then "$\Delta a=b$" is valid; but, rather, that the theory in question--namely, (N)+(C)+(Eq)--is an inadequate theory for an operator which is meant to express vagueness (or lack thereof), as no statement is, in the sense of this operator, vague.

It might also be suggested that Evans intends to appeal to some modal claim regarding 'Δ'. He writes that, "if Δ determines a logic at least as strong as S5", then "$\Delta(\neg a=b)$" is derivable from "$\nabla a=b$".[16]

As a version of the characteristic axiom of S5, i.e., "$\Diamond A \rightarrow \Box\Diamond A$", we may take:

(5Δ) $\nabla A \rightarrow \Delta\nabla A$

We also need to appeal to the following distribution principle:[17]

---

[16] Evans. As Bob Stalnaker pointed out, Evans's remarks here are influenced by the 'slip' mentioned earlier, which is to say that he seems to be best interpreted as discussing not "Δ" but "$\Box$".

[17] David Lewis pointed out to me that the standard distribution principle, which would allow the inference from '$\Delta(A \rightarrow B)$' to '$\Delta A \rightarrow \Delta B$' is invalid. To see this, just let A be the falsum. Then '$A \rightarrow B$' is true, and so determinate; similarly, A is false, so 'ΔA' is determinate; but then 'ΔB' is true, whatever B is.

To prove the restricted version, we use the equivalence between $\Box$ and Δ. '$\Delta(A \rightarrow B)$' is equivalent to '$\Box(A \rightarrow B) \vee$

$$\frac{\Delta(A \rightarrow B)}{A \,\&\, \Delta A \rightarrow \Delta B}$$

The argument is then as follows:

$\nabla a{=}b \rightarrow \neg a{=}b$	Evans's argument
$\Delta(\nabla a{=}b \rightarrow \neg a{=}b)$	Necessitation (analogue)
$\nabla a{=}b \,\&\, \Delta\nabla a{=}b \rightarrow \Delta(\neg a{=}b)$	(1) Distribution
$\nabla a{=}b \rightarrow \Delta\nabla a{=}b$	(5Δ)
$\nabla a{=}b \rightarrow \nabla a{=}b \,\&\, \Delta\nabla a{=}b$	(2) Last step, PC
$\nabla(a{=}b) \rightarrow \Delta(\neg a{=}b)$	(1,2) PC
$\therefore\ \nabla a{=}b \rightarrow \Delta a{=}b$	(Eq), PC

That, indeed, is contradictory.

What sort of justification can be given for (5_{Δ}), however? The most natural which comes to mind is the following: Every proposition is either (definitely) true, (definitely) false, or (definitely) neither true nor false. We may take "ΔA" to be (definitely) true if, and only if, A is either definitely true or definitely false; otherwise, it is (definitely) false. Similarly, "∇A" is (definitely) true if, and only if, A is (definitely) neither true nor false. Hence, (5Δ): If it is indeterminate whether A, it is (definitely) true that A is neither true nor false; hence, it is determinate whether it is neither true nor false; hence, it is determinate whether it is indeterminate whether A.[18]

$\Box\neg(A \rightarrow B)$'; hence, by distribution, PC: '$(\Box A \rightarrow \Box B) \vee \Box(A \,\&\, \neg B)$'; so, '$(\Box A \rightarrow \Box B) \vee (\Box A \,\&\, \Box\neg B)$'; hence, by PC, '$\Box A \rightarrow (\Box B \vee \Box\neg B)$'. But the consequent is just 'ΔB' and '$\Box A$' is equivalent to '$A \,\&\, \Delta A$'. Thus: '$\Delta A \,\&\, A \rightarrow \Delta B$'.

[18] I should thank Bob Stalnaker for suggesting this as a possible justification; the suggestion greatly improved this section of the paper. In previous drafts, I had found myself rather lost for a justification, since I was concentrating instead upon (the equivalent) "$\nabla\Delta A \rightarrow \Delta A$". At first sight, it is difficult to see why one should accept this principle.

This justification of (5Δ), however, is one which one who maintains that there are vague objects has no reason to accept. Recall that we may define an operator '□' as follows:

□A ≡df A & ΔA

Note, then, that "□A" is either definitely true or definitely false: For either A is definitely true, definitely false, or definitely neither true nor false. If A is either definitely true or definitely false, then "ΔA" is true; so "□A" is definitely true or definitely false, as A is true or false. Similarly, if A is definitely neither true nor false, then "ΔA" is false, so "□A" is false. That is: "□A" is (definitely) true if, and only if, A is definitely true; otherwise, it is (definitely) false. The justification for (5Δ) thus also provides a justification for this principle:[17]

Δ□A

It follows that, by making use of the operator '□', we have the means for speaking about our (by hypothesis) vague subject matter *with no vagueness whatsoever, using an opera-*

Nonetheless, it is valid, given the suggested interpretation of 'Δ', since the antecedent is necessarily false.

[17] A formal derivation of this schema can be given, but it is somewhat space-consuming, due to the fact that 'Δ' is here the primitive operator. If we introduce '□' as our primitive operator and define 'Δ' as earlier, the proof is rather easier. For 'Δ□A' is equivalent to '□□A ∨ □¬□A'; i.e., to '□□A ∨ □◊¬A', which is provable in S5.

tor which is transparent.[20] All one need do is take care to insert '$\Box$' before anything one writes or says, and whatever vagueness may have affected the original sentence will be removed: All one's utterances will be (definitely) true or (definitely) false.

But that is a possibility which one who holds that there are vague objects has reason to reject: If the objects themselves are responsible for the vagueness of (identity-) statements containing names of them, then *all* our talk about such objects must be, in principle, vague. Surely the picture proposed, that there is vagueness in reality--*any* sort of vagueness, whether that of properties or of objects--could hardly be better explained than in terms of the claim that the vagueness which characterizes our talk about such objects is an essential feature of it, one which can not be eliminated merely by the introduction of as-yet-unheard-of operators into the language. For it is not our language which is responsible.

That is just to say that on a conception of vagueness according to which there is 'vagueness in reality', vagueness is ineradicable. While it will, on such a view, be possible to introduce operators which "strengthen" vague

---

[20] Note that Pelletier's argument, in terms of many-valued logic, relies upon the same sort of claim: His J-operators may be defined in terms of '$\Box$', subject to S5. In a slightly different, but hopefully self-explanatory, terminology: $J_t A \equiv_{df} \Box A$; $J_f A \equiv_{df} \Box\neg A$; $J_u A \equiv_{df} \neg\Box A \& \neg\Box\neg A$. It is then not too difficult to derive a contradiction from "$J_u(a=b)$". See Pelletier, pp. 488-90.

statements--for example, "Definitely"--no such operator can *eliminate* vagueness: If A is vague, so, in principle, is "Definitely: A".[21] But the semantic assumptions required to justify (5Δ) are strong enough to justify the introduction of operators, like '$\Box$', which eradicate vagueness from vague statements. Indeed, the assumption that every (apparently vague) statement is either definitely true, definitely false, or definitely neither true nor false amounts to the assumption that vagueness is eradicable and therefore begs the question against one who maintains that there are vague objects.[22]

In any event, it is hardly likely that Evans intended to appeal to this sort of modal principle. For he says, recall, that "*if* Δ determines a logic at least as strong as S5", then "$\Delta a=b$" is derivable from "$\nabla a=b$". So he is not intending to appeal to any such claim as part of his original argument. The fact that we can derive "$\neg a=b$" from "$\nabla a=b$" is the real problem: The remark about S5 is but an aside. I pursue it only to show that that avenue is definitely closed.

---

[21] See Michael Dummett, "Wang's Paradox", in his *Truth and Other Enigmas* (Cambridge MA: Harvard University Press, 1978), pp. 248-68, at p. 257.

[22] It is worth emphasizing that this argument shows that one who maintains that there are vague objects should *not* attempt to provide a semantics for vague statements in terms of a many-valued logic. (Note that this claim depends upon the results of sections 5 and 6.)

Pelletier makes precisely the opposite suggestion, for reasons I do not understand: See p. 482.

But we ought nonetheless to be extremely puzzled by this last remark from Evans's paper. Surely, if, as Evans says, "¬a=b" *contradicts* the original assumption that "∇a=b", then "¬a=b" must be at least as strong as "¬∇a=b": If one statement *contradicts* another, then it must at least imply the negation of that other statement. If so, then, for whatever reason, "¬∇a=b"--i.e., "Δa=b"--must follow from "¬a=b". But why then does Evans say that it is only *if* the logic governing 'Δ' is at least as strong as S5, then we can derive "Δa=b"? To this question, I can give no definitive answer: But I think that Evans was trying to express a quite different distinction between what he can and what he can not prove, to which we now turn.

4. Whence the Contradiction

The most natural suggestion to make at this point is that it is not the axiom (N) but the rule of inference (N*) which is valid:

$$\frac{A}{\Delta A}$$

Such a rule is surely valid: If A is true, then it is indeed determinate whether A is true. That is all that is required of a valid rule of inference: That its conclusion be true whenever its premises are true.

The point of introducing the rule (N*) is to get the effect of (N) without its disadvantages. Hence, we must renounce conditional proof: For if conditional proof is valid, we shall be able to derive (N) from (N*). Similarly, we must renounce proof by cases: For if proof by cases is valid, we shall once again be able to demonstrate the validity of "ΔA" without appeal to the transparency of 'Δ':

$$\frac{\begin{matrix}[A]\\ \Delta A\end{matrix}\qquad\begin{matrix}[\neg A]\\ \Delta\neg A\\ \Delta A\end{matrix}\qquad A \vee \neg A}{\Delta A}$$

(Note that '[A]' indicates that A has been discharged.) And, again, if proof by *reductio* is valid:

$$\frac{\begin{matrix}[\neg A]\\ \Delta\neg A\\ \Delta A\end{matrix}\qquad \neg\Delta A}{A}$$

Hence, $\neg\Delta A \vdash' A$. Therefore, by substitution: $\neg\Delta\neg A \vdash' \neg A$. But, "$\neg\Delta\neg A$" is equivalent to "$\neg\Delta A$", by (Eq). Hence, $\neg\Delta A \vdash' \neg A$. Hence, $\neg\Delta A \vdash' (A \,\&\, \neg A)$; so, by *reductio*, again: $\vdash' \Delta A$.

Thus, if any one of conditional proof, proof by cases, and proof by *reductio* is valid,[23] we shall be able to show that "ΔA" is valid, without appeal to the transparency of 'Δ'. Just as in the case of (N), we shall yet be without an argument that there are no vague objects.

We must, therefore, abandon conditional proof and its kin: More precisely, we must renounce appeal to (N*) within

[23] If we assume the validity of classical propositional logic, it is not difficult to show that proof by cases, conditional proof, and proof by *reductio* stand or fall together.

so-called subordinate deductions; within, that is, deductions from premises which may subsequently be discharged. Such rules are sometimes called "auxiliary" rules of inference; I shall refer to rules such as (N*) as 'rules of deduction'. We may thus interpret Evans as having intended that the rule (N*) be valid as a rule of deduction.[24]

With the rule (N*) in hand, we can complete the derivation of the contradiction (omitting the lambda-notation):

$\nabla a=b$	Premise
$\neg\nabla a=a$	(R)
$\neg a=b$	(LL)
$\Delta(\neg a=b)$	by (N*)
$\neg\nabla a=b$	by (D)
$\nabla a=b \ \& \ \neg\nabla a=b$	first and last lines

Contradiction. Because we have been forced to renounce application of (N*) within proofs by *reductio*, however, we can not infer that $\neg\nabla a=b$. Even given the hypothesis that '∇' is transparent, we can not prove, *via* (N*), that "$\Delta a=b$" is valid; what we can do is derive a contradiction from "$\nabla a=b$".[25]

---

[24] Oft-expressed worries about the validity of conditional proof thus prove relevant.

Such rules have a place in other contexts: For example, one can formulate consistent theories of truth using such rules. See Vann McGee, "Applying Kripke's Theory of Truth", *Journal of Philosophy* LXXXVI (1989), pp. 530-9. See also Harvey Friedman and Michael Sheard, "An Axiomatic Approach to Self-Referential Truth" (draft), in which the term "auxiliary rule of inference" appears in a similar context.

[25] We can give models for this language as follows. Let the underlying structure of the models be that for a quantified version of S4, with the domain fixed, in the sense that, if an object exists in one world, it exists in all. (We may, for the moment, abstract from the problem of existence.)

Assuming, for the moment, that the (amended) formal argument Evans has presented is one his opponent must accept, the only problem with his argument is now the final step: Namely, that by which he passes from the intermediate conclusion that no statement of the form "∇a=b" can be true, to the ultimate conclusion that there can be no vague objects. We shall discuss this step after we discuss Evans's appeal to the Indiscernibility principle.

5. Formulating the Indiscernibility Principle

Given the utility we have found the notion of a rule of deduction to have in this context, one might well seek to defend the view that "∇a=b" might be true by denying Evans's appeal to the Indiscernibility Principle, in the form in

---

Instead of taking truth to be truth at some 'actual' world, we define truth as truth in all worlds. Define "ΔA", as usual, as "□A ∨ □¬A". Then (Eq) is obvious. Take (D) as the definition of the dual. One may also verify that (LL) holds (on the assumption that "a=b → □a=b" is valid).

Suppose A is true. Then A is true in all worlds; so "□A" is true at all worlds; so "ΔA" is true at all worlds; so "ΔA" is true. Hence, (N*) is valid.--Conditional proof, however, clearly fails: "A → ΔA" is not valid, since A may be true at one world, but not true at another.

"¬Δa=b" is not satisfiable. For suppose that "¬Δa=b" is true at some world w. Then "¬□¬a=b" is true at w. Hence, there is some world w', accessible from w, at which "a=b" is true. But "a=b → □a=b" is valid; hence, "□a=b" is true at w'. And so, "Δa=b" is true at w'. Hence, "¬Δa=b" is not true at all worlds; so, "¬Δa=b" can not be true.

Nonetheless, "Δa=b" is not valid. Let there be two worlds, w and w'. Take w' accessible to w, though not conversely. Let "a=b" be false at w; true, at w'. Then "Δa=b" is not true at w and so is not true.

It can be shown that S4+(N*), which I call "V4", is complete with respect to this class of models.

which it is required for his argument. In this section, we shall look at the prospects of such a move.

Earlier, we recorded this principle in the form:

(LL) $a=b \ \& \ \lambda x(Fx)(a) \rightarrow \lambda x(Fx)(b)$

Given the transparency of '$\Box$' and 'Δ', it is then easy to derive the two schemata:

$a=b \rightarrow \Box a=b$

$a=b \rightarrow \Delta a=b$

Both of these principles have been questioned.[24] Consider, for instance, the latter schema: Suppose that "a=b" is neither (definitely) true nor (definitely) false; then "$\Delta a=b$" is false (or, at least, not true); hence, plausibly, "$a=b \rightarrow \Delta a=b$" is not true. No instance of this conditional can possibly be false, since, if "a=b" is true, so is "$\Delta a=b$": But it does not follow that the conditional is valid.

Hence, (LL) itself need not be a valid schema: If "a=b" is neither definitely true nor definitely false, then "Fa" might be true though "Fb" is neither true nor false. (Above, we took the predicate "$F\xi$" to be "$a=\xi$".) Thus, both the antecedent and the consequent might be neither true nor false; plausibly, the conditional is then itself neither true nor false. To assume the validity of (LL) would thus

---

[24] See, for example, B.J. Garrett, "Vagueness and Identity", *Analysis* XLVIII, 3 (1988), pp. 130-4.

appear to beg the question against one who maintains that there are vague objects.

One who maintains that there are vague objects may adopt, instead of (LL), the rule of deduction (LL*):[27]

$$\frac{a=b \qquad Fa}{Fb}$$

After all, if **a** *is* **b**, then **a** and **b** must share all their properties: So, if it is true that a=b, then, if it is true that, say, △Fa, it must also be true that △Fb.

If appeal to this rule *alone* is allowed, Evans's proof fails. For we can not prove that, if "Fa" is true and "¬Fb" is true, then "¬a=b" is true. That is, we can not derive the rule (II*):

$$\frac{Fa \qquad \neg Fb}{\neg a=b}$$

We might try to do so as follows:

$$\frac{\dfrac{Fa \quad [a=b]}{Fb} \qquad \neg Fb}{\neg a=b}$$

But this proof by *reductio* is invalid, since appeal to (LL*), a rule of deduction, is invalid within subordinate deductions. In principle, then, one may accept the validity of (LL*) while denying that of (II*). Thus, if only the rule of deduction (LL*) is accepted as valid, we can not show

---

[27] Henceforth, I omit the predicate-abstraction operators, since we are assuming that "λx(Fx)(a)" is equivalent to "Fa", in the cases which interest us.

that no sentence of the form "∇a=b" is true. Denial of the validity of (LL) and its replacement by (LL*) will block the derivation of a contradiction from "∇a=b".[28]

However, Evans does not require appeal to (LL) itself to derive the contradiction. Rather, he requires only appeal to the rule (II*). What he needs is to be able to derive "¬a=b" from "¬∇a=a" and "∇a=b". The rule (II*) would license this transition.[29]

One who wishes to defend the view that sentences of the form "∇a=b" might be true must deny, as we have seen, not only that (LL) is valid, but also that (II*) is valid. But it is difficult to see on what ground the denial is to be

---

[28] We can give models for such a language. Let the underlying structure be that for a quantified version of S5, *without* the assumption that a=b → □a=b. We require only that, if "Fξ" does not contain '□', then, if "a=b" is true at a world, "Fa ↔ Fb" is also true at that world. We again define truth as truth in all worlds: Hence, (N*) is valid.

It is straightforward to prove, by induction on the number of occurrences of '□' in "Fξ", that "□a=b → (Fa ↔ Fb)" is valid, for *any* predicate "Fξ". Hence, if "a=b" and "Fa" are (absolutely) true, then, since "a=b & Fa" is true at all worlds, so must "Fb" be true at all worlds. Hence, (LL*) is valid.

We may show simultaneously that (II*) fails--and so is independent of (LL*)--and that "¬□a=b & ¬□¬a=b" is satisfiable. Let there be two worlds w and w′. Let "a=b" be true at w; false, at w′. Then, of course, "□a=a" is true at both w and w′; but "¬□a=b" is also true at both w and w′. Hence, both "□a=a" and "¬□a=b" are true, though "¬a=b" is not true, since "a=b" is true at w. Furthermore, "¬□¬a=b" is true at both w and w′. Hence, "¬□¬a=b" is true and so "¬□a=b & ¬□¬a=b" is true.

[29] There is some reason to think that Evans was aware of this problem. If (LL) were the principle to which he was appealing, he could simplify the proof. Viz.: a=b → Δa=b; hence, ¬Δa=b → ¬a=b; so, ∇a=b → ¬a=b. But who knows?

made. It is one thing to argue, as we did earlier, that if "Fa" is true and "Fb" is not true, then "a=b" need not be false, but need only fail to be true. Such an argument is sufficient to call the validity of the schema (II)--i.e., "Fa & ¬Fb → ¬a=b"--into doubt. (For both the antecedent and the consequent might then be neither true nor false.) This argument should remind us of the argument for rejecting (LL), which we discussed earlier: If "a=b" is neither true nor false, then "Fa" might be true, though "Fb" too is neither true nor false. That is to say, roughly, that if it is indeterminate whether **a** is **b**, it might be similarly indeterminate whether they 'share all their properties': There might be biconditionals of the form "Fa ↔ Fb" which are neither true nor false.

It is another thing to suggest that, if it is indeterminate whether **a** is **b**, it might, in fact, be *false* that they share all their properties, that there might be some predicate "Fξ" such that "Fa" is true though "Fb" is false.[30]

The talk of 'properties', which I have used heuristically, is, of course, rather slippery. We need now to remove the appeal to the notion of a property.

The important disanalogy between the rejection of (LL) and the rejection of (II*) is that the latter depends upon

---

[30] Note that, if so, the semantic counterpart of (LL), which states that the truth-value of "a=b" is the same as that of the infinite conjunction of all biconditionals "Fa ↔ Fb", fails.

the presence of the operators '□' and '△' in the language. Now, if we assume that the language contains only, as it were, 'ordinary' predicates (which do not contain these operators), then there is no reason to question the validity of (II*).[31] (If, for example, "a is red" is true and "b is red" is false, then "a=b" is false.) One may yet wish to reject (LL), for reasons like those just discussed. *That* ground for the rejection of (LL) does not depend upon the presence of such operators as '□' and '△'; it requires *only* the claim that "a=b" itself may be of indeterminate truth-value, for that claim *entails* the corresponding claim that, if so, biconditionals of the form "Fa ↔ Fb" may be of indeterminate truth-value.

The rejection of (II*), on the other hand, depends upon the presence of such operators as '□' and '△',[32] upon the assumption that they are transparent, *and* upon certain assumptions about the truth-values of sentences containing such operators. If we *assume* the transparency of such operators, if we so explain '□' that, if "Fa" is not true, then "□Fa" is *false*, *and* if we assume that "a=b" might be neither true nor false, then we shall find ourselves compelled to

---

[31] As was reflected in the model developed for such a language above.

[32] More precisely, the case against (II*) depends upon specific assumptions about what sorts of predicates the language contains. The assumption that the language contains operators like '□' and '△' is one such assumption, appeal to which is natural in this context.

reject (II*). (Evans shows us why.) But it is not clear that we are entitled to make such assumptions.

6. That the Re-formulation of Indiscernibility Amounts Not to a Reply to Evans But to Capitulation

The intuition, which I hope some share, that any "real property" is subject to Leibniz's Law, at least in the form of (LL*) and (II*), finds a theoretical justification here. Leibniz's Law may require re-formulation,[33] in the form of (LL*) and (II*), in the case of languages whose semantics allow for a (non-trivial) distinction between rules of deduction and more ordinary rules of inference (in, that is, cases in which the deduction theorem fails *anyway*). But it ought not be re-formulated due to the presence of sentential *operators* of certain sorts: The question whether a new operator which we wish to introduce is transparent ought to be answered by determining whether, if it is taken to be transparent, Leibniz's Law remains valid, in whatever form it was taken to be valid *before* the introduction of the new operator. Surely it is quite special pleading to argue that, conversely, we are so convinced that this new operator is transparent that we *must* revise Leibniz's Law.

---

[33] I am not suggesting here that Leibniz's Law ought to be re-formulated at all.

I do not know whether the independence of (LL*) and (II*) from (LL) can be demonstrated, if the underlying logic is classical. (It is fairly easy to establish the independence result if we do not require that the underlying logic be classical.)

To put the point differently: I argued above that Evans's assumption that (LL) is valid begs the question whether there are vague objects. What I am now arguing is that the assumption that (II*) is valid does *not* beg that question. It is, of course, true that the assumption that (II*) is valid, even if the predicate in question contains '□' or '△', is *inconsistent* with the claim that "▽a=b" might be true: But that does not imply that its assumption begs the question whether "▽a=b" might be true. (If it did, it would be impossible to argue at all.)

The invalidity of (LL) is an immediate consequence of the assumption that "a=b" might be neither true nor false. The validity of (II*), on the other hand, is inconsistent with the following trio of claims: First, that "a=b" might be neither true nor false; Secondly, that "□A" and "△A" are false, if A is not true; and, Thirdly, that '□' and '△' are transparent operators. It is not at all obvious that one is entitled simultaneously to make stipulations about the transparency of an operator like '△' *and* to make stipulations about the truth-values of sentences which contain it.

On the contrary, it would seem that one ought explain such an operator--settle how the truth-value of a sentence "△A" containing it is determined by that of "A" itself--and then *ask* whether it is transparent. Or, conversely, one ought settle upon the transparency of the operator and then *ask* how, consistent with its transparency, the truth-conditions of sentences containing it may be explained.[34] To answer either of these questions, one must make reference to that form of Leibniz's Law which is properly taken to be valid prior to the introduction of the new operator: Hence, *if* the new operator is transparent, it will be subject to whatever form of Leibniz's Law is valid for sentences which do not contain it.

The notion of transparency in use here, is not, of course, self-explanatory: One might well wonder whether it is any less slippery than the notion of a property. One might similarly wonder whether the reliance upon the analogy with the introduction of a new operator should be trusted: Perhaps the remarks about the introduction of a new operator are just irrelevant to the situation we face when an operator is already in general use. It is therefore worth re-emphasizing the role the notion of transparency is playing in this discussion.

The difficulty is that the rejection of (II*) can not help but raise the question whether '□' and '△', as they must then be understood, are not merely *epistemic* operators.

---

[34] Operators like '△' form a special class, since they are truth-functional--or, at least, are intended to be by one who adopts the view we are considering. The sorts of remarks being made here apply also to other sorts of operators, however, and their force may be more apparent in such cases.

The general problem is reminiscent of questions Dummett has often raised concerning whether certain logical laws are *in harmony with* other laws. See his *The Logical Basis of Metaphysics* (Cambridge MA: Harvard University Press, 1991) for an extended discussion of this problem.

As was said earlier, what is really troubling Evans is the following: Of course it is true that some identity-statements are vague; but, if it is vague whether **a** is **b**, that is not because **a** and **b** are vague objects. Rather, the vagueness of the statement is a product of our epistemic limitations and of the vagueness of our language. Formally, Evans's opponent is committed to the transparency of (some) such operators as "It is determinate whether...". If what transparency requires is itself in dispute, this formulation will help us less than it otherwise might. What we may ask instead is whether the rejection of (II*) is bound to reinforce Evans's true worry: We may ask, that is, whether the claim that sentences containing '□' and '△' are not subject to (II*) is bound to invite the charge that, if so, such operators must be epistemic ones.

And, indeed, it is: For there are plainly epistemic (or otherwise intensional) operators which are subject to (LL*), but for which (II*) fails. For example, let '**θ**A' be read as "Linguistic conventions so far laid down and the non-semantic facts together determine that A".[36] Plainly, if A is true, then '**θ**A' is true. So (N*) is valid. Similarly, if "a=b" and "Fa" are true, then "Fb" is true. (If a=b and linguistic-conventions-plus-reality determine that Fa, then linguistic-conventions-plus-reality determine that Fb.) But

---

[36] This interpretation was suggested by remarks made by McGee, in a rather different context. See p. 537.

(II*) fails: If linguistic-conventions-plus-reality determine that Fa, but it is not the case (i.e., it is false) that linguistic-conventions-plus-reality determine that Fb, it does not follow that **a** is not **b**. Rather, it follcws only that it is not the case that linguistic-conventions-plus-reality determine that **a** is **b**. Further linguistic conventions which we might lay down--i.e., a more precise specification of what we mean by "a" and "b"--might decide the question of their identity either way.[36]

Given such an operator, then, there might well be some true sentences of the form "$\neg\Box\neg a=b$ & $\neg\Box a=b$"--i.e., of the form "$\nabla a=b$". But one who accepts the legitimacy of this operator is in no way committed to the existence of vague objects: So far as this particular operator is concerned,[37] the 'vagueness' of a statement is but a matter of our not having stipulated sufficiently many linguistic conventions to determine its truth-value. Vagueness, so far as this operator is concerned, is indeed a product of epistemic and linguistic phenomena.

Thus, rejection of the validity of (II*) would seem to amount to acceptance of Evans's claim that, if some sentence

[36] Note, first, that the rule $\Box Fa, \Box Fb \vdash' \neg a=b$ is just equivalent to (II*) in the presence of (N*). Note, secondly, that the rule $Fa, Fb, \vdash' \neg\Box a=b$ is that which one attracted to this sort of view might well propose as a replacement for (II*).

[37] I say "so far as this particular operator is concerned" because one who maintains that there are vague objects may nonetheless accept its legitimacy.

of the form "∇a=b" is true, then the operator '∇' is an epistemic one. For Evans, or one who agrees with him, may freely accept that *there are* operators which are subject to (LL*), though not to (II*). If so, then, of course, there may be some true sentences of the form "It is vague whether a=b". But Evans had not intended to deny that there are vague identity-statements: He thus had no need to deny that there may be operators subject to (LL*), though not to (II*). What he must deny, rather, is that, so understood, "It is vague whether..." is transparent or non-epistemic: And, indeed, at least one such operator is plainly epistemic.

I admit that I am hedging my bets. It is, in my opinion, implausible that any operator which is not subject to (II*) can be shown to be 'non-epistemic'. But it is difficult to argue that there is no *possible* interpretation of '□' and '△' which would serve the purposes of one who wished to defend the view that "∇a=b" might be true, 'in a non-epistemic sense'. To do so would, as has become clear, require the resolution of some sticky issues involving the notion of a 'property' and the related notion of transparency. However, it should also have become clear that any attempt to defend this view will have to negotiate a number of obstacles. These obstacles are not formal: We have, after all, seen precisely what is required of a formal system compatible with such a view. But neither does the existence of such a formal system guarantee the coherence of the view.

7. The Existence of Vague Objects

For the purposes of this section, I shall assume that the principle of Indiscernibility is valid in its classical form: $a=b \ \& \ Fa \rightarrow Fb$. That is, I shall assume that, despite the caveat at the end of the last section, it is not possible to defend the view that there are vague objects by way of the rejection of the validity of (II*).[38] If so, then one who wishes to defend the existence of vague objects must accept as valid Evans's derivation of a contradiction from "$\nabla a=b$"; she must, therefore, grant Evans that there neither are nor could be any true sentences of the form "$\nabla a=b$"; the question is whether she must not also grant that "$\Delta a=b$" is valid.

Whether Evans's argument shows that there can be no vague objects may now seem to be but a terminological question. We know what Evans's argument shows and what it does not show: It does show that there is no true sentence of the form "$\nabla a=b$"; it does not show that "$\Delta a=b$" is valid. If we identify the view that there are vague objects with the view

---

[38] As mentioned in the last section, one could assume the validity of (LL*) and (II*), and Evans's argument would still succeed. However, the only reason given for the abandonment of (LL) itself is that "a=b" might have some truth-value other than true or false, an assumption I have now rejected.

that there are (or might be) true sentences of the form "∇a=b", then Evans has shown that there are no vague objects. If, on the other hand, we identify the view that there are vague objects with the view that "Δa=b" is not valid, then he has not.

But the dispute is not merely verbal. Evans certainly took himself to be arguing a metaphysical point, namely, that there can be no vague objects. Any argument for such a conclusion must rest upon some characterization of the nature of the dispute; in this case, it rests upon a characterization of the view that there are vague objects. Evans's view, I suggest, can only have been that one who maintains that there are vague objects is committed to the claim that there may be true sentences of the form "∇a=b". Indeed, he opens his paper by saying that one who maintains that there are vague objects is committed to maintaining that it can be "a *fact*" that a particular identity-statement is of indeterminate truth-value.[37] This is not an unnatural way to understand the view: The dispute does not concern whether some identity-statements are vague; the view that there are vague objects is the view that the vagueness in question is due not to language but to the nature of reality itself; and that view may be explicated as the view that it might be a 'fact' that some identity-statements are of indeterminate truth-value.

---

[37] Evans.

Nonetheless, in the remainder of this section, I shall argue that there is a view, which can plausibly be identified as committed to the existence of vague objects, which is *not* committed to the possibility that there might be truths of the form "$\nabla a=b$". I shall thus argue that Evans's claim to have proven that there are no vague objects fails, even though the (formal) argument he gives is valid. For we need not accept Evans's characterization of this view, and there is an alternative which can be independently motivated.

It is worth considering, for a moment, an objection to my claim that we need to resolve the question how to characterize the view that there are vague objects. The objection is that we can extend Evans's argument to show that "$\Delta a=b$" is valid. Evans has, it is now being granted, shown that we may deduce a contradiction from "$\nabla a=b$". Hence, no sentence of the form "$\nabla a=b$" can possibly be true (since no contradiction is true). Hence, since every statement is either true or false, (each instance of) "$\nabla a=b$" must be false; it follows that (each instance of) "$\Delta a=b$" must be true, and so that "$\Delta a=b$" is, as a schema, valid.

This argument is naturally understood as appealing to the principle that every statement[40] is either true or false. It is this principle that licenses the crucial in-

[40] The word 'statement' here is used in the sense of 'sentence fit to be uttered assertorically' and so in the sense of 'sentence fit to be assigned truth-value'.

ference from the non-truth of "▽a=b" to its falsity. Plainly, one who wishes to reject Evans's argument must reject this principle, the principle of Bivalence.[41] More precisely, she must reject the claim that every statement *about vague objects* (and, in particular, every statement of the form "△a=b") is either true or false. That is to say, one who rejects Evans's argument must reject the principle of Bivalence, as it applies to statements about vague objects.

Plausibly, the principle of Bivalence is yet more intimately connected to our subject. The (metaphysical)[42]

---

[41] Bivalence is often said to be the principle that every statement is *determinately* either true or false. The qualifier 'determinately' does help, I think, to convey what is intended--i.e., that Bivalence is a stronger principle than is Excluded Middle. But, on the other hand, I do not think that, ultimately, it really helps us to *explain* the principle of Bivalence.

The use of the word 'determinately' in this context should not be confused with the use of the same word in this paper.

My own view is that Bivalence can not be distinguished from Excluded Middle until the notion of truth is itself explained as a notion of semantic theory, rather than, say, directly in terms of Convention T. The problem is thus to explain why the notion of truth is needed in semantic theory, why the explanation directly in terms of Convention T will not suffice to explain the notion as it is there needed, and, finally, how the notion, as it is needed in semantic theory, is connected to the intuitive notion of truth.

[42] The relevance of Bivalence to such issues has, of course, been argued by Michael Dummett, in a variety of places. See, for example, his "Realism", in *Truth and Other Enigmas*, pp. 145-65.

Note that one who holds such a view need not deny that "a=b ∨ ¬a=b" is valid: The rejection of Bivalence need not commit one to rejection of Excluded Middle, unless one's notion of (absolute) truth distributes over disjunction.

view that there are no vague objects is itself plausibly explained as the view that the 'boundaries' of any given object are perfectly determinate; because its boundaries are determinate, the identity of such an object--whether it is **b**, say--is also perfectly determinate. Every identity-statement is therefore either (definitely) true or (definitely) false, is of determinate truth-value: Thus, the view that there are no vague objects is committed to the claim that the principle of Bivalence holds for all *identity-statements*.[43]

The view that there are vague objects is simply the denial of the view just explained: Hence, it is committed only to the claim that not every identity-statement is of determinate truth-value. The conclusion may be reinforced by reflection upon the picture characteristic of the view that there are vague objects: The 'boundaries' of such objects are, so to speak, fuzzy. Because the boundaries of these objects are fuzzy, the identity of such objects may itself be fuzzy: Because the objects do not themselves have determinate boundaries, an identity-statement containing names of

---

Since truth, in this sort of case, is unlikely to do so, acceptance of Excluded Middle is consistent with rejection of Bivalence.

[43] One may hold this sort of view even if one maintains that Bivalence fails for such statements as "The table is red". One who holds this sort of view may accept a many-valued semantics: Identity-statements, however, will take on only the values True and False.

such objects need not have a determinate truth-value, need not be determinately either true or false.

It follows that the view that there are vague objects need be committed to no more than the invalidity of "$\Delta a=b$": For, given that any such view is incompatible with Bivalence, one can consistently hold that, though not every instance of "$\Delta a=b$" is true, no instance is false.

The difficulty is not, however, to motivate the claim that the view that there are vague objects is committed to the denial of the principle of Bivalence. The view that there are vague objects has usually been explained in terms of the claim that identity-statements might be neither true nor false, that is, might have some *truth-value* other than True or False. In general, however, to claim that statements may be neither true nor false, may have some intermediate truth-value, is to allow for the introduction of an operator '∇' interpreted as follows: A sentence "∇A" is true if, and only if, A has some truth-value other than True or False, and is false otherwise. This sort of view, which proceeds *via* a *many-valued* semantics, is thus committed to the view that statements of the form "$\nabla a=b$" may be true: For the view just is the view that identity-statements may be neither true nor false.

That view, we have seen, is probably not tenable. But it is not the only alternative to the view that every statement is either true or false. An additional alternative to

Evans's view that there are no vague objects is a view committed both to the denial of Bivalence and to the denial of the principle of Multi-valence, the principle that there are (usually finitely) many truth-values and that each statement has some one of these.[44] One who is attracted to the view that there are vague objects is naturally drawn to the view that some identity-statements may be neither true nor false. Such a view reflects the commitment to the rejection of Bivalence but retains a commitment to Multi-valence: That, or so I am now arguing, is the source of its defeat.

Any view which rejects not only Bivalence but also Multi-valence, which rejects the idea that every statement has some particular truth-value, constitutes a substantial departure from the view that every identity-statement is either true or false. To reject the claim that every statement has some one of however many possible truth-values is to reject the claim that what truth-value a statement has is *independent of our knowledge and of our capacities for knowledge*. The view is not that there is something other than True or False for sentences to be: It is rather that our model of truth and falsity, as objective properties of sentences, whose possession of various truth-values is epistemically unconstrained, fails to apply in this case.

---

[44] It is again usual to add the qualifier "determinately" here. I think that the term "Multi-valence" originated with Michael Dummett, but I am not sure.

The view that there are not two, but three (or more), truth-values can only seem comparatively familiar.

We have, at least, a model of such a more unusual view in the Intuitionistic philosophy of mathematics.[45] An Intuitionist rejects the principle of Bivalence and so denies that every statement is either true or false. Yet she also maintains that no statement can be neither true nor false. Her view is emphatically *not* that, objectively speaking, some statements are true (i.e., provable), others false (i.e., refutable), and yet others neither true nor false (i.e., neither provable nor refutable). Her view is that statements of mathematics do not merit the sort of objectivity which we naturally accord to them: We may speak only of what is provable, and we may speak of what is provable only in terms of what we can prove, of what we might be able to prove.

Whether the alternative to Evans's view which I have sketched is even ultimately explicable depends upon whether it is possible to formulate a semantic theory which would

---

[45] One might well want to say that Intuitionistic real numbers, or more generally Choice Sequences (which appear primarily in Intuitionistic Analysis), are vague objects. (Identity for natural numbers is decidable, so Bivalence holds here.)

Wittgenstein too hints at such a picture, when he writes: "And if you say that the infinite expansion must contain the pattern ϕ or not contain it, you are so to speak shewing us the picture of an unsurveryable series reaching into the distance. But what if the picture began to flicker in the far distance?"--*Remarks on the Foundations of Mathematics*, 3rd ed., tr. by G.E.M. Anscombe (Oxford: Blackwell, 1978), Part V, section 10.

accord with such a view. I have not even attempted to present such a semantic theory here, and I do not know how to do so. Nonetheless, the analogy with Intuitionism is meant as a promissory note: The formulation of an alternative to the view that the vagueness of identity-statements is a product of the vagueness of our language requires the development of a view which is similar to Intuitionism.[46]

Short of providing such a semantic theory, we can best understand the nature of this alternative view by considering by what sorts of arguments it might be motivated.

We may envisage, in the first instance, an argument which parallels arguments for Intuitionism and other forms of Anti-Realism.[47] Bivalence is to be abandoned, perhaps,

---

[46] There is a formal similarity between the views which is worth mentioning here. One who holds such a view will probably maintain the validity of "$\neg\Box\nabla a=b$". Suppose the usual definition of 'Δ' in terms of '$\Box$'. Then Evans's proof shows that "$\nabla a=b \rightarrow \neg a=b$". So, necessitating and distributing, we have (1) "$\Box\nabla a=b \rightarrow \Box\neg a=b$". By (T$\Box$), (2) "$\Box\nabla a=b \rightarrow \nabla a=b$"; and by the definition of 'Δ' and (C), (3) "$\Box\neg a=b \rightarrow \neg\nabla a=b$". So by (1) and (3), "$\Box\nabla a=b \rightarrow \Delta a=b$"; so, conjoining with (2), "$\Box\nabla a=b \rightarrow \nabla a=b \;\&\; \neg\nabla a=b$"; hence, by PC, "$\neg\Box\nabla a=b$".

This may well be compared to the Intuitionist's acceptance of the validity of "$\neg\neg(A \vee \neg A)$".

The similarities of which I speak should not be taken to imply that the alternative is committed to the abandonment of classical predicate logic.

It is also worth noting that a proof like Evans's can be given for any sentence "$\nabla\nabla\ldots\nabla a=b$". So "$\nabla\nabla\ldots\nabla a=b \rightarrow \neg a=b$" is valid and we shall thus also be able to show that "$\neg\Box\nabla\nabla\ldots\nabla a=b$" is valid, for any finite string of ∇'s.

[47] I refer, of course, to the well-known arguments due to Michael Dummett. See his "The Philosophical Basis of Intuitionistic Logic", in *Truth and Other Enigmas*. See also the helpful explication of these arguments in the "Introduction" to Crispin Wright's *Realism, Meaning, and Truth* (Oxford: Blackwell, 1986), pp. 13-29.

because the identity of objects of any sort is essentially connected to human practices: It is our language and related social institutions which provide criteria for the identity of objects, and those criteria do not in all cases decide (or provide for the decision of) all questions of identity. The objects whose identity is not decided by the criteria provided by the practice of speaking our language are the vague objects: Standard Anti-realist considerations might then lead one to hold that, not only are we sometimes unable to decide the identity of a vague object, whether say **a** is **b**; there is, moreover, nothing here which we do not know. The counterargument would, as in the case of Intuitionism, presumably be that the apparently undecided cases are *really* decided, that every identity-statement *must* be either true or false. But there is at least precedent for resistance to this sort of claim.

However, it need not be the general sorts of considerations which motivate Anti-realism which motivate the proposed alternative to Evans's view. It may be essential to the view that there are vague objects that criteria for the identity of such objects be connected to social practices in some *more specific* way than, as required by standard Anti-realist arguments, meaning is in general (supposed to be) connected to social practices. Paradigmatically, vague objects are artifacts, objects of human creation, in quite an ordinary sense.[46] This sort of fact may well impress one who is attracted to the view that, say, there really are such things as ships and clocks, over and above mereological fusions or collections of clock- and ship-parts (appropriately arranged). One might suggest, for example, that it is only because people create clocks, or use them for certain purposes, that we need distinguish between the clock and the parts which make it up, between the history of the clock and the history of the parts.[47]

Indeed, one *might* find oneself attracted to the idea that, were there no people, no minds, or were there no practice of telling time or of sailing, there would be no such things as clocks or ships (though there might be collections of ship-parts, arranged just as the parts of a ship are arranged). If so, one might find it an attractive idea that the truth about the identities of such objects just *can not* transcend the criteria for their identity which are contained in our social practices (even if that sort of principle does not generally attract one): There is, as it were, nothing more to the identity of such objects than what

---

[46] Mountains would likely fall outside the scope of such a view. Mountains are vague, in the sense that it is undetermined what a mountain's precise boundary is. But one might think that that is quite inessential, that we are talking *imprecisely* about the physical stuff which not only makes up (constitutes) the mountain but which really *is* the mountain. Compare David Wiggins on constitution, in his *Sameness and Substance* (Oxford: Blackwell, 1980), pp. 43-4 and elsewhere.

[47] Compare Wiggins, pp. 90-9, 124-6.

is provided by our language and other social institutions, for such things are what they are only in virtue of certain of our social institutions.

8. Closing

That, of course, is but an outline of the kind of considerations which might attract one to the view that there are vague objects. For our purposes here, it is not important that they should be convincing: I make no claim to be convinced myself. The important point is that nothing about the fundamental motivation of the view commits one to the existence of truths of the form "$\nabla a=b$". To what it commits one is, again, the rejection of Bivalence and Multivalence, and so, in my opinion, the rejection of Realism. The view is that vague objects are objects of human creation in a not so ordinary sense: Vague objects are mind-dependent.[80]

It is for this reason that I have consistently spoken of the view that there are vague objects, rather than of the view that vagueness is 'real' or of the 'reality of vague-

---

[80] Dummett has suggested a similar-sounding view (though I do not claim anything by way of interpretation): "Realism about vagueness is anti-realism about the world". See his "Reply to Wright", in *Essays on the Philosophy of Michael Dummett*, ed. B. Taylor (Dordrecht: Maritnus Nijhoff, 1987), p. 229.

Vagueness has never fit very well into Dummett's analysis of metaphysical issues. The above remark is, in fact, taken from a discussion of just this sort of problem. This paper does, I think, help to bring the problem of vagueness into the fold.

ness'.[51] The view that there are vague objects is incompatible with a Realist treatment of statements containing names of them: One who is committed to Realism about the material world and all that it contains is therefore committed to denying that there are vague objects. *That* is what Evans's argument, it seems to me, really shows: That there are no vague objects which are mind-independent; that Realism about the physical world is incompatible with the view that there are vague objects.[52]

---

[51] As does Pelletier, *passim*. The phrase also fails to distinguish 'Realism' about vague objects from 'Realism' about vague properties.

[52] I should like to thank George Boolos, Jim Higginbotham, Paul Horwich, and Bob Stalnaker for their encouragement and criticism. I should also like to thank Prof. Michael Dummett for related discussions. Reflection on his now revised William James Lectures, *The Logical Basis of Metaphysics*, inspired many of the underlying ideas of this paper.

Most of my discussion of Francis Jeffry Pelletier's paper has been quite critical. I should like to emphasize my debt to it.

APPENDIX

I mentioned earlier that there are logics well adapted to the purposes of the view that there are vague objects. I shall sketch the completeness proof here for these logics: The language contains sentence-letters, terms, n-place predicates, the usual sentential operators, and identity, along with $\Box$ and Δ. The extension to predicate logics poses no difficulty of principle.

We formalize these logics within a generalized natural deduction system. We write sequents not just as Γ:A, but as $(\Gamma;\Delta)$:A. Here Γ is a set of premises; Δ, a set of *hypotheses*. Intuitively, a premise is something we assume to be true, from which we deduce the conclusion; an hypothesis, on the other hand, is something we assume for the sake of argument. Rules like &-introduction are largely unchanged:

$$\frac{(\Gamma;\Delta):A \qquad (\Gamma';\Delta'):B}{(\Gamma,\Gamma';\Delta,\Delta'):A \,\&\, B}$$

Rules which discharge premisses are now written as follows:

$$\frac{(\Gamma;\Delta,A):B}{(\Gamma;\Delta):A\rightarrow B} \qquad\qquad \frac{(\Gamma;\Delta,A):\bot}{(\Gamma;\Delta):\neg A}$$

The sentence to be discharged must be an *hypothesis*, not a premise. We may then have other rules, which will be the rules of deduction, such as the following, for '$\Box$' read as 'Definitely':

$$\frac{(\Gamma;\Delta):A}{(\Gamma,\Delta;\emptyset):\Box A}$$

All hypotheses on which 'A' depends are, thus, converted to premises by this rule. Hence, the rule can not possibly occur within a subordinate deduction, since subordinate deductions discharge only hypotheses.

Let '□' be subject to the laws of any complete, compact, normal modal logic, which we may call Ω, suitably formalized as a natural deduction system of the usual sort. We define a logic $V\Omega$ as follows. Formalize Ω itself as a natural deduction system of this new sort, and add to the logic the rule of deduction mentioned above, which I shall call V□.

Basic sequents are of the form $(\emptyset;A):A$ or $(A;\emptyset):A$. (We allow also for a rule, analogous to thinning rules, which moves sentences from the set of hypotheses to that of premises.) Every proof in Ω may thus be converted to a proof in $V\Omega$ by replacing basic sequents, in Ω, by basic sequents, of the former sort in $V\Omega$, and by replacing applications of the rules in Ω by their related rules in $V\Omega$. (Since no rule of Ω will introduce sentences into the set of premises, the $V\Omega$-rules which discharge premises remain applicable.) Note, importantly, that if $(\Gamma,\Delta):A$ is provable in Ω, then, not only is $(\emptyset;\Gamma,\Delta):A$ provable in $V\Omega$, but $(\Gamma;\Delta):A$ is provable in $V\Omega$. For no sentence $p\in\Gamma$ *is* discharged; hence, no such sentence figures as hypothesis of a subordinate deduction.

Take, as models of the new language VΛ, standard models of Λ (with respect to which Λ is complete and, which, therefore are strictly characteristic for Λ), but define truth as truth in *all* possible worlds; VΛ is complete with respect to these models.

The proof is fairly straightforward. First, just the same formulas are provable in the two systems: For, if V□ is ever applied to a sequent with non-empty antecedent, the result is a sequent with non-empty premise; and no premise can be discharged. If V□ is applied to a sequent with empty antecedent, it can be replaced by an application of necessitation. It follows that every VΛ-provable formula, being Λ-provable, is valid in all models of Λ and, given that it follows that it is true at each world in every such model, it is true in all models of VΛ. On the other hand, if a sentence A is valid in all models of VΛ, it is, plainly, valid in all models of Λ; hence, it is provable in Λ (since Λ is complete) and so in VΛ.

The proof that these models are strictly characteristic for VΛ is a bit more complicated. (The models are strictly characteristic if, and only if, a sequent is valid when and only when provable.) We say that a sequent $(\Gamma;\emptyset)$:A is valid if, whenever each sentence $P\in\Gamma$ is true at each world in a model, A is true at each world in that model. (Note that I shall discuss, shortly, the more general case of provable sequents of the form $(\Gamma;\Delta)$:A.)

Suppose that $(\Gamma;\emptyset):A$ is provable in $V\mathfrak{N}$. Consider an application of the rule $V\Box$, above which there is no other application of it:

$$\frac{(\Gamma';\Delta'):A'}{(\Gamma',\Delta';\emptyset):\Box A'}$$

Thus, $\Gamma',\Delta':A'$ must provable in $\mathfrak{N}$. Since $\mathfrak{N}$ is normal, $\Box\Gamma',\Box\Delta':\Box A'$ is also provable in $\mathfrak{N}$ (here, if $\Gamma=\{G,H,\ldots\}$, $\Box\Gamma=\{\Box G,\Box H,\ldots\}$). We now replace the proof of $(\Gamma',\Delta';\emptyset):A$, in $V\mathfrak{N}$, by this proof of $\Gamma',\Delta':A$ in $\mathfrak{N}$. By induction, since proofs are finite (and well-founded), if $(\Gamma;\emptyset):A$ is provable in $V\mathfrak{N}$, there is some **n** such that $\Box^n\Gamma:A$ is provable in $\mathfrak{N}$.

Suppose that each sentence in Γ is true in some arbitrary model of $V\mathfrak{N}$. If p is such a sentence, then, for any **k**, $\Box^k p$ is true at each world, since p is true at each world in the model. Hence, if we now construe the model as a model for $\mathfrak{N}$ (choosing some 'actual' world *a*), for each $p\in\Gamma$, $\Box^n p$ is true at *a*; hence, A must be true at *a*. Thus, A must be true at all worlds (since the 'actual' world was arbitrary); hence, A is true in the (original) $V\mathfrak{N}$-model. And so the class of such models is faithful to $V\mathfrak{N}$.

Conversely, suppose that $(\Gamma;\emptyset):A$ is valid in $V\mathfrak{N}$. Then, in any model in which each sentence in Γ is true at every world, A is true at every world. It follows that $\Gamma,\Box\Gamma,\ldots,\Box^n\Gamma,\ldots:A$ is valid in $\mathfrak{N}$. (Proof: Let $\mathfrak{M}$ be a model of $\mathfrak{N}$, with *a* the actual world. Suppose that the antecedent is satisfied at *a*. Consider the sub-model $\mathfrak{M}|a$ of $\mathfrak{M}$, which consists of worlds w for which there is a sequence $a=b_0 R b_1 R \ldots R b_n = w$ connecting w to a. Suppose that, for each n, $\Box^n\Gamma$ is true at a. Then since each world $w \in \mathfrak{M}|a$ is some finite distance n from a, it follows that each sentence in Γ is true at each world in $\mathfrak{M}|a$. Considering $\mathfrak{M}|a$ as a model of $V\Omega$, then, we have that A is true at each world, since each sentence in Γ is true at every world and $(\Gamma;\emptyset)$:A is valid in $V\Omega$. Hence A is true at a.)

But Ω is compact. So there is some **k** such that $\Gamma,\ldots,\Box^k\Gamma$:A is valid in Ω. Hence, it is provable in Ω, since standard models for Ω are strictly characteristic for it. Hence, $(\Gamma,\ldots,\Box^k\Gamma;\emptyset)$:A is provable in $V\Omega$, and it is easy to construct a proof, in $V\Omega$, of $(\Gamma;\emptyset)$:A. We need only take our proof in Ω, and append it to as many $V\Box$ steps as are require to take us, for each sentence p in Γ, from p to $\Box^k$p.

So the class of models is strictly characteristic for $V\Omega$.

We may now define a more general notion of validity for sequents of the form $(\Gamma;\Delta)$:A. We say that a sequent of this form is valid if, whenever each sentence $p \in \Gamma$ is true at each world in a model, A is true at each world at which each sentence $q \in \Delta$ is true.

It is now easy to see that our original class of models, given this definition of validity, is strictly characteristic for $V\Omega$. Since no sentence $p \in \Delta$ can be an hypothesis of a sequent which is the basis of an application of $V\Box$, not only may we conclude that, for some n, $\Box^n\Gamma, \Box^n\Delta : A$ is provable in Ω, we may conclude that $\Box^n\Gamma, \Delta : A$ is provable in Ω. (For the induction step of our original proof operated only on applications of $V\Box$, and so only on sentences which become premises.) Since Ω is compact, there is some finite $\Delta' \subset \Delta$ such that $\Box^n\Gamma, \Delta' : A$ and, hence, $\Box^n\Gamma : (\&\Delta') \rightarrow A$ is provable in Ω (where "$\&\Delta'$" means the conjunction of all sentences in Δ'), by the deduction theorem for Ω.

Now suppose that $\mathfrak{M}$ is a model of $V\Omega$ in which each sentence in Γ is true at each world. As earlier, $\Box^n p$ is true at each world, for each $p \in \Gamma$. Hence, we have that $\&\Delta' \rightarrow A$ is true at each world (since $\Box^n\Gamma : \&\Delta' \rightarrow A$ is provable in Ω and, again, the actual world is arbitrary). And so, if [illegible] is a world at which each $q \in \Delta$ is true, since $\Delta' \subset \Delta$, $\&\Delta'$ is true at w, so A is true at w. Hence $(\Gamma;\Delta) : A$ is valid, and the class of models is faithful to $V\Omega$.

Conversely, suppose $(\Gamma;\Delta) : A$ is valid. As earlier, it follows that $\Gamma, \ldots, \Box^n\Gamma, \ldots, \Delta : A$ is valid in Ω. By compactness, again, we have that, for some n and for some finite $\Delta' \subset \Delta$, $\Gamma, \ldots, \Box^n\Gamma : \&\Delta' \rightarrow A$ is valid and hence provable in Ω. Hence, $(\Gamma;\emptyset) : \&\Delta' \rightarrow A$ is provable in $V\Omega$; it is therefore easy to construct a proof of $(\Gamma;\Delta') : A$ in $V\Omega$. So $(\Gamma;\Delta) : A$ is provable in $V\Omega$, and the class of models is strictly characteristic for $V\Omega$.

Paper II

Whether Structure May Be Misleading:

Wright on Reductionism

0. Opening

Ontological questions arise, perhaps, nowhere more frequently, nor more intractably, than where we are concerned with abstract objects: Directions, numbers, letter- and word-types, and the like. On the one hand, the fact that expressions which purport to refer to abstract objects function in much the same way as names of concrete objects might incline one toward the view that there are such objects. On the other hand, philosophers have found abstract objects to be, among other things, epistemologically problematic. Such objects apparently do not have causal powers and can, therefore, neither be perceived nor be known by their effects: It thus becomes an important question how, if such objects do exist, we can know anything about them. Moreover, if abstract objects have no causal powers, it is not clear why science, say, should have any use for them: It is at least plausible that no causal explanation must make reference to entities which have no causal powers; and, if not, science need not, and therefore ought not, recognize the existence of such entities.[1]

Those who have been impressed by such epistemological considerations, if they have not rejected the existence of abstract objects entirely, have at least thought an Ontology

[1] For a good expression of this line of thought, see Hartry Field's *Science without Numbers* (Oxford: Blackwell, 1980).

virtuous in so far as it is parsimonious, i.e., is committed to the existence of as few different *types* of objects--and in particular, abstract objects--as possible. These epistemological concerns, however, have rarely been taken to constitute a conclusive argument for the view that there are no abstract objects. Rather, philosophers have taken themselves to need to show that we can, so to speak, do without names for objects of one sort or another: The task which confronts such a view is to show just how we can--or, indeed, if it is correct, how we *do*--do without reference to, say, numbers.

Russell's Theory of Descriptions provides, and was surely taken as providing, a model for such demonstrations. Russell did not explicitly define the word "the", nor the phrase "the King of France"; rather, he showed us how systematically to replace sentences in which they occur by sentences in which the quantifiers 'all' and 'some' occur. Similarly, one might think, to show that we can do without reference to numbers, one need only show how systematically to translate sentences containing expressions which purport to refer to numbers into sentences which contain no such expressions; into sentences which, instead, contain only expressions which refer to, or quantify over, epistemologically less problematic entities.

Such a translation is, classically, to be accomplished by means of a Contextual Definition. Crispin Wright takes as his standard example of a Contextual Definition a version of Frege's definition of names of directions. The definition is motivated by the observation that the direction of **a** is the same as the direction of **b** if, and only if, **a** is parallel to **b**: What appears to be a relation of identity between directions is, in some sense, just the relation of parallelism between lines. A sentence which says something about a 'direction' is really just a sentence which says something about a given line (and, by implication, any line parallel to it). Hence, we have the following:[2]

> dir **a** = dir **b** if, and only if, **a** $\parallel$ **b**
> F(dir **a**) if, and only if, f**a**

Here, 'dir a' is to be read 'the direction of **a**'; 'fξ' is a predicate of lines which is a *congruence* with respect to parallelism[3] and which is suitably related to 'Fξ' (so that the truth-conditions come out right).

The important feature of the relation, parallelism, is that it is an equivalence relation: The fact that it is an equivalence relation guarantees that "=", as it occurs in sentences of the form "dir **a** = dir **b**", has the formal pro-

---

[2] Crispin Wright, *Frege's Conception of Numbers as Objects* (Aberdeen: Aberdeen University Press, 1983), pp. 29-30.

[3] A predicate 'Fξ' is a *congruence* with respect to a relation '$\eta R \xi$' if, and only if, $(\forall x)(\forall y)(Fx \,\&\, xRy \rightarrow Fy)$. A relation '$\eta R \xi$' is an equivalence relation if it is reflexive--if $\forall x(xRx)$--symmetric--$\forall x \forall y(xRy \equiv yRx)$--and transitive--$\forall x \forall y \forall z(xRy \,\&\, yRz \rightarrow xRz)$. If '$\eta R \xi$' is an equivalence relation, then it follows that a predicate 'Fξ' is a congruence with respect to it if, and only if, $\forall x \forall y(xRy \rightarrow [Fx \equiv Fy])$.

perties of identity. This, together with the fact that each predicate '$f\xi$' is a congruence with respect to parallelism, guarantees that a form of Leibniz's Law is valid, namely:

$$\forall x \forall y[\text{dir } x = \text{dir } y \mathbin{\&} F(\text{dir } x) \rightarrow F(\text{dir } y)]$$

Note that "$F\xi$" here is required to be a predicate defined in accordance with the Contextual Definition given above.

There is nothing wrong with this sort of Contextual Definition. Wright properly emphasizes the need to define, not just *identity*, but *predicates* of directions as well. The definition does, however, mask an important fact: Namely, that a Contextual Definition may, in principle, be given in terms of *any* equivalence relation, that there need not be any simple (or primitive) predicates (in the language in which the Definition is given) which are congruences with respect to that relation. That is no obstacle to the production of *complex* (or defined) predicates which are congruences with respect to it.[4] To give the form of Contextual Definition, in the more general case, we may assume that the predicates, on the right-hand side, in terms of which the predicates on the left-hand side are defined, are of the form "$(\forall x)(xR\xi \rightarrow fx)$", where "$\eta R\xi$" is the equivalence relation in terms of which the definition is given.

---

[4] The distinction between simple and complex predicates, which I am using here, is intuitively clear but difficult to explain. For more on the distinction, see Michael Dummett, *Frege: Philosophy of Language*, 2nd ed. (London: Duckworth, 1980), pp. 28-33.

In the case of directions, then, the more general form of the Definition is:

$F(\text{dir } \mathbf{a}) \equiv_{df} (\forall x)(x \parallel \mathbf{a} \rightarrow fx)$

And, in the more general case:

$F(\text{fnc } \mathbf{a}) \equiv_{df} (\forall x)(xR\mathbf{a} \rightarrow fx)$

Here, "fnc ξ" is a functional expression, like "dir ξ"; "$\eta R \xi$" is an equivalence relation such that fnc **a** = fnc **b** if, and only if, **a**R**b**; "fξ" is, again, a predicate suitably related to "F_ξ" which, now, *need not* itself be a congruence with respect to '$\eta R \xi$'.

For the purposes of this paper, I am going to understand the term "Reductionism" as a common name of such views as hold that, *if* a Contextual Definition of this sort can successfully be given, then the names which have been eliminated from the class of sentences in question do not in fact refer.

It is worth re-emphasizing that Reductionism, as I understand it, is not committed to the view that the *best argument* against the existence of objects whose names are capable of elimination is just that those names are eliminable. The best argument may well be an epistemological one, that, even if there were such objects, we could know nothing of them. Nonetheless, few Reductionists would want to suggest that we abandon Arithmetic, because there are no numbers, or Linguistics, because there are no types: Hence the importance of showing how we can make sense of Arithmetic without supposing ourselves to refer to Numbers; of Linguistics, without supposing that we refer to types. This sort of Reductionism is thus here being understood as committed to the view that, *if* we can do without reference to objects of a given sort, then there are no such objects, even though that does not tell the whole story.

Wright has argued that Reductionism is "fundamentally misguided".[5] He has argued, in particular, that a version of the Context Principle, originally due to Frege,[6] provides sufficient material from which to construct a justification of the view that there are abstract objects. In this paper, I am going to argue that, as I understand Wright's view, it is mistaken. The arguments by means of which he defends, say, the existence of directions, despite the Contextual Definition of names of them, are powerful enough to defend the existence of objects, names of which may be introduced by Contextual Definition, which just do not exist.

---

[5] Both Wright and I attempt to avoid the thorny question when a Reductionist should take a Contextual Definition successfully to eliminate names. Such blatantly circular attempts as

type(a) = type(b) iff a is a word of the same type as b

should not be. If I am correct, the best arguments for the non-eliminability of names of abstract objects present a problem *even for this sort of circular 'definition'*. That constitutes a strong sense in which worries about what constitutes a 'successful' definition are irrelevant to this debate. See "Trans-sortal Identification", in this thesis.

[6] Gottlob Frege, *The Foundations of Arithmetic*, tr. by J. Austin (Evanston IL: Northwestern University Press, 1980), p.x.

1. Understanding Contextual Definitions

We are supposing, for the sake of argument, that names of directions may be eliminated and, indeed, could have been introduced by means of the Contextual Definition discussed in the last section.

The least restrictive possible view would hold that, if a speaker understands such a Contextual Definition, then she is able to refer to directions. But, plainly, it is not sufficient for a speaker to be able to refer to abstract objects that she be introduced, by means of a Contextual Definition, to terms which purport to refer to such objects; not, at least, if all that she is able to do is to use the Definition as a scheme of translation. A speaker might well come to grasp the scheme of translation embodied in the Contextual Definition for directions without knowing that parallelism is an equivalence relation; if she did not know this, she would not know that "=", as it occurs in statements of the form "dir **a** = dir **b**", has the formal properties of identity. To adapt a famous phrase, our speaker can not refer to directions if she does not know what determines the identity of a given direction; and she can not know that if she does not know that "=", in this use, is (or at least has the formal properties of) the identity-sign.

For similar reasons, our speaker must know that the structure of the Definition guarantees that, if "=" is so understood, Leibniz's Law holds. She must, that is, know that the predicates, in terms of which these predicates "$F\xi$" are defined, are congruences with respect to parallelism.

Wright emphasizes that his view requires that the terms introduced by a given definition be able sensibly to flank the identity-sign.[7] Presumably, it is not essential that the sign for identity be the same in all parts of the language, that there be such a thing as "the" identity-sign. What Wright has in mind, I think, is rather that the sign which occurs, as it were, where one would expect an identity-sign to occur, has the formal properties of identity; that, so far as the definition is concerned, an identity-sign *might just as well* occur there. We have now seen that this amounts to a requirement that the relation in terms of which the Defintion is given be an equivalence relation and that the predicates thus introduced be congruences with respect to it.

Wright speaks, in most cases, of what is required if there are to *be* certain abstract objects. I shall here content myself with discussion of the related question what is required if a given speaker is to be able to refer to such objects. If one understands a given Contextual Definition and knows that the relevant relation is an equivalence relation, and if she knows that the defined predicates are congruences with respect to it; then I shall say that she has a *Theoretical Understanding* of that Definition. I hereby

---

[7]Wright, pp. 149-52.

attribute to Wright the view (or a view which entails) that, if a speaker has a Theoretical Understanding of a Contextual Definition, then she understands the terms so defined to be *able* to refer to (ordinarily abstract) objects. Or again: A speaker who has a Theoretical Understanding of a Contextual Definition is, herself, *able*, so far as her conceptual resources are concerned, to refer to (ordinarily abstract) objects by means of such names.[6]

Whether a given term does refer, or whether our speaker is *in fact* able to refer, to such an object is a question outside the domain of philosophy *per se*. One way of understanding what is, additionally, required can be extracted from an idea of Wright's: The question is wh ther there are any *true* sentences which contain the term.[7] *That* question is

---

[6] Wright does not explain his view in anything like these terms. I should mention that he does make certain remarks which conflict with this interpretation of his views: Namely, that it is unclear to him why a philosopher might seek to defend "a general policy of contextual definition"--p. 9. As the view just attributed to him entails that just such a policy would be quite reasonable, I do not see how to reconcile his views with this remark. On the other hand, it is not entirely clear that Wright means to reject this view at all.

In any event, there is a great deal of tension between this remark and other aspects of his view, already, as we shall see.

[7] We probably need to say something like "true *simple* sentences". 'Simple' is just the best way to exclude sentences of the form "Such-and-such does not exist". (Wright does not emphasize this point and I am not sure he would agree.)

My use of the notion of a simple sentence, here, should *not* be taken as having any connection with my use of the notion of a simple *predicate*, earlier, nor with the related notion of a *variant* predicate, which I explain shortly.

to be answered by the science (in a very loose sense) whose province such questions are. In the case of, say, the Contextual Definition of directions, it is geometry which will answer the question; in other cases, it will be other fields.

This distinction only becomes important in other contexts; but it is worth mentioning it here. As it is not of great import to our discussion of Wright's view,[10] however, I shall ignore this, and an earlier, complication, since most terms *able* to refer *will* refer to abstract objects. I shall thus speak of Wright's view as the view that a speaker who has a Theoretical Understanding of a Contextual Definition understands the terms so defined *to refer* to abstract objects; or I shall say that such a speaker is able to refer to such objects.

---

[10] Indeed, it is not really clear that Wright would accept the distinction. Wright seems to think that there is just no additional question, once the Contextual Definition has been given, whether the objects in question exist: See *FCNO*, pp. 146-53. Admittedly, no general principle is laid down: But how else is one to understand "[T]here seems to be a kind of incoherence in the idea that a line might lack a direction, a geometric figure a shape..."? Especially is this so when it has just been said that a comparable situation exists "*whenever* an equivalence relation between things of one sort is taken as necessary and sufficient for identity of things of another sort" (p. 148, my emphasis).

2. Initial Difficulties with Wright's View

A Reductionist may offer, it seems to me, three sorts of objections to Wright's view. The first, classical objection is simply that we do not need to ascribe reference to the 'names' which have been introduced by Contextual Definition. This objection, as Wright emphasizes,[11] itself prompts the question just when we *do* need to ascribe reference to names. To answer this question, we must ask what theoretical *work* the notion of reference does for us. An object, the referent of a name, is primarily that of which predicates are true or false: We require the notion of reference to an object, primarily, because a simple sentence "Fa" must be explained as being true if, and only if, the object to which 'a' refers satisfies the predicate "Fξ".[12]

That said, we may re-state the objection: It is just not sufficient, for us to be justified in ascribing reference to a singular term, that that expression function, syntactically, as a singular term. What is required, instead, is that--in a sense as yet unexplained--that expression function, semantically, as a *constituent* of (at least some) sentences in which it occurs. As a first approximation, we may say that a speaker does not understand a term to refer to an object unless she understands

[11] Wright, *FCNO*, section v, esp. pp. 31-6.

[12] For a discussion of this point, as contrasting with Quine's (one time?) view of the matter, see Chapter 15 of Dummett's *Frege*.

that term, not as a kind of code for some hidden quantificational structure, but as a semantically unitary expression.[13]

One who is attracted by this sort of objection need not deny that the notion of an object is primarily to be explained in terms of the notion of a proper name. She need not, that is, deny the Fregean thesis that the notion of an object is in part a theoretical notion, one which is in part explained by its role in semantic theory. Indeed, the objection would seem largely to *rest* upon a similar claim: The point of having a notion of reference for sub-sentential constituents of sentences--for words in general and for names in particular--in a Semantic Theory, is as part of an explanation how the constituents of a sentence contribute to the determination of its truth-value. If an expression functions as a singular term, and if it functions as a constituent of certain sentences in which it occurs, then the Theory can do no other than assign it, presumably, an object as its reference (unless, again, it is merely empty). If, on the other hand, what is syntactically a singular term does *not* function as a constituent of sentences in which it occurs, then it is wholly otiose to assign it a reference.[14]

---

[13] Compare Wright, *FCNO*, pp. 67-9.

[14] And here we may want to see, for example, Dummett's discussion of abstract nouns in just this light: See ***Frege***, pp. 72-80.

The second sort of objection which a Reductionist might offer is the following: Granted that a speaker does not have a conception of (say) directions as objects[15] if she does not know that parallelism is an equivalence relation, how can her knowledge that it *is* an equivalence relation give her the ability to refer to directions? It is true that, if she has such knowledge, she will be able, as it were, to use the symbol "=", in such contexts as "dir a = dir b", as if it were the identity-sign: But the question is, how can the mere knowledge that "=", in such contexts, has the formal properties of identity justify an ascription of an ability to refer to directions? granted that, without such knowledge, one lacks this ability?

As we said earlier, it is presumably of no great importance whether the same *sign* (say, "=") is used in this case and in other cases. So suppose that there is, in a given language, no single sign of identity, which occurs both in more usual contexts and in this newer one. Suppose, further, that our hypothetical speaker knows that parallelism is an equivalence relation and that

$$(\mathfrak{X}) \quad \forall x \forall y[x \parallel y \; \& \; F(\text{dir } x) \rightarrow F(\text{dir } y)]$$

is valid. Since the mere replacement of "x ∥ y" by "dir x = dir y" will not give our speaker the ability to refer to directions if she did not already have it, it is the knowl-

[15] I borrow this turn of phrase from the title to Wright's book.

edge that $\mathfrak{T}$ is valid which, in addition to the knowledge that parallelism is an equivalence relation, is required if she is to have a Theoretical[14] Understanding of the Definition, if she is to be able to refer to directions. But how can *this* sort of knowledge serve to distinguish those who can from those who can not refer to directions?

The third objection is really a combination of the first two. Sentences of the form '$\forall x(x \parallel \mathbf{a} \rightarrow fx)$' may well have been in common use before sentences of the form 'F(dir **a**)' were introduced. Suppose so. If we were to learn that parallelism is an equivalence relation, we might well come to know that, for any **a**, **b**, and '$f\xi$',

$$(\mathfrak{T}')\ \mathbf{a} \parallel \mathbf{b} \rightarrow \{(\forall x)(x \parallel \mathbf{a} \rightarrow fx) \leftrightarrow (\forall x)(x \parallel \mathbf{b} \rightarrow fx)\}\ .$$

We may well notice that $\mathfrak{T}'$ bears a formal resemblance to Leibniz's Law: We may then start re-writing '**a** ‖ **b**' as 'dir **a** = dir **b**'; we may re-write the constituents of the biconditional as "F(dir **a**)" and "F(dir **b**)". We may, indeed, record this scheme of abbreviation in a Contextual Definition. But all we are doing, at this late stage, is merely *re-writing* those sentences: What is important must be the knowledge which we have, which makes it *possible* to give the Contextual Definition, not the Definition itself.

---

[14] Hence the name "Theoretical": The knowledge which is required is capable of being explicitly formulated and could, indeed, be had even in the absence of the Definition. Similar remarks will apply to the predicate "$F\xi$" and its replacement of the original complex predicate.

It should be emphasized here that the problem is not that Wright (or anyone else) is committed to the view that an understanding of a *Contextual Definition* is the crucial ingredient in one's conception of directions (say) as objects. That would be absurd. What is important, according to Wright, is whether one understands that sentences of the form "$(\forall x)(x \parallel a \rightarrow fx)$" constitute a *kind* of sentence: Whether one knows that it is sentences of this *kind* to which one's knowledge concerning the 'identity of directions'--or, the truth-values of statements of the form "dir a = dir b"--is relevant. What the Contextual Definition does is to record that these sentences are of a kind and to introduce a special notation for them. What is problematic, from the point of view of a Reductionist, is, again, the claim that *knowledge* of a certain sort is alone sufficient to give one the ability to refer to directions. (That knowledge, again, is the knowledge that parallelism is an equivalence relation and the knowledge that Leibniz's Law--in whatever form, that of $\mathfrak{X}$ or $\mathfrak{X}'$ or some other form--is valid.)

The problem may, perhaps, be illustrated as follows: Let us introduce an equivalence relation by arbitrarily pairing all individuals who were alive on 21 March 1990. The relation 'ηPℓ' is defined as the smallest equivalence relation which holds between individuals in a pair. We may now introduce, as names of what I shall call by the common name "poursons", terms of the form 'pour **a**'. We do so by means of the Contextual Definition:

$$F(\text{pour } \mathbf{a}) =_{df} (\forall x)(xP\mathbf{a} \rightarrow fx)$$

Since '$\eta P\xi$' is defined as an equivalence relation, and since I know that, if **a**P**b**, then F(pour **a**) if, and only if, F(pour **b**), I have a Theoretical Understanding of this Definition. But do I now have the ability to refer to poursons?[17]

Wright is fully aware that his views commit him to a rich Ontology, one which admits all sorts of strange and wonderful objects. In itself, that is no objection. But I for one am at least inclined to think that there are no such things as poursons. And, if that were not trouble enough, we could have chosen *any* equivalence relation to construct this example: At the very least, there are not distinct sorts of objects corresponding to every (extensionally distinct) equivalence relation. Surely *something* must be wrong.[18]

---

[17] What sort of thing is a pourson? What can be said about them? Suppose that I have been paired with Dan Quayle. Then I know that the statement "Red-headed (pour(rh))" has been defined as equivalent to "$(\forall x)(xP(rh) \rightarrow$ red-headed(x))"; moreover, I know the latter to be false, since Dan Quayle does not have red hair. Hence, I know that "Red-headed(pour(rh))" is false.

One can now amuse oneself to no end with such sentences.

[18] One might wish to reply that poursons are just equivalence classes. I am not going to discuss this sort of move here. First, a discussion of it belongs with a more general discussion of the Julius Caesar problem, since the identification of, say, directions and classes is problematic, if it is, in the same way that the identification of numbers with persons would be. Secondly, I do not want to make any assumptions about the existence of sets, which are, of course, themselves abstract objects (if they are anything).

Now, one may well be *further* inclined to think that the *reason* that there are no such things as poursons is because we can eliminate names of poursons, by Contextual Definition. One of the great attractions of Reductionism is that it provides a safe haven from the ontologically explosive consequences of a view like Wright's. But whether it is the only serious alternative to Wright's view is another question.

3. That a Theoretical Understanding Might Suffice: Theoretical Understanding and the Context Principle

What good, theoretical reason can there be for ascribing Reference to names introduced by Contextual Definition? Why is it not sufficient to state the truth-conditions of such sentences in terms of the Contextual Definition itself? We shall be considering, in this section, a strategy for answering this question, on the assumption that speakers have a Theoretical Understanding of the Definition.

Before we do so, it is worth remarking that we shall not be considering a strategy which may seem promising:

The difference is one of emphasis only, in any event. My view is that, if we refer to poursons at all, then the objects to which we refer are *sui generis*, distinct from all other sorts of objects to which we can otherwise refer. We may not refer to poursons, but refer to objects of some other kind by means of terms of the form "pour(a)": But the question is whether we refer, by means of such terms, to objects which are *sui generis*; and the claim that we do not is just the claim that, in the sense in which poursons are, if they exist, *sui generis*, there are no such objects.

Namely, to argue that the knowledge which is required, if a speaker is to have a Theoretical Understanding of a Contextual Definition, is *precisely* the knowledge required if she is to grasp a criterion of identity for the objects whose names are introduced by the Definition. Such a move would answer one of the main, underlying worries which troubles my sort of Reductionist: For one would then be maintaining that the *knowledge* whose possession is distinctive of a Theoretical Understanding is *linguistic* knowledge; and, if so, we can begin to understand how possession of such knowledge can determine whether one can refer to objects of a given sort.

However promising the prospects of this strategy, though, it ultimately fails, for reasons I can not discuss in detail here. Later, we shall see some of the reasons it fails: My objections to it concern precisely the Theoretical, or *explicit*, character of the knowledge which it requires of speakers.[19]

It might be said that we have so far wholly and wrongly ignored the most important question: What is the *effect* of the possession of the knowledge constitutive of a Theoretical Understanding of a Contextual Definition? On the one hand, of course, our speaker simply grasps the scheme of translation embodied in the Definition. If *that* were all she understood, she would not be able to refer to the 'new'

---

[19] For further discussion of criteria of identity, in this connection, see "Trans-sortal Identification".

objects. But whatever else she may or may not know, however, she does know that the terms 'dir **a**', 'dir **b**', and so on, and the predicates "$F\xi$", "$G\xi$", and so on, make some regular contribution to the truth-conditions of sentences in which they occur. She knows, for example, that the predicate "$F\xi$" is always to be translated, say, as "$(\forall x)(x \parallel \xi \rightarrow fx)$"; and she knows that, whenever 'dir **a**' occurs as argument of a predicate, '**a**' will occur as argument of the translation of the predicate.

Her knowledge that 'copies' is an equivalence relation and that predicates of the form "$(\forall x)(x \parallel \xi \rightarrow fx)$" are congruences with respect to it has yet greater effect. Our speaker knows, for example, that if 'F(dir **a**)' is true and 'dir **a** = dir **b**' is true, then 'F(dir **b**)' is true. She knows, that is, that, if **a** is parallel to **b**, then the question whether every line parallel to **a** satisfies '$f\xi$' is *just the same question* as whether every line parallel to **b** satisfies '$f\xi$'.

To put the point differently: She knows that, whatever contribution 'dir **a**' makes to the determination of the truth-value of 'F(dir **a**)', then, if **a** is parallel to **b**, 'dir **b**' makes just the same contribution. She knows, that is, that, so far as the truth-value of the sentence is concerned, it matters not at all if '**a**' is replaced by '**b**'--if **a** is parallel to **b**. This, of course, is just to re-emphasize that 'dir **a**' and 'dir **b**' are intersubstitutable, *salva veritate*, in such circumstances; and our speaker, if she has a Theoretical Understanding of the Definition, knows that they are.

The ascription of a capacity to refer to an object by means of a given expression is, indeed, only intelligible if the speaker knows that that expression makes some uniform contribution to the determination of the truth-values of sentences which contain it. But our speaker *does* know that. Moreover, Reductionism, as we discussed earlier, properly emphasizes that reference can not be ascribed to names except insofar as we conceive of those names as semantic constituents of sentences, as making a regular contribution to the determination of the truth-values of sentences. But if that is why we are in the business of assigning reference to names, we must ask ourselves another question: Given that, if **a** ∥ **b**, 'dir **a**' and 'dir **b**' are intersubstitutable and that our speaker knows it, what reason can there be to assign *different* references to 'dir **a**' and 'dir **b**'?

It is at best pointless to ascribe different references to expressions which make the same contribution to the determination of the truth-values of such sentences: Indeed, it is theoretically unjustifiable. The Ontology of this (part of the) language ought not be any more rich than is required to explain the behavior of the sentences in question. If we ascribe different references to 'dir **a**' and to 'dir **b**', then it is utterly obscure why there *could not* be a predicate, say 'Qξ', which, though **a** and **b** were parallel, dir **a** satisfied but dir **b** did not. It is not an accident, a result of the impoverishment of the language, that there is no such predicate: There can be no such predicate, if speakers understand sentences 'about directions' in accord with the Contextual Definition.[20] Hence we can not ascribe different references to 'dir **a**' and 'dir **b**'.

The *general* principle for which I have just argued might well deserve the name "The Context Principle".[21] What it states is that the notion of reference is subject to certain theoretical constraints deriving from the role it plays in semantic theory. Frege says that one ought not to ask after the meaning (or reference) "of a word in isolation, but only in the context of a proposition".[22] Why not?

---

[20] Surely it is now time to clear up any worries about intensional contexts. One may have been saying to oneself that I must be, perhaps understandably and legitimately, setting such contexts aside for now. But, in fact, the Contextual Definition we are considering *requires* that we do so. Since the form of the definition requires that we quantify into the predicates 'fξ' on the right-hand side, those predicates must themselves be extensional, with respect to the argument-place bound. This in turn guarantees that the predicates "$(\forall x)(xR\xi \rightarrow fx)$" are extensional.

[21] The great advantage of this interpretation of the Principle, to my mind, over such interpretations as that due to Wright, is that it once again makes the Principle one which is *relatively uncontroversial*, which is motivated by broadly accepted, indeed now common, features of Frege's philosophy of logic. A plausible interpretation of the Context Principle must explain why Frege thought that no argument was required for it.

[22] Frege, p. x. Of course, Frege had yet to distinguish Sense from Reference when he wrote the *Grundlagen*. See Dummett, *Frege*, pp. 192-6, 494ff., and his *The Interpreta-*

Because the ascription of reference to names only makes sense as part of the explanation of the truth-conditions of sentences containing those names, as part of a particular *theoretical project*. The Context Principle amounts to an injunction never to lose sight of the wider theoretical context in which the notion of reference--placed there by Frege--has its home; and, as I have interpreted it, it states constraints which this theoretical context puts upon our use (in semantic theory) of the notion of reference.

As I am now understanding it, the Context Principle has two parts. Firstly: The ascription of reference to an expression is only intelligible if that expression makes a regular, or uniform, contribution to the determination of the truth-values of sentences in which it occurs. This first claim entails that the *general* explanation of, say, the notion of the reference of a name can *only* be given in terms of the kind of contribution which names make to the determination of the truth-values of sentences: For it is only because names make such a contribution that a notion of reference is needed for them at all.[23]

---

tion of Frege's Philosophy (London: Duckworth, 1981), pp. 369-74, 380-87.

[23] Other of Frege's theses also fall into place here. Most of these follow from the fact that, for Frege, the word "object" is now to be *explained* in such a way that objects are, precisely, the referents of names. (See Dummett, *Frege*, p. 471.) The claim that "the referents of our words are what we talk about" can now be seen, indeed, as an expression of

Secondly, the ascription of different references to expressions which make the *same* contribution is unjustifiable. This part of the principle is a sort of converse of Frege's view that the sense of every name must include a criterion of identity for the object which is its referent. It states, instead, that *if* speakers have some criterion for the identity of the referents of some class of names, if speakers treat such names as intersubstitutable *salva veritate* if a certain circumstance obtains, and so as co-referential if that circumstance obtains, then, if that circum-

---

Frege's realism: Namely, that "what we talk about"--namely, people, stars, and so forth--are the referents of our words--are what contribute to the determination of the truth-values of sentences.

It is, in my opinion, simply a mistake to worry that the notion of *reference*, as I have expounded it here, is neither the notion of reference to which we ordinarily appeal, when we ask, philosophically, whether a term refers, nor the notion of reference we use in ordinary language, when we ask to what someone is referring. For, first, this notion of reference is an unabashedly theoretical one, which relates to the ordinary notion in the usual way, as the physicist's notion of temperature relates to our ordinary one.

Secondly, there is now welcome space for debate about the *nature* of reference in any given case: Frege's view that "the referents of our words are what we talk about" may be adopted, opposed, or whatever, in *individual* cases. Dummett's discussion of the Context Principle, at pp. 499ff. of *Frege*, is best understood as raising just this sort of question. And, indeed, Dummett must be correct about at least this much: That the Context Principle alone does not, and can not, even begin to address the more *metaphysical* worries one might have regarding whether abstract objects are, as Frege would have it, mind-independent or are, as Dummett himself suggests (regarding a specific sort of abstract object), mind-dependent, 'mere reflections of language'.

stance in fact obtains, the names *are* co-referential (unless, like names of demons, they fail to refer at all).[24]

The importance of this Principle to our problem can not be over-stated. Consider again the Contextual Definition of identity-statements containing names of directions:

dir **a** = dir **b** if, and only if, **a** ∥ **b**

What, according to a Reductionist, are the references of 'dir **a**' and 'dir **b**'? The question may seem plainly unfair, but it is not: These are expressions of the language, expressions which make a regular, uniform contribution to the determination of the truth-conditions of sentences of the language. Our semantic theory must explain the truth-conditions of sentences which contain these expressions, just as it explains the truth-conditions of any other sentence. In the broad sense in which I am using the word 'reference' here, to explain what contribution 'dir **a**' makes to the determination of the truth-values of sentences in which it occurs *just is* to explain what its reference is.

The most immediate temptation is to say that 'dir **a**' refers to **a**; that 'dir **b**' refers to **b**; and that it is the use of '=' which is misleading. The symbol '=' is ordinarily used for the identity-relation: In this case, however, it is

---

[24] One should not confuse the use of 'criterion' in 'criterion of identity' with the use made by Wittgenstein. As I am using the term, a 'criterion' of identity is just a sufficient condition for identity. The notion of a *criterion* has no epistemological biases built into it. (The notion of a criterion of identity, on the other hand, has non-trivial epistemological aspects.)

being used as a symbol for the relation of parallelism. (As for the apparently functional expression 'dir ℓ', there are lots of options.)[28] This construal of 'dir **a** = dir **b**', however, is inconsistent with the Context Principle, as interpreted above: If **a** is parallel to **b**, then 'dir **a**' and 'dir **b**' make the same (semantic) contribution to sentences in which they occur. Hence, if, as we are assuming for argument, **a** is parallel to **b**, it is theoretically otiose, and unjustifiable, to assign different references to 'dir **a**' and to 'dir **b**'. Hence, that proposal fails.

So, we may now conclude, 'dir **a**' and 'dir **b**' refer to the same entity; and so, all we need conclude now is that 'dir **a**' and 'dir **b**' refer to the same object.

It would be nice if, granted the cogency of this argument, we could declare Wright the victor and go home, as once I was happy to do. But we can *not* pass from the intermediate conclusion that 'dir **a**' and 'dir **b**' have the same reference, that they make the same semantic contribution, to the claim that they refer to the same object. One is plainly tempted to make that inference: The expressions 'dir **a**' and 'dir **b**' are indeed singular terms. This temptation may well feed a sense that all is too easy so far, as well: One may well misplace one's criticism, arguing that the notion of *reference* is here too formal, too thin. But it is not the

[28] Among them, that it is a name of the identity-function; that it is semantically inert, but serves as a kind of reminder; and so forth.

generalized notion of reference, sometimes known as the notion of *semantic value*,[24] which is at fault: The fundamental difficulty with Wright's view--with the view that a Theoretical Understanding suffices for possession of an ability to refer to abstract objects--is that this inference, from co-referentiality to co-reference to some *object*, is unjustifiable.

4. Two Kinds of Predicates

We saw earlier that there are a number of reasons to be dissatisfied with Wright's view. In particular, we saw that his view leads to wild ontological proliferation; the diagnosis of the cause of this proliferation is now our chief task. Moreover, we saw that there is some question how the knowledge which Wright's view requires a speaker to have, if she is to be able to refer to abstract objects, can possibly play the role assigned to it. How can whether one knows that a relation is an equivalence relation and whether one knows that certain predicates are congruences with respect to it determine whether one is able to refer to objects of a given sort?

The short answer to the latter question, one would expect, is that the knowledge in question constitutes grasp of a criterion of identity for the objects in question. As

---

[24] See Michael Dummett, *The Interpretation of Frege's Philosophy*, Ch. 7.

we saw earlier, possession of such knowledge amounts precisely to knowledge that, since 'fnc **a** = fnc **b**' is defined to be true just in case '**a**R**b**' is true (where '$\eta R \xi$' is the relevant equivalence relation), '=', in this context, is--or, has the formal properties of--identity; and that, since each of the predicates "$F\xi$", which can occur in a sentence of the form "F(fnc **a**)", is defined as equivalent to a predicate which is a congruence with respect to '$\eta R \xi$', Leibniz's Law holds.

Whether possession of such knowledge suffices for grasp of a criterion for the identity of the objects in question is, however, not yet decided by these considerations: One might well be troubled by what seems an excessively *formal* characterization of what one needs to know if one is to grasp a criterion of identity. Moreover, even if such objections were answered, we should still need to ask ourselves whether, *so explained*, grasp of a criterion of identity for a sort of object necessarily confers an ability to refer to those objects.

As I intend to avoid the notion of a criterion of identity, so far as is possible, I shall say no more about it here. What is important for the moment is that the question, whether speakers who have a Theoretical Understanding of a Contextual Definition are therefore able to refer to abstract objects, may be raised directly whether or not the claim that they are is defended by means of an appeal to the notion of a criterion of identity.[27]

We need to consider a distinction which may seem quite distant. For the purposes of our discussion here, it will be more convenient to present it in the context of a discussion of a Contextual Definition other than that of names of directions. Let us consider a language which contains names of (physical) books, of word-types,[28] and so forth; but which contains no names of book-types or, as I shall call them, works (in the sense of a work of literature). We may define an equivalence relation, 'η copies ξ', by stipulating that **a** copies **b** if, and only if, the book **a** contains precisely the same (type) words as the book **b**, in the same order. We may thus introduce expressions of the form 'work **a**' and a range of predicates suitable for use with such names, by means of a Contextual Definition.

There are two quite different sorts of predicates which speakers of the augmented language might understand. On the one hand, there are predicates like 'ξ contains the word "fantasy"'; on the other, there are predicates like 'ξ is

---

[27] I should argue that the distinction I draw in the remainder of this section is also relevant to this issue. But we can not pursue that problem now.

Of course, if speakers who have a Theoretical Understanding are not therefore able to refer to abstract objects, that may itself give us good reason not to characterize the notion of a criterion of identity in this way.

[28] I am making no assumption that there *are* any such things: It is quite irrelevant whether these 'names' actually refer or are capable of elimination.

such that all copies of it (i.e., all books which copy it) have a torn page'. Both predicates are, of course, congruences with respect to the equivalence relation 'copies'. Intuitively, however, the predicates are quite different: One is tempted to say something like, "The former predicate expresses a *property* of works; the latter does not".

This talk of 'properties' is notoriously slippery, though heuristically useful. I can not explain this distinction in full generality: What is required, for present purposes, is that we understand the distinction between predicates which express properties and those which do not *as it arises in the cases in which we are interested*, namely, cases of names and predicates introduced by Contextual Definition.

The phenomenon is not, however, limited to such cases. One is just as tempted to say that such predicates as 'has only blue-eyed children' or 'has a son who is in London' do not express properties of a person in the same sense that 'has blue eyes' or 'is now in London' do. Indeed, a difference precisely analogous to that we discussed above arises with respect to the sentences "The father of RH is 48 years old" and "The father of RH has only male children". Now, *one* feature of the latter sentence is that, ordinarily, in order to determine its truth-value, one must know whether the father of RH has any other children and, if so, who they are. In the case of the former sentence, however, one does not, ordinarily, need such knowledge; to determine its truth-value, one does, naturally, need to know who RH is, who his father is, and so forth; but one does *not* ordinarily need to know whether he has any paternal half-siblings and, if so, who they are.

Similarly, to determine the truth-value of the sentence "The work of which this book is a copy has only copies which have torn pages",[29] one ordinarily must know which other books copy it. But to determine the truth-value of the sentence "The work of which this book is a copy contains the word 'fantasy'", one does not, ordinarily, need to know any such thing: One need only have a look at the book in question or, indeed, any book which copies it.

What distinguishes the two sorts of predicates which I am here discussing is, thus, precisely this: Whether, to determine the truth-values of sentences containing such predicates, one is ordinarily required to know whether and which other books copy some given book. I shall say that a predicate of (say) works is *variant* if one may ordinarily determine the truth-value of a (simple) sentence containing it without knowing whether and which books copy some given

---

[29] It is worth remarking here that such 'constant' predicates can be formed by means of quantifiers other than 'all'. For example: "ξ is such that some copy of it is owned by Quine"; "ξ is such that most copies of it are in libraries"; "ξ is such that exactly five copies of it have torn pages". All such predicates are congruences with respect to 'copies'.

Reflection on such predicates may help persuade one of the intuitive basis of the distinction drawn here.

book. I shall say that a predicate is *constant* otherwise. The generalization of the distinction to predicates of other sorts of objects, names for which are defined in terms of other equivalence relations, should be obvious.[30]

By saying that one must *ordinarily* know whether, and if so which, other books copy **a** to determine whether each copy of it has a torn page, I mean to recognize that, one may, on any given occasion, be able to determine that not all copies of some work have a torn page without knowing whether any other books copy the given copy **a**--if **a** itself has no torn page--or without knowing which other books copy **a**--if, say, **b** copies **a** and has no torn page. In general, however, one's ability to determine whether each copy of **a** has a torn page depends upon one's ability to determine which books copy **a**. There is a generally (or universally) applicable procedure for determining whether each copy of **a** has a torn page, and

---

[30] It would be nice to be able to formulate a more general distinction here. One promising way to do so would be as follows. Every object is capable of being identified in a variety of different ways. Now, there are certain predicates of, say, people such that one can, in general, determine whether a particular person satisfies the predicate if one identifies the person in any of a variety of ways. In particular, one does not, in such cases, need to be able to identify the person in any way *other* than the given way. So, for example, I can determine how tall John Doe is if am able to identify him as John Doe, or as John Smith, or as the Grim Reaper, or however. To determine whether the father of John Doe has only blue-eyed children, however, I need to know--basically--whether it is possible to identify him in various other ways: As, say, the father of Jane Doe.

There is plainly work to do to make such a distinction work, but, if it did work, it would be possible to derive the version given in the text from the more general version.

the application of this procedure requires one to determine which other books copy **a**. On the other hand, while there is such a procedure for determining whether a given work contains the word "dog", there is *also* a procedure which does not require one to determine which other books copy **a**.[31] It is the existence of such a generally applicable procedure in the one case, though not in the other, which distinguishes the two sorts of predicates.

With the use of the word 'ordinarily' explained in this way, we may continue to use it as it was used above.

5. That a Theoretical Understanding Will Not Suffice: Reference and Logical Type

I said earlier that, though expressions, introduced by a Contextual Definition of which a speaker has a Theoretical Understanding, refer, that, indeed, though expressions which one would expect refer to the same entity do, we can not pass from this claim to the conclusion that these expressions refer to the same *object*. With the distinction between variant and constant predicates in place, we are now in a position to see why.

---

[31] There is an obvious connection here with Dummett's distinction between 'direct' and 'indirect' means of verification. See his "What Is a Theory of Meaning? (II)", in G. Evans and J. McDowell, eds., *Truth and Meaning* (Oxford: Oxford University Press, 1976), pp. 67-138, at pp. 115ff.

The difficulty concerns cases in which most of the predicates which a speaker understands are constant.[32] Let us, then, consider a case in which we can give a Contextual Definition and in which, plausibly, most predicates we do in fact understand are of this sort. The Definition in question is one we have already discussed, namely, that of names of poursons. Recall that the definition is given in terms of an equivalence relation '$\eta P \xi$', defined as the smallest equivalence relation holding between arbitrarily paired individuals who were alive on 21 March 1990. The definition is then of the form:

$$F(\text{pour } \mathbf{a}) \equiv_{df} (\forall x)(xP\mathbf{a} \rightarrow fx)$$

There are many predicates of poursons which we can thus define: One, "Brown-haired(ξ)", is defined as equivalent to "$(\forall x)(xP\xi \rightarrow$ Brown-haired$(x))$". This predicate is clearly constant: In order to determine whether "Brown-haired(pour (George Bush))" is true or not, I must know who is paired with George Bush. The same can be said of most other predicates which we can define.[33]

---

[32] The use of the term 'most' here will be further explained below.

[33] One may have been wanting to remark that there seems to be a connection between the notion of a variant predicate and the notion of projectibility: Perhaps the reason there are no variant predicates of poursons is because there are no predicates of *persons* which project over the relevant equivalence classes. See Sylvain Bromberger, "Types and Tokens in Linguistics", in A. George, ed., *Reflections on Chomsky* (Oxford: Blackwell, 1989), pp. 58-89.

Now, I have argued above that the expression 'pour **a**' has a reference. I have also argued that, if **aPb**, then 'pour **a**' and 'pour **b**' have the same reference. But, on the other hand, I do not want to allow that 'pour **a**' refers to an object. So what is this common reference?

What does it *look* like their common reference is? What is, from our present perspective, peculiar about the name 'pour **a**' is that, in order to determine the truth-value of a sentence of the form "F(pour **a**)", one must, ordinarily, know just which other person has been paired with **a**; what is peculiar about names *like* 'pour **a**' is that one must ordinarily know the contents of the relevant equivalence class {x: pour(x)=pour(a)} in order to determine whether a sentence containing the name is true. The sort of expression which typically induces this sort of requirement is a quantifier; in particular, a quantifier restricted to a given equivalence class. For example, the sentence "All copies of *Grundlagen* have a torn page" contains such a quantifier; it is because the sentence contains such a quantifier that one can not, ordinarily, determine its truth-value unless one knows just which copies of *Grundlagen* exist.

There are complications here. For each predicate "fξ", the predicate "(∀x)(xPξ → fx)" does project over the relevant equivalence classes. After all, it is a *logical truth* that it is a congruence with respect to "ηPξ". Whether the idea is salvageable, I do not know.

One might suggest, therefore, that the true logical form of "F(pour **a**)" is:

$\mathrm{Pour}_{\varkappa}(\mathbf{a})(fx)$

Here, "$\mathrm{Pour}_{\varkappa}(\xi)(\Phi x)$" is an operator which forms a quantifier from a term: It is equivalent to "$(x)(xP\xi \rightarrow \Phi x)$". This quantifier is *itself* a congruence with respect to '$\eta P\xi$', for the same reason that predicates of the form "$(x)(xP\xi \rightarrow fx)$" are. Hence, when 'pour **a**' is understood as a quantifier, 'pour **a**' and 'pour **b**' *do* have the same reference, if **a** is parallel to **b**. The two expressions do not refer to the same object, for they are not names. Rather, in the usual Fregean parlance, these expressions refer to the same *second-level concept*, as is their lot, their being quantifiers. Such an account thus takes full notice of the *effects* of a Theoretical Understanding of the Definition: And if so, there is no theoretical justification for taking 'pour **a**', and other such 'names', to refer to (abstract) objects.

However, the fact, if it is one, that there is some sort of quantificational structure in "F(pour **a**)" does not entail that 'pour **a**' must *itself* be read as a quantifier: We may interpret "Brown-haired(pour **a**)", not as suggested above, but as would be obvious in the case of "All copies of *Grundlagen* have a torn page". In the latter case, the most natural interpretation would be:

$(\forall x)(\text{work } x = \textit{Grundlagen} \rightarrow \text{torn-page}(x))$

Hence, in this case, we may try:

$$(\forall x)(\text{pour } x = \text{pour } \mathbf{a} \rightarrow \text{brown-haired}(x))$$

Here, 'pour **a**' is treated as a singular term: The quantificational structure is located not in 'pour **a**' itself, but in the predicate "Brown-haired(ξ)".

I said earlier that the difficulty concerned cases in which *most*--rather than *all*--of the predicates speakers understand are constant. We can now see why. Identity is *itself*, in the obvious extended sense, a variant *relation*: To know whether "pour **a** = pour **b**" is true, one must, indeed, determine whether **aPb**: But one does not, in general, need to know whether any objects other than **b** bear P to **a**.[34] It is certainly true that one may provide a quantificational construal of identity-statements containing names of poursons. Viz.:

$$\text{pour } \mathbf{a} = \text{pour } \mathbf{b} \text{ iff } (\forall\Phi)[\text{Pour}_{\text{x}}(\mathbf{a})(\Phi x) \leftrightarrow \text{Pour}_{\text{x}}(\mathbf{b})(\Phi x)]$$

But the motivation for such a construal is absent in this case: We are assuming, for the moment, driven to construe sentences containing constant predicates as having a quantificational structure by the fact that such sentences do contain constant predicates. There is no such requirement in this case.[35]

---

[34] This point is slightly obscured by the nature of the relation '$\xi P \eta$': For exactly one other object bears P to **a**. But the general point should be clear.

[35] There is a general argument that such names can not be eliminated, though for reasons different from those given by Wright: See "Trans-sortal Identification".

The claim that the ,xpressions 'pour **a**' are quantifiers thus seems dubious. But if identity is the only variant (one- or many-placed) predicate one understands, then the only contexts one understands, in which 'pour **a**' occurs, are identities: These contexts include not only identity-*statements* but (partially) open identities embedded in more complex sentences. Nonetheless, the only (primitive or simple) predicate one understands which is applicable to names of poursons is the identity-sign.

There is thus very little which we speakers can say about poursons. Any sense one might have had that, whatever these objects are, they are very peculiar, now vanishes: Such objects have neither psychological nor physical properties; indeed, the *only* properties they seem to have are identity with and distinctness from one another. Such objects can play almost no role in our thought. More to the point, there seems little reason to deny that the objects to which we refer, if we refer to any, by means of names 'of poursons' are just equivalence classes. For consider the functional expression 'poureq(ξ)' defined as:

poureq(**a**) $=_{df}$ {x: **a**Px}

Clearly, poureq(**a**) is identical with poureq(**b**) if, and only if, pour(**a**) is identical with pour(**b**).

We should, I suggest, be utterly stumped if asked to say what in our use (or understanding) of names of the form 'pour(**a**)' distinguishes our use of them from our use of names of the form 'poureq(a)'. If nothing does, then it is difficult to see how we can defend the claim that names of the two kinds refer to objects of different kinds. (And, as was said earlier, the claim that names of poursons refer to equivalence classes is, so far as I am concerned, equivalent to the claim that there are no poursons, in the sense in which poursons are supposed to be *sui generis*.)

I can not pursue this point very far here.[36] The general claim upon which this argument depends is that, if a speaker is to grasp a criterion of identity for names of a given class, then she must understand a wide range of variant predicates fit for use with those names. The notion of a variant predicate is, recall, closely connected to the intuitive notion of a *property* of an object: So, intuitively, the claim is that one's understanding of a criterion of identity for names of objects of some sort depends upon one's understanding what sorts of properties those objects are to be understood as having or failing to have. Our conception of the kinds of properties such objects may have informs our conception of the *kind* of objects these are: Our conception of the sort of objects to which names in a given class is informed by an *Ideology* about those objects.--One might well say that, at least, how 'robust' the objects seem

---

[36] See "Trans-Sortal Identification" for a discussion of it. My approach to these problems is quite different in that paper, so this point is also treated quite differently there.

to be is closely connected with how 'robust' our Ideology about those objects is.

I shall have to leave this claim undefended here: Nonetheless, the strategy of the argument is easily conveyed, and, as should be clear, it is closely related to what was said earlier about the case of poursons. Consider, again, the Contextual Definition of directions: The Definition introduces expressions of the form 'dir ℓ', with which we associate particular identity-conditions. There are a variety of other functional expressions which, extensionally (and indeed necessarily), have the same identity-conditions. Among these are 'the line through the Origin parallel to ℓ', 'the set of all lines parallel to ℓ', 'the angle at which ℓ intersects the x-axis', and so forth.

Despite the fact that these functional expressions all have the same identity-conditions, we conceive of the last three as referring to distinct sorts of objects: The former, to a line; the next, to a set; the last, to a real number. In what does it consist that the three refer to different sorts of objects? One might say that the *sortal* predicates which occur in the different functional expressions provide for the distinction: But the question is precisely *how* they do so. And one plausible, partial answer to this question is this one: Our conception of the sorts of properties which the line through the origin parallel to **a** has is quite different from our conception of the sorts of properties which the set of all lines parallel to a has; our Ideology about the former is quite different from our Ideology about the latter.[37]

If that is correct, then, an understanding of the sort of object to which such an expression refers depends upon a grasp of the sorts of properties such objects may have: That is, upon an understanding of a variety of variant predicates of such objects. The argument against Wright's view would then be complete, and, moreover, the foundations for an argument for the existence of abstract objects *and* for their distinctness, in general, from sets would have been laid: For the Ideology which we associate with, say, directions or letter-types have is quite different both from the Ideology we associate with sets and from that we associate with concrete objects of any sort.

6. Meaning and Understanding, and Variant and Constant Predicates

Thus far, I have merely *suggested* that expressions introduced by Contextual Definition, of which speakers have but a Theoretical Understanding, are not the names they appear to be, but are, instead, quantifiers. In the next two sections, I shall argue for this view. I shall argue that,

---

[37] There are a number of similarities between these ideas and Wittgenstein's discussion of ostensive definition, by which they were inspired. See *Philosophical Investigations*, tr. by G.E.M. Anscombe (Oxford: Blackwell, 1958), §§28ff.

in cases in which a speaker understands some variant predicates of the objects in question, there is a difference between her understanding of sentences which contain variant predicates and those which contain constant ones. This difference must then be reflected in the semantic theory for this (part of the) language. *If* sentences which contain variant predicates are, as Wright's view should have it, to be accorded a simple (subject-predicate) structure, sentences which contain constant predicates can not be. Thus, even if there are no variant predicates which speakers understand, sentences which contain constant predicates can yet not be accorded a simple structure, but must be accorded a quantificational structure.

It is worth reminding ourselves how the distinction between variant and constant predicates was drawn. A variant predicate of a sort of object is one with respect to which speakers have certain sorts of abilities: In particular, speakers are required to be able, ordinarily, to determine the truth-value of a sentence "F(fnc **a**)", containing such a predicate, without knowing which, if any, other objects bear the relevant equivalence relation to **a**. A predicate is constant otherwise. Note that the distinction is one between sorts of *predicates*: We are to consider a variety of sentences of the form "F(fnc **a**)" to determine whether "Fξ" is variant or constant. That is just to say that the abilities which distinguish these two sorts of predicates relate to sentences of a certain form: Hence, if our semantic theory is to take account of this distinction, it must distinguish sentences of the *form* "F(fnc a)"--where "Fξ" is a variant predicate--from those of the *form* "C(fnc a)"--where "Cξ" is not.

It is also worth remarking, here, that the distinction is drawn in terms of abilities which speakers may or may not have. I am not *assuming* that the distinction is one between predicates of which speakers have one sort of *understanding* and those of which they have some other sort of understanding. Furthermore, *if* the distinction is, in fact, a distinction between sorts of understanding, that, in itself, does not in any way entail that possession of one sort of understanding or the other *consists in* possession of just the abilities in terms of which the distinction has been drawn; nor does it entail that it consists in the possession of any other abilities whatsoever. The argument here is entirely independent of the answers to such questions.

It should not, however, be assumed that, because the distinction is drawn in terms of abilities which speakers may or may not have, it can *not* be a distinction between sorts of understanding.

The importance, to certain sorts of metaphysical views, of the claim that understanding is constituted by possession of certain (linguistic) abilities has wrongly made philosophers who are not enamored of those arguments wary of *any* introduction of linguistic abilities into discussions of ontology or metaphysics. However, whether or not such arguments should stand, they are motivated by a genuine insight: Namely, that it is *the linguistic abilities which speakers have whose possession theories of meaning are intended to explain*. However one thinks semantics ought to be done--*a la* Hintikka, Lewis, Davidson, Dummett, or a host of others--the task is essentially the same: To explain speakers' possession of the linguistic abilities whose possession their linguistic behavior manifests.[38]

The implications of this point are difficult to ascertain: Hence the debate over anti-realism. Fortunately, however, it is one of its relatively immediate implications which is crucial here. Suppose that there are two classes of sentences, both having the same surface structure, with respect to which speakers have the same linguistic abilities, except that speakers have some linguistic ability with respect to sentences in the former class though not with

---

[38] Use of the notion of *manifestation*, it is worth saying here, also invites the suspicious reaction of which I am complaining. The notion is, however, not in any way peculiar itself. As I use it, and, indeed, as Dummett himself uses it, to say that one's behavior *manifests* the possession of a linguistic ability is just to say that that behavior *reveals* that one has that ability; or, that *evidences* that one has the ability; or, that it provides good reason to believe that one has the ability. *Why* it does so is another question, a possible, indeed, popular, answer to which is that possession of that ability is causally implicated in production of that behavior. There are, of course, *other* answers to the question, one of which Dummett has made famous.

respect to those in the latter. No theory of meaning which fails to explain this difference can be adequate: Any adequate theory must explain why speakers possess the ability in question with respect to the one class though not with respect to the other. Hence, no adequate theory can treat sentences in the two classes in the same way: There must be some difference in the *way* speakers understand sentences in the two classes (for the difference, by hypothesis, concerns an ability speakers have, generally, with respect to a class of sentences). For, if a uniform treatment were adequate to explain speakers' possession of the ability in the former case, it would of necessity also be adequate to explain their possession of that same ability in the latter case, a case in which the ability is, by hypothesis, absent.

The claim for which I have just argued may thus be stated as follows: If a speaker has linguistic abilities with respect to one class of sentences which she does not have with respect to another, then she must understand sentences in the two classes differently. This principle entails that the distinction between variant and constant predicates, on which we are focusing, is a distinction of which any semantic theory must take notice. In order to draw this conclusion, however, we must show that the abilities, in terms of whose possession the distinction was drawn, are in fact *linguistic* abilities: For the argument just given depends upon the supposition that speakers have different *linguistic* abilities *vis-a-vis* the relevant classes of sentences; and the argument promised is concerned solely with the different abilities which speakers have *vis-a-vis* sentences containing variant and constant predicates.

A full argument that these abilities are linguistic ones would depend upon the defense of a general distinction between linguistic and non-linguistic abilities, which, sadly but not surprisingly, I am in no position to present. I should urge that the fact that these abilities concern how speakers are, quite generally, able to determine the truth-values of a broad range of sentences makes it quite plausible that these abilities are linguistic, if any are. Moreover, I should claim that the distinction is crucial to a correct explanation of the notion of a criterion of identity.[39] But I am not going to pursue that point here.

Instead, I shall offer a more direct, though somewhat speculative, argument that the distinction between variant and constant predicates is one of which a theory of meaning must take notice. The argument appeals to a principle about the meanings of predicates: Namely, that evidential relations among sentences are *relevant* to the theory of meaning; that predicates "Fξ" and "Gξ", which bear different evidential relations to other predicates, differ in meaning, in Sense. (As earlier, this principle in no way entails that

---

[39] See, again, "Trans-sortal Identification".

the notion of meaning is *itself* to be *explained* in terms of the notion of evidence.)

To say that, in order to determine the truth-value of "F(dir **a**)", one need not know which, if any, other lines are parallel to the line **a**, is precisely to say that there is a very strong evidential relationship between a sentence "f**a**" and "F(dir a)"--where "fξ" is a simple predicate.[40] Were there, then, a predicate "Cξ", which was constant, however strongly *equivalent* to "Fξ", the two would have to differ in at least this respect: Namely, that "Fξ" should bear stronger[41] evidential relations to certain sentences than should "Cξ". For, by hypothesis, it is not the case that knowledge of the truth of any sentence of the form "f**b**"--where **b** is parallel to **a**--suffices for knowledge that "C(dir **a**)", since, otherwise, one could determine whether "C(dir **a**)" was true by determining whether or not any such sentence was true.

Now I am not claiming that there must always be differences, in this respect, between any given variant predicate and some *particular* constant predicate in a given

---

[40] The predicate "fξ" is, in general, just that predicate in terms of which "Fξ" itself was defined: I.e., the "fξ" of "$(\forall x)(xRa \rightarrow fx)$".

A variant predicate is, thus, one which is defined in terms of a predicate *itself* a congruence with respect to the relation in question. (But note, not every such predicate is variant: Consider "$(\forall x)(xRa \rightarrow (\forall y)(yRx \rightarrow fx))$".)

[41] "Stronger", of course, because "Cξ" may bear *some* evidential relation to "fξ".

language. Rather, I am arguing that there is a *sort* of sentence to which sentences containing variant predicates bear strong evidential relationships, to which sentences containing constant predicates do not. What distinguishes variant predicates from constant ones, from this point of view, is that, in general, variant predicates bear a strong evidential relation to *such* sentences--particular simple sentences to which they are closely related: Constant predicates on the other hand, do not. And for this reason, I claim, there is a difference between the *sorts* of meaning which the two sorts of predicates have.

7. That a Theoretical Understanding Will Not Suffice (VI): Quantification and Logical Form

Any adequate semantic theory must, therefore, treat variant predicates, as a class, differently from constant predicates. The question which we must now address is just how the theory must treat them. In order to argue, at last, that a speaker who has a Theoretical Understanding of a Contextual Definition need not be able to refer to abstract objects, we need to argue that any adequate theory must accord sentences containing constant predicates a quantificational structure.

To show this, I am going to argue that, if there are both constant and variant predicates in the language, the theory must accord sentences which contain constant predicates a quantificational structure. This argument depends upon quite general features of constant predicates and how those features distinguish them from variant predicates, not upon any specific relationship between constant predicates and the variant predicates which, I am supposing, are also present in the language. The argument, that is to say, is designed to show that, because constant predicates are unlike variant predicates, the two must--in general--be treated differently.

The argument is, again, being offered against the view that speakers who have a Theoretical Understanding of a Contextual Definition are, thereby, made able to refer to abstract objects. Hence, we may assume that (simple) sentences containing variant predicates are to be assigned a simple, subject-predicate structure: The semantic theory is assumed to treat a sentence such as "The work of which **a** is a copy contains the word 'fantasy'" by assigning it the structure of "F(work **a**)": That is, the sentence is, according to the theory, true if, and only if, a particular (abstract) object--namely, the work of which **a** is a copy--satisfies the predicate "ξ contains the word 'fantasy'".

There will, of course, be other sentences which contain constant predicates of works, such as "The work of which **a** is a copy has only copies which contain torn pages". Since this sentence does contain a constant predicate, it can not be treated as are sentences which contain variant predicates. The most plausible way to treat it is to assign it the semantic structure of "$(\forall x)(\text{work } x = \text{work } \mathbf{a} \rightarrow fx)$". And, of course, there are sentences, like "All books which are copies of *Grundlagen* have a torn page", in English. Such sentences have a *plainly* quantificational structure: The difficulty, so far as I can see, is thus not to show that sentences containing constant predicates have a quantificational structure. Rather, the difficulty was to find a way of distinguishing them without *assuming* that they have such a structure.

Treating sentences which contain variant and constant predicates in these different ways offers an explanation of the differences between them. On the one hand, sentences which contain variant predicates are assigned a simple structure: Simple sentences, those of the form "Fa", paradigmatically are susceptible to verification without any particular knowledge concerning other objects:[42] Plainly, the semantics of these sentences does not lead one to *expect* that one should need any such knowledge. What typically needs to be done, in such a case, is to identify the object in question and to determine whether it satisfies the predicate: In the case of "The work of which **a** is a copy contains the word 'fantasy'", we need to identify the work in question--to do which we must identify some *one*, *any* one of its

---

[42] It is worth comparing with this point what Wiggins calls the "Only **a** and **b**" condition on identity: See his *Sameness and Substance* (Oxford: Blackwell, 1980), pp. 94ff.

copies--and determine whether it contains the word 'fantasy'.

On the other hand, sentences which contain constant predicates contain a quantifier, whose domain is restricted to other copies of the work in question. Since the truth-value of the sentence depends upon what truth-value each of its instances has--upon whether, in particular, each copy of the work satisfies some given predicate--one ordinarily will need to know, in order to determine the truth-value of the sentence, whether and which other copies of the work exist.

Thus, drawing on our earlier discussion, if speakers do not understand certain sorts of variant predicates fit for use with a class of names introduced by Contextual Definition, then the only contexts in which speakers understand the use of those names are identities; for the variant predicates speakers do understand are derived from the Definition of identity itself, and sentences containing constant predicates have a quantificational structure. As was said above, there is reason to be skeptical that, if that is all speakers understand, they are able to refer, by means of such names, to objects to which they were not already able to refer. But, for the moment, it is perhaps sufficient to note that the distinction between variant and constant predicates has a relevance to our Ontological problems and that it is not clear how Wright can make a place for it.[43]

8. Closing

I have argued here that Wright's attempted defense of the view that there are abstract objects fails. There are cases in which, as I have put it, speakers understand only 'constant' predicates of certain (supposed) objects; in such cases, we are unable to justify the claim that the expressions introduced by the Contextual Definition, singular terms though they may appear to be, are *proper names*. We are, that is, unable to justify the claim that these expressions require to be treated, in a semantic theory, as expressions fit to refer to objects. Indeed, so to treat them would be to treat them as like expressions *variant* predicates of which speakers understand; it would therefore be to treat them in such a way as to make possible an explanation of speakers' possession of just those abilities characteristic of an understanding of variant predicates in a case in which they have no such abilities.[44]

---

[43] For more on the relevance of the distinction, see "Trans-sortal Identification".

[44] I should like to thank George Boolos, Sylvain Bromberger, Michael Dummett, Jim Higginbotham, Tom Kuhn, Jim Page, and Bob Stalnaker for comments upon and criticism of earlier drafts of this paper. I should also like to thank those who attended readings of earlier versions at the Wolfson Colloquium, in Oxford, and at the graduate colloquium at the Massachusetts Institute of Technology.

Paper III

Trans-sortal Identification

0. Opening

Nominalism, in its modern form, is the view that there are no abstract objects. We may refer to its more specific relatives as various forms of Reductionism, the view, in each case, that there are no (abstract) objects of some specific kind. Hence, we may speak of Reductionism about Number Theory, the view that there are no Numbers, and of Reductionism about linguistic objects, the view that there are no letter- or word-types. It is, of course, possible to hold a Reductionist view about some subject-matters, while rejecting the corresponding views about other subject-matters: Nominalism is thus Reductionism about every class of sentences in which occur terms purporting to refer to abstract objects.

Both Reductionism and Nominalism come in a variety of flavors. Following Dummett, I shall refer to the part of our language upon which any specific ontological dispute is focused as the Disputed Class of sentences. We may make, first, a distinction between those positions according to which statements in the Disputed Class have some truth-value, some such statements typically being true, and others false, and those positions which deny this claim. We may refer to views of this latter sort as Fictionalist views: For it is often said, by those who defend such positions, that, for example, statements of Number Theory should not be thought of as up for evaluation as true or false; rather, our talk of Numbers should be thought of as a convenience, a fiction, which facilitates the process of drawing inferences within science.[1]

I shall not be concerned, in this paper, with Fictionalism.[2] The Reductionist views which I shall discuss here are thus committed to the view that some statements of, say, Number Theory are true; others, false. The chief problem then facing such a view is to explain *how* such statements may be true, despite the fact that there are no objects of the sort to which names occurring in them purport to refer.

While the motivation for Reductionism may differ from case to case, and from philosopher to philosopher, this problem is constant. One form of Reductionism, which we may call Semantic Reductionism, is *motivated* by the observation that it may be possible to eliminate names of, say, letter- and word-types from sentences in which they occur: The

---

[1] See here Hartry Field, *Science Without Numbers* (Oxford: Blackwell, 1980). There are plainly a variety of options here: One may say, for example, that statements of Number Theory are, *strictly speaking*, false, and then attempt to explain their apparent utility in terms of the idea of a convenient fiction. On the other hand, one might want to appeal to some notion of truth-within-a-fiction, so that some statement like 'In the fiction of Number Theory, 2+2=4' would be true. However such views are developed, however, no statement of Number Theory will, on any such view, be *true*.

[2] For a discussion of such views, see Bob Hale, *Abstract Objects* (Oxford: Blackwell, 1987), Ch. 5.

elimination shows that no names which purport to refer to abstract objects actually occur in the sentences in the Disputed Class. There is thus no need to take seriously the idea that we refer to such objects and, correspondingly, no need to recognize their existence.

A different sort of Reductionism, which we might call Epistemological Reductionism, is motivated by epistemological difficulties connected with abstract objects. Abstract objects, such as Numbers, are often said to be causally inert: If so, then they can neither be perceived (perception presumably being a causal notion), nor known by their effects: How, then, even if there are such objects, can we know anything about them? Now, we are assuming that there are a variety of true statements of Number Theory, some of which there is no reason to doubt we *know*: Hence, if we had some account of how such statements could be true, though there were no Numbers, and how we could know the truths they express, we could avoid these epistemological difficulties. Classically, such an account would be provided by means of precisely the sort of analysis discussed earlier, the object being to show that 'names of abstract objects' are nothing of the sort; that they merely disguise the true logical form of sentences in which they occur; that they may be eliminated, leaving us without any need for a notion of reference to abstract objects.

Thus, whatever the motivation for Reductionism, the question how statements containing what purport to be names of abstract objects can be true (and which an elimination of names of abstract objects would answer) arises. I do not mean to be suggesting that the problem can only be resolved by means of an elimination of purported names of abstract objects. What must be eliminated is *reference* to abstract objects: The notion of reference to abstract objects might be eliminable, even if names 'of abstract objects' are not. Names which purport to refer to abstract objects might refer, not to abstract objects, but to objects of some less problematic sort. I shall argue in the first section of this paper that names of abstract objects *can not* be eliminated and that, therefore, Reductionism of any sort other than Fictionalism is committed to the elimination *only* of reference to abstract objects, to the view that names 'of abstract objects' refer, but to objects which are not abstract.

The traditional opponent of Nominalism is Platonism,[3] the view that there are abstract objects of various kinds. We may distinguish between views according to which the existence of abstract objects of any given kind is largely independent of empirical matters, and views which according

---

[3] This term is used by Crispin Wright and, following him, by Bob Hale. See Wright's *Frege's Conception of Numbers as Objects* (Aberdeen: Aberdeen University Press, 1983), p. xviii, and Hale's *Abstract Objects*.

to which abstract objects of a given sort may exist, or may not exist, as the empirical facts may have it.[4] The former view, it should be emphasized, is not committed to the claim that the existence of all *particular* abstract objects must be necessary: And, indeed, if the existence of, say, word-types is dependent upon the existence of their tokens, then one who held this view could well deny that it is necessary that word-types exist. The important point, however, is that, according to this view, *if* there are any word-tokens, then, necessarily, there are word-types. There is, that is to say, no empirical question about the existence of word-types as such; there is only the question about the existence of their tokens. According to the latter view, however, there is an additional, non-philosophical question whether there are any word-types, even if the existence of their tokens is granted.

Crispin Wright's argument against Reductionism is that names of and reference to abstract objects can not be eliminated. He considers, as a way of eliminating them, what are commonly known as Contextual Definitions. Such a Definition does not provide an explicit definition of names of abstract objects: Rather, it shows how to translate sentences which contain names of abstract objects into sentences which do

---

[4] Compare here Dummett's discussion of the distinction between what he calls Aristotelian and Strawsonian conceptions of the existence of abstract objects, in his *Frege: Philosophy of Language*, 2nd ed. (London: Duckworth, 1981), pp. 501-4.

not. For example, once we have seen that the direction of one line is the same as that of another line if and only if those lines are parallel, we may offer the following elimination of names of directions:[c]

$$\text{dir } a = \text{dir } b \equiv_{df} a \parallel b$$
$$F(\text{dir } a) \equiv_{df} fa$$

Here 'dir a' is to be read 'the direction of a'; '$f\xi$' is a predicate which is a *congruence* with respect to the (equivalence) relation of parallelism[d] and which is suitably related to '$F\xi$' (so that the truth-conditions come out right). Note that these two conditions on '$f\xi$' guarantee, first, that the Definition of identity preserves its formal properties (reflexivity, symmetricity, and transitivity) and, secondly, that names of 'the same direction' are intersubstitutable *salva veritate* within statements 'about directions': I.e., they guarantee that Leibniz's Law is valid in the extended language (assuming that it was valid in the original language).

---

[c] See Wright, pp. 29ff. See also Gottlob Frege, *The Foundations of Arithmetic*, 2nd. rev. ed., tr. by J.L. Austin (Evanston IL: Northwestern University Press, 1980), §65.

[d] To say that a relation '$\xi R \eta$' is an equivalence relation is to say that it is reflexive--$\forall x(xRx)$--symmetric--$\forall x \forall y(xRy \leftrightarrow yRx)$--and transitive--$\forall x \forall y \forall z(xRy \,\&\, yRz \rightarrow xRz)$. A predicate '$F\xi$' is a congruence with respect to a relation '$\xi R \eta$' if, and only if, $\forall x \forall y(Fx \,\&\, xRy \rightarrow Fy)$. If '$\xi R \eta$' is an equivalence relation, then '$F\xi$' is a congruence if, and only if, $\forall x \forall y(xRy \rightarrow (Fx \leftrightarrow Fy))$. The notion of congruence can be extended to functions and many-placed predicates in obvious ways.

In general, a Contextual Definition is of the form:[7]

fnc a = fnc b $\equiv_{df}$ aRb
F(fnc a) $\equiv_{df}$ fa

Here, 'fnc ξ' is the (functional) expression to be eliminated; 'ξRη' is an appropriate equivalence relation; the predicates 'fξ' are suitably related to the various predicates 'Fξ' which occur on the left-hand side and are congruences with respect to the relation 'ξRη'.

Wright argues, for reasons I shall not discuss here,[8] that no Contextual Definition can succeed in eliminating reference to abstract objects. His argument, as I understand it, commits him to the view that there are abstract objects of various sorts corresponding to every equivalence relation. The point may, perhaps, be seen most clearly if we conceive of Contextual Definitions, not as a way of *eliminating* names of abstract objects, but as a way of *introducing* names of abstract objects. So, consider the equivalence relation 'ξ was born on the same day as η' (understood as true only of pairs of people) and the Definition:

dap(a) = dap(b) $\equiv_{df}$ a was born on the same day as b
F(dap a) $\equiv_{df}$ fa

(The predicates 'fξ' are, of course, required to be congruences with respect to the equivalence relation.) The

---

[7] Of course, other parts of the Definition will be required to deal with two- and many-place predicates. But the generalization poses no special problems.

[8] See Wright. See also my "Whether Structure May Be Misleading", *forthcoming*, for criticism of the view I am about to attribute to Wright.

functional expression 'dap ξ' may be read 'the day-person of ξ'. Wright's view, as I understand it,[9] is that there are such objects as day-persons: For his view is that, if expressions in a given class function, so far as syntax is concerned, as singular terms, and if those expressions occur in some true sentences,[10] then the terms in question *refer*.

Now, it seems to me implausible that there are any such objects as day-persons. I shall be arguing that *some but not all* such Contextual Definitions successfully introduce names of abstract objects, names for abstract objects of some particular sort. The apparent implausibility of the claim that there are day-persons does not, however, constitute an argument against the view there are. One might want to suggest, for example, that names of day-persons are not names of some kind of object which is *sui generis* but are,

---

[9] I should say that there are certain tensions in Wright's position which speak against this interpretation. See, for example, his remark, on p. 9, that "it is, admittedly, obscure why any philosopher might endorse a general policy of contextual definition". The force of the remark is not, however, clear, and I do not, in any event, know how to interpret Wright so as to avoid committing him to this claim.

One possibility would be to take his views on the Caesar problem as relieving him of a commitment to this claim, so that, in *some* cases, we do not need to take the names introduced to refer to *abstract* objects but may take them to refer to concrete objects or, perhaps, to sets, rather than to objects of some new, distinct type. But, as I shall remark later, I do not know how to apply Wright's views on this subject. See Wright, §xiv, esp. pp. 116-17.

[10] It is probably best to say 'true simple sentences' here, since the occurrence of names of day-persons in e.g. "dap a does not exist" will not entail the existence of day-persons.

rather, names of *equivalence classes*, i.e., of sets; hence, there are such objects as day-persons, since these sets do exist.

In order to sustain this view, we should need some principled way to distinguish between the sorts of cases in which names of abstract objects, such as day-persons, may be taken to be names of equivalence classes and the sorts of cases in which they must be taken to be names of objects which are *sui generis*. One might hold, say, that names of words and letters refer to words and letters, not sets, but that names of day-persons refer to sets, not to day-persons. But this view is but a re-formulation of the view for which I shall be arguing. My view is that there are no day-persons, by which I mean that there are no objects, which have the identity-conditions stipulated in the Contextual Definition and which are related to persons as directions are related to lines. I mean to deny that there are any day-persons, by which I mean to deny that there are any objects which are *sui generis* and which we may take to be the referents of the names introduced by the Contextual Definition. My view thus differs from that just stated only in emphasis: For to say that names of day-persons name, not day-persons, but sets is to say that there are no day-persons, in the sense in which day-persons are of a kind unto themselves.

Thus, if we distinguish among names of abstract objects whose referents are *sui generis* and those which name equivalence classes, we are committed to holding that, say, day-persons might not exist. If we make no such distinction, we may hold that day-persons are *sui generis* and exist merely in virtue of the fact that 'ξ was born on the same day as η' is an equivalence relation. (For that will guarantee that 'dap a = dap a' is true, for each name 'a'.) I shall argue against this view in section three.

There is, however, another way of avoiding a distinction between names of abstract objects whose referents are *sui generis* and those whose referents are sets. One might hold, as Quine seems to hold, that *all* abstract objects are sets, that, once we have decided that we must allow reference to sets, we do not need to countenance reference to abstract objects of any other sort.[11] This Quinean view is thus a third alternative here: The other two alternatives are (what I call) *Naive* Platonism, the view that names of abstract objects (introduced by Contextual Definition) always refer and refer to objects which are *sui generis*; and the view for which I have said I shall argue, *Neutral* Platonism.[12]

---

[11] See here Chapter 7 of Quine's *Word and Object* (Cambridge MA: MIT Press, 1960).

[12] The former tag is meant to suggest a similarity to Moore's views. The latter is meant to suggest that this variety of Platonism is neutral on the question whether *there are* any abstract objects of any given type, that question being one whose decision depends upon the outcome of certain investigations which are the responsibility, not of Philosophy, but of Mathematics, Psychology, or whatever the relevant discipline.

1. That Names of Abstract Objects Are Ineliminable

Arguments for the ineliminability of names of Abstract Objects have classically focused upon the equivalence relation in terms of which the Contextual Definition is given. Conversely, work towards the elimination of such names has usually consisted in attempts to find a suitable equivalence relation. On the one hand, an appropriate equivalence relation is easy to define. In the case of works of literature, for example, one can appeal to the equivalence relation 'ξ is a copy of the same work of literature as η', thus allowing the following Definition:[13]

work a = work b $\equiv_{df}$ a is a copy of the same work of literature as b

Apparently, however, such a Definition is circular: Reference to works is not eliminated from the left-hand side, since reference is made to works on the right-hand side.

Such a Definition plainly preserves the meaning of the sentences which occur on the left-hand side. In general, however, it has proved difficult to find a Definition which was not circular in this way and which preserved meaning. Hence, debate often concerned just what sorts of conditions

---

[13] To avoid confusion, I shall refer to works of literature--in the sense in which there is only one *Macbeth*--as *works* and refer to the physical objects which are copies of them as *books*. I say that one book *copies* another if they are copies of the same work.

such a Definition should meet: Synonymy is too strong; material equivalence, too weak.

Such problems too proved intractable, and I do not intend to revive them here. I remind us of them only to emphasize that my argument *abstracts* from such issues: It does not turn on any such consideration, and any resolution of these problems is irrelevant to the fundamental difficulty facing any purported elimination of names of abstract objects.

If we are to avoid these classical problems, we must give an argument for the ineliminability of names of Abstract Objects which applies to *any* Contextual Definition; that is, an argument which could be formulated directly in terms of the general form of Contextual Definition. We must, that is, assume that the Contextual Definition offered meets any constraints upon such a Definition which could possibly be laid down and prescind from any considerations concerning the character of the equivalence relation or the possibility of finding predicates suitable to appear on the right-hand side.

We have yet to say anything about sentences which contain quantifiers which purport to range over abstract objects. Wright discusses some such quantifiers when he discusses Contextual Definition. Sentences which contain universal quantifiers, for example, may be treated via the following sort of schema:[14]

$$\forall x(Fx) \equiv_{df} \forall x(fx)$$

Existential quantification may be handled similarly:

$$\exists x(Fx) \equiv_{df} \exists x(fx)$$

With such Definitions of universal and existential quantification, along with the Definition of identity, one may define numerically definite and indefinite quantifiers in the usual way.

Our conception of works of literature as objects is connected with our understanding of the domain which they constitute, the domain over which we take our quantifiers to range.[15] But the character of this domain *as* a domain of works of literature can not be captured by our understanding of quantifiers such as the universal and existential ones: The truth-conditions of sentences in which they occur are not appropriately sensitive to the character of the domain. A quantifier which *is* so sensitive is "Most", and it is upon sentences containing quantifiers such as "Most", "Few", "At least two-thirds", and so forth, that I want to focus our attention.

---

[14] Wright, p. 30.

[15] Such a view is implicit in Frege and explicit, though in a much different form than here, in W.V.O. Quine. See his "On What There Is", in his *From a Logical Point of View,* 2nd ed. rev. (Cambridge MA: Harvard University Press, 1980), pp. 1-19.

I am going to assume, for the purposes of argument, that the language in which the Contextual Definition is being given contains *primitive* quantifiers of this sort, which range over the sort of object to which reference is made on the right-hand side of the Definition. So, for example, in the case of the Definition of names of works of literature, I am assuming that the language contains primitive quantifiers "Most books", "Few books", and so forth. The question is how we can define quantifiers such as "Most works" in terms of these ones.

Consider the sentence "Most works are long". The following analysis plainly will not do:[16]

Most works are long iff most books are long.

For most works might be long, though there are many more copies of short works than of long ones, so that most books are not long. What is required is that we select, for each work of literature, some *representative copy* of it, and formulate the right-hand side so that it says that most of *those* books are long. Where '$\xi C \eta$' is to be read 'ξ is a copy

---

[16] I am assuming here that it makes perfectly good sense to speak of a (physical) book as being long, as containing the word "dog", and so forth. I do think that this makes perfectly good sense. But, however that may be, the requirement that we prescind from problems attaching to the Definition of identity and of simple sentences such as "*Grundlagen* contains the word 'fantasy'" requires that we assume that there is some way to make sense of such talk.

of the same work as η′, the right-hand side may be taken to be:[17]

$$\exists\phi[\forall x\forall y(\phi x = \phi y \leftrightarrow xCy) \;\&\; (\text{Most } x)(\exists y(x = \phi y);\ \text{long}(x)]$$

The generalization to other sentences involving 'Most', and to sentences involving similar quantifiers, is obvious enough.

One worry about this sort of Definition might focus upon the use of second-order quantifiers. It is true enough that 'Most' is itself not a first-order quantifier. But one might be struck by the fact that we have assumed that primitive quantifiers such as "Most books" are already present in the language: One would not have thought that "Most works are long" was any *more* involved with second-order notions than is "Most books are long". But the argument need not rest upon such intuitions.

The point concerns, rather, the *character* of the second-order notions to which we have had to appeal. Consider the right-hand side of the Definition again:

---

[17] 'Most' is a binary quantifier. A sentence '(Most x)(Fx; Gx)' is to be read "Most Fs are Gs".

Equivalently, of course, the quantifier which ranges over first-order functions may be replaced by a second-order quantifier which ranges over (what Frege called) Concepts:

$$\exists\Phi[\forall x\exists y(\Phi y) \;\&\; \forall x\forall y(\Phi x \;\&\; \Phi y \rightarrow \neg xCy) \;\&\; (\text{Most } x)(\Phi x;\ \text{long}(x)]$$

Similarly, the second-order variable may be replaced by one which ranges over sets.

I know of no proof that there is no way to represent sentences like 'Most works are long' within an otherwise first-order language, whose quantifiers range over books, and which contains a primitive quantifier 'Most'.

$$\exists\phi(\forall x\forall y(\phi x = \phi y \leftrightarrow xCy) \,\&\, (\text{Most } x)(\exists y(x = \phi y); \text{long}(x))$$

We have here introduced a quantifier whose range is restricted to functions which satisfy the following conditions:

1. The function is defined for all books
2. The function does not distinguish copies of the same work
3. The function does distinguish copies of different works

We may simplify the Definition by introducing a fixed, primitive first-order function, call it '$W(\xi)$', which satisfies these three conditions. That is, we may introduce a first-order functional constant subject to the following axiom:[18]

$$\forall x\forall y(W(x) = W(y) \leftrightarrow xCy)$$

We may now re-write the right-hand side, in the general case, as follows:

$$(\text{Most } x)[\exists y(x = W(y)); fx]$$

One can not but notice the similarity between this analysis and this one:

$$(\text{Most } x)(\exists y(x = \text{work } y); Fx)$$

This last is the natural, Platonistic analysis of "Most works are long", since '$\exists y(\xi = \text{work } y)$' may be read '$\xi$ is a work'.

---

[18] I have omitted a constraint above and have not included it in this axiom: Namely, that the value of the function, for any given book, is some book which copies it. This constraint is not important for our purposes. Properly, however, the axiom should read: $\forall x\forall y([W(x) = W(y) \leftrightarrow xCy] \,\&\, xC\phi(x))$.

An understanding of the Contextual Definition of identity for a class of names issues in a conception of the *domain* over which quantifiers are intended to range, and it is upon that conception which our understanding of statements such as "Most works are long" draws. Our understanding of the Definition issues immediately in an understanding of these statements, no further explanation being required.[19] The *explanation* of our understanding of such sentences must proceed in terms of an explanation of our understanding of some such function as 'W(ξ)': Whatever else the Contextual Definition succeeds in doing, it *does* succeed in introducing such a function. (Indeed, this point, once stated, may seem obvious.)

Once appeal to such a function has become necessary, there can be no objection to its invocation in more familiar contexts. For example, we may give the following Contextual Definition of identity and of simple sentences:[20]

$$\text{work } a = \text{work } b \equiv_{df} W(a) = W(b)$$
$$F(\text{work } a) \equiv_{df} f(W(a))$$

This sort of Definition has quite definite advantages over Definitions like those we considered earlier. First, a

---

[19] One might wonder whether such a remark can possibly apply to Contextual Definitions, such as that of names of Numbers and Directions, which introduce names of infinitely many objects. 'Most' does not, of course, have any natural interpretation in such cases. This does not matter, however. In the case of Numbers, for example, we may focus upon such sentences as "Most numbers less that 12 are composite".

[20] Note, of course, that 'W(ξ)' is still subject to the non-logical axiom 'W(a) = W(b) iff aCb'.

Definition of this sort fully respects the apparent semantic structure of the sentences on the left-hand side. Secondly, the Definition makes the validity of intersubstitutivity depend, not upon some special feature of the sorts of predicates which occur on the right-hand side, but upon the fact that expressions which purport to refer to the same object *do* refer to the same object.

There is thus no possibility of eliminating the *functional expression* 'the work of which ξ is a copy' via such Contextual Definition. The right-hand side of these Definitions must make use of a functional expression with formal properties much like those of the functional expression which occurs on the left-hand side. The Definition therefore accomplishes nothing, if the goal was to eliminate *names* of abstract objects--expressions of the form 'work a'--occurrences of functional expressions such as "the word of which ξ is a token", and so on. The replacement of 'work(ξ)' by 'W(ξ)' hardly constitutes progress.

Nonetheless, it should be clear that this argument does *not* show that the notion of *reference* to abstract objects is not eliminable. For we saw above that, even should names of abstract objects prove ineliminable, there is an alternative course open to a Reductionist. The substitution of 'W(ξ)' for 'work(ξ)' should be seen, not as an attempt to eliminate an expression, but as reflecting a treatment of names of works as referring *to books*, rather than to works. That is to say: The substitution of '$W(\xi)$' for 'work(ξ)' suggests, in the syntax, an elimination of reference to abstract objects in the semantics.

The problem with which we are now left may seem familiar: For it is misleading to say that, on such a view, names of works do not refer to works. *Of course* names of works refer to works, works being whatever names of works name: The view is better explained as the view that works *are* books. The dispute concerns the range of the function 'work(ξ)': Another option would be to take the range to consist of equivalence classes, of sets of books. Our problem thus concerns, quite generally, when it is possible to identify objects, which purport to be of one sort, with objects of some other sort. Are works books? Are they sets? Are words sets? Are countries directions? Are people natural numbers?

This problem is Frege's famous Julius Caesar problem.[21] For the problem is how we are to decide questions about the identity and distinctness of objects of apparently different kinds: We may, borrowing a term from Michael Dummett, call this the problem of Trans-sortal Identification.

---

[21] See Frege, §§56, 66-8.

2. Reference, Semantics, and Model Theory

The argument given in the last section does not, as I have just emphasized, decide *to what* names, which purport to refer to abstract objects, in fact refer. After all, the argument is purely semantical in nature: It shows that we require to appeal to a function which restricts the domain to one of a certain *cardinality*. No purely semantical argument can show more, for no such argument can decide what *constitutes* the domain. So far as concerns this argument, the referents of names 'of works' might be books, sets, works, or candlesticks, so long as we can find a way to fix the cardinality of the domain they constitute.[22]

The view that we can 'take' the references of names of works to be sets is therefore entirely trivial (as, indeed, is the view that we can 'take' the references of these names to be works) unless there are some sorts of constraints upon what we may take the references of names of a given type to be. That is to say, the view is trivial unless the notion of *reference* is distinguished from the notion of the *value* of an expression in a model, for *of course* we can take the values of names of works, in some model, to be (almost) anything we like. By saying that such views are trivial,

---

[22] That there is no formal obstacle to doing so or to defining the various predicates appropriately is part of the point of Hilary Putnam's papers "Models and Reality", in his *Realism and Reason* (Cambridge: Cambridge University Press, 1983), and "Model Theory and the 'Factuality' of Semantics", in *Reflections on Chomsky*, ed. A. George (Oxford: Blackwell, 1989), pp. 213-32.

I mean to be claiming that no such view poses a threat to any form of Platonism, unless it incorporates a distinction between reference and the value of an expression in a model: For a Platonist should not be taken to be arguing anything other than that we can, and do, refer to certain abstract objects--such as letters, works of literature, and numbers--in whatever sense we can, and do, refer to other sorts of objects, be these sets, persons, or electrons.

We do not, therefore, need to resolve the question *whether* it is possible to distinguish the notion of reference from that of value in a model in order to resolve the dispute with which we are presently concerned: The viability of Reductionism depends upon the validity of such a distinction. For suppose that there is no way to distinguish the two notions, that there are no constraints upon what we may take the referents of names 'of persons', 'of Numbers', or 'of letters' to be. Then there is no sense in which we can, or must, take names 'of letters' to refer to certain sets or inscriptions which accords any special place to *sets* or to *inscriptions*: In precisely the same sense, we may take names 'of letters' to refer to letters or to Numbers. And in the same sense, we may take names of people to refer to people or to Numbers.

It is worth pausing here to reflect upon the question which is now guiding our discussion and the argumentative strategy which I shall deploy in answering it. The question is what, if anything, constrains what we may take the references of names of a given kind to be. Intuitively, names of people refer to people; names of rivers, to rivers; names of cities, to cities; and so forth. Either these intuitions are substantive or they are not. Any view which entails they are not, that there is no 'deep' sense in which names of people refer to people, rather than to objects of some other kind, I shall call 'deflationary'. On such a view, either there is *no* sense in which names of people refer to people, rather than, say, places; or, while there is a sense in which names of people refer only to people, this fact is a trivial consequence of some philosophical thesis concerning the notion of reference.

The view we discussed above, that there is no distinction between the notion of reference and the notion of interpretation in a model, is an instance of the former sort of deflationary view. My argument concerning it is that it will not serve the purposes of a Reductionist. My view about other deflationary strategies is similar. Consider, for example, Putnam's reply to this problem: Names of people do refer to people, and refer to the very people to whom we think they refer; but this is merely because "we don't intend" names of people to refer to anything but people.[23] And if that is why names of people refer to people, surely we may also say that we do not intend names of letters to

---

[23] See his "Models and Reality", p. 24.

refer to anything other than letters. We intend names of letters to refer neither to people, nor to inscriptions, nor, for that matter, to sets.[24]

These remarks about Putnam's solution also indicate the second strand of the argumentative strategy I am employing. It is my view that we can say rather more about what fixes the sort of object to which names in a given class refer: As I said, that does not need to be established here, since Reductionism is committed to this view. As deflationary as Putnam's views are, he does say something in response to our leading question: And what I argued was that, in so far as his response works, it distinguishes reference to *abstract* objects both from reference to concrete objects and from reference to abstract objects of *prima facie* distinct kinds. In general, then, the claim is that any way of distinguishing the kinds of objects to which, say, names of people, of rivers, of places, and so forth, refer will *also* distinguish the kinds of objects to which names of Numbers, of letters, of works of literature, and so forth, refer; it will distinguish the sort of object to which, say, names of letters refer both from the sorts of objects to which names of concrete objects refer and from the sorts of objects to which names of other sorts of abstract objects refer.

---

[24] Similar remarks also apply to views, such as that of Davidson, which hold that the notion of reference is purely theoretical. See his "Reality Without Reference", *Dialectica* 31 (1977), pp. 247-58.

One might want to suggest that we can not say, so to speak, piecemeal what constrains the sort of object to which terms in a given class refer; that the only constraints which we may place on the references of names of some type are *universal*, in the sense that the constraint applies simultaneously to each class of names. For example, one might hold that the semantic theory as a whole must provide the simplest possible account of the truth-conditions of the sentences of some given language and that this constrains the sort of object to which a name may refer. This kind of view may well be correct: But, again, it would not appear to be available to a Reductionist. If there is one such theory, there are many, which can be derived from it by replacing reference to objects of one sort with reference to objects of some other sort.[25] Of course, this may well be accepted, if not stressed, by one who holds a position of this kind: It may be said that the appropriate conclusion is that reference is inscrutable, that ontology is relative. But, if the references of names of abstract objects are inscrutable, that is, for present purposes, fine with me, so long as it is recognized that names of abstract objects are not, in this regard, any worse off than are names of concrete objects.[26]

---

[25] Putnam's remarks also apply at this sort of point. See his "Model Theory and the 'Factuality' of Semantics", again.

[26] Compare Wright's discussion of Benacerraf, in §xv.

Two other points should be made about this Holistic view, however. First, that the theory should be as 'simple' as possible can not be taken to entail that the ontology of the theory should be as parsimonious as possible. Whatever the virtues of parsimony elsewhere (say, in physics), I for one know of no very good, independent argument for the virtues of parsimony in semantics. To invoke such a conception would be to beg the question against the Platonist, who is not obviously party to such a view. Secondly, the discussion in this paper might well be read as a discussion of the relative virtues of a theory which appeals to a notion of reference to abstract objects (in addition to reference to sets) and a theory which does not. The former may well give us the simpler, though not the more parsimonious, theory.

The point of the discussion so far should be clear enough: A Reductionist must hold that we may take the referents of names of abstract objects to be, not abstract objects, but objects of some other sort, in a sense in which we may *not* take names of, say, persons to refer to Numbers. Any such view is committed to demonstrating that there is some kind of *special* problem for the Platonist here: And, to show this, one must distinguish the notion of reference from that of value in a model. One must, that is, show that there are certain constraints upon what we may take the referents of names of a certain kind to be and show that we may, *nonetheless*, take names of abstract objects to refer either to sets or to representatives of those sets.

A Reductionist of the sort I have labelled an Epistemological Reductionist might reply here that I have mischaracterized her position. Her view is that abstract objects pose a special epistemological problem, that no account is possible of how we can have knowledge about abstract objects, even if there are any such things. The goal is not to show why we must do away with abstract objects: Rather, the goal is to show how we can *make sense* of talk of words or Numbers, without supposing ourselves to refer to abstract objects. The claim is that we *may* take names of abstract objects to refer to concrete objects or to sets, and that we may not take names of, say, people to refer to, say, works, because of the epistemological difficulties such an identification would pose. Nor does taking the referents of names of persons to be people raise comparable problems.[27]

I do not mean to ignore this sort of view: But I want to consider it as a view which is committed to a particular sort of *solution* to the problem we are discussing. Quite generally, one may say that such a view holds that we may not take names in a given class to refer to objects of a given kind unless it is possible to give some account of how

---

[27] I should thank Bob Stalnaker for emphasizing and helping me to understand this reply.

we may have knowledge of such objects. Presumably, we can all agree with that claim: But one who is attracted to Epistemological Reductionism usually is so attracted because she has quite specific epistemological views. In particular, many of those who have argued that there is no coherent account of how we could have knowledge of Numbers, or of words, or of works of literature have so argued on the basis of a *causal* theory of knowledge, which has gone hand-in-hand with a causal theory of reference.

In *this* context, however, the causal theory, whether of knowledge or of reference, suffers, if not from circularity, then from specificity--if, I emphasize, it is intended to function as an objection to the *possibility* of supplying an acceptable epistemology or theory of reference where abstract objects are concerned. The causal theory of reference, for instance, has no antecedent claim to be a theory of reference at all: The theory is developed, motivated, explained, and defended wholly in terms of examples which fundamentally concern concrete objects. If abstract objects do not cause anything, then, in so far as the sorts of examples which motivate the causal theory can be formulated at all, no *causal* theory is going to resolve them; and, if such examples can not be formulated, neither the causal nor any other such theory will be required to resolve them. It is a fallacy of hasty generalization to develop a theory on the basis of examples which concern objects which unproblematically *do* have causal powers,[28] to generalize the theory to one about reference *in general*, and then to argue that this general theory raises insuperable problems for the view that we can refer to objects which do not have causal powers.

Similar remarks apply to the causal theory of knowledge.[29]

Moreover, not all abstract objects are, in the required sense, causally inert. Consider, for example, the statement, "John believes that p because he read the *Grundlagen*". John's reading of the *Grundlagen* is an event in which the *Grundlagen* figures, and this event causes something. It might be suggested that John *really* believes that p because he read a certain copy of the *Grundlagen*, because he had a causal interaction with some physical book. But that may not be true, unless the 'book' is spatially discontinuous, the mereological fusion of parts of different copies of the *Grundlagen*. And furthermore, the claim that John's belief was caused by a specific copy of the *Grundlagen* is far too

---

[28] I do not, of course, hereby commit myself to the intelligibility of any such use of the notion of an object's having causal powers.

[29] For more specific objections, see Wright, §xi-xii and Hale, Chs. 4, 6. I do not mean to be cavalier about the problem of providing a coherent epistemology for mathematics or for any other sort of talk about abstract objects. Nor do I mean to be cavalier about the problem of explaining what *does* fix the references of names of abstract objects. The former problem I am not going to be able to discuss here, however. The latter problem just is our topic.

specific: For he would have had the same belief no matter which copy of the *Grundlagen* he had read.

Moreover, the causal theorist ought not to get carried away with such manoeuvers, unless she is to find herself committed to the claim that reference to people is impossible, on the grounds that no-one ever interacts causally with a person, but only with some person-stage, or even only with some clump of matter at some time.[30] It is natural to take such objects as persons to be among the objects which may figure in the sorts of events which are causes of, among other things, our beliefs. We naturally take our beliefs about, say, John Doe to be caused by events in which John Doe figures, because they are sensitive to how things stand with *John Doe*, with that particular person or organism, not one part or one stage of him. But if that is the sort of thing we must say if we are to allow persons to be the references of names of them, if the causal theory of knowledge or of reference is to be compatible with the claim that we have knowledge about and refer to persons, then it is clear enough how a corresponding story could be told about the *Grundlagen*: Our beliefs about *it* are sensitive, not to

---

[30] We shall return to this view at the end of the paper. For the moment, I am assuming that we do require an account of what distinguishes the sort of object to which names of people refer from the sort of object to which names of rivers refer.

how things stand with some copy of it, but to what is said in *any* copy of it.[31]

These last remarks may also be directed against a quite different deployment of causal notions. One might suggest that, by appeal to causality, we may fix the sorts of objects to which names of concrete objects refer. If so, and if no such answer can be given in the case of names of abstract objects, then we have thereby established the right sort of difference between names of concrete objects and names of abstract objects: Reference, in the latter case, is free-floating in a way in which reference is *not* free-floating in the case of names of concrete objects. In the case of names of abstract objects, one might then say, we can take their references to be of any sort we like, so long as we get the cardinality of the domain right and suitably define the relevant predicates: But we can not similarly take the referents of names of concrete objects to be anything we like.

My argument against such a view does not depend upon any claim to the effect that the appeal to causality will not work in the case of names of concrete objects. What I have argued are two points which it is worth re-emphasizing here. First, it is important to state such a view, as it has been stated here, as incorporating a claim that, *if* no

---

[31] Sylvain Bromberger defends just such a view about linguistic types in his "Types and Tokens in Linguistics", in A. George, ed., pp. 58-89.

answer can be given to the question what fixes the sort of object to which names of abstract objects refer, then we require no notion of reference to them. The supposed fact that abstract objects do not have 'causal powers' can not, for the reasons given above, constitute reason to think no such answer can be given. And secondly, there is some initial reason to think that the kind of answer offered, by our new causal theorist, in the case of concrete objects can be seen as a special case of a more general answer to our question, which would also fix the sorts of objects to which names of abstract objects refer: This is part of the point of the discussion of the causes of beliefs such as those caused by my reading of the *Grundlagen*.

Lacking any detailed proposal for a causal account of what fixes the sort of object to which various names of concrete objects refer, we can not show in any detail that such a view must allow generalization to an account of what fixes the sort of object to which names of abstract objects refer. Moreover, we can not show, in any detail, that any such generalization must entail that names of (what seem to be) abstract objects are not names of concrete objects. Instead, I shall be suggesting a particular account of what does fix the sort of object to which names in a given class refer: Plausible causal accounts are compatible with this more general view.

3. The Existence of Abstract Objects

Before embarking on that project, however, we need to lay the foundations for the discussion. In the course of doing so, I shall present an argument against the view that, for any given Contextual Definition, the names introduced by means of it refer and refer to objects which are *sui generis*. My view is thus that, while *there are* abstract objects of some sorts, there *are not* abstract objects of other sorts.

Consider again the Contextual Definition of names of day-persons:

$$\text{dap } a = \text{dap } b \equiv_{df} a \text{ was born on the same day as } b$$
$$F(\text{dap } a) \equiv_{df} fa$$

What sorts of things may be said about day-persons? Well, one suitable predicate '$f\xi$' is 'all persons born on the same day as ξ have red hair': So the sentence 'Red-haired(dap **a**)' is defined in terms of 'All persons born on the same day as **a** have red hair'. We may form arbitrarily many similar predicates, not only by means of the (restricted) quantifier 'all persons born on the same day as ξ', but by making use of such (restricted) quantifiers as 'some person born on the same day as ξ', 'most persons born on the day before ξ', and so forth.

Now, one might want to say that 'Red-haired(ξ)' is rather peculiar, that, intuitively, it does not express a *property* of day-persons (whatever 'day-persons' may be). And there is a distinction between sorts of predicates which we use with various names which is important here.[32] Consider, for instance, the sentences:

The father of John is six feet tall.
The father of John has only blue-eyed children.

Intuitively, the former sentence attributes a *property* of persons to the father of John; the latter does not. My intuitions about the following sentences are similar:

The work of which **a** is a copy contains the word 'dog'.
The work of which **a** is a copy has only copies which have a torn page.

Again, it is tempting to say that, while the former sentence attributes a *property* of works of literature to a work, the latter does not.

This distinction needs to be made more precise: The notion of a property can serve no more than a heuristic purpose here.

One feature which distinguishes predicates which occur in sentences of the former sort from those which occur in sentences of the latter sort, is that, to determine whether a sentence like "The work of which **a** is a copy contains the word 'dog'" is true, one need only look at any given copy of the work in question: One need only look at **a** itself or at any book which copies **a**. On the other hand, to determine whether a sentence like "The work of which **a** is a copy has only copies which have a torn page" is true, one, ordinarily, must know whether, and if so which, other books copy

[32] I have discussed this distinction in more detail elsewhere. See my "Whether Structure May Be Misleading".

a. And, again, in order to know whether the former sentence is true, one does *not*, ordinarily, need to know whether any other books copy **a** or, if so, which books copy it.

By saying that one must *ordinarily* know whether, and if so which, other books copy **a** to determine whether each copy of it has a torn page, I mean to recognize that, one may, on any given occasion, be able to determine that not all copies of some work have a torn page without knowing whether any other books copy the given copy **a**--if **a** itself has no torn page--or without knowing which other books copy **a**--if, say, **b** copies **a** and has no torn page. In general, one's ability to determine whether each copy of **a** has a torn page depends upon one's ability to determine which books copy **a**. There is a generally (or universally) applicable procedure for determining whether each copy of **a** has a torn page, and the application of this procedure requires one to determine which other books copy **a**. On the other hand, while there is such a procedure for determining whether a given work contains the word "dog", there is *also* a procedure which does not require one to determine which other books copy **a**.[33] It is the existence of such a generally applicable procedure in

[33] There is an obvious connection here with Dummett's distinction between 'direct' and 'indirect' means of verification. See his "What Is a Theory of Meaning? (II)", in Gareth Evans and John McDowell, eds., *Truth and Meaning* (Oxford: Oxford University Press, 1976), pp. 67-137, at pp. 115ff.

the one case, though not in the other, which distinguishes the two sorts of predicates.

With the use of the word 'ordinarily' explained in this way, we may continue to use it as above.

Let us call those predicates like "contains the word 'dog'" *variant* predicates; other predicates, like "has only copies which have a torn page", *constant* predicates.[34] A constant predicate (of works) is, thus, a predicate which is such that, in order to determine whether a given work falls under it (i.e., in order to know whether a given sentence of the form "C(work a)" is true), one must ordinarily know whether, and if so which, other books copy some given book. A variant predicate, on the other hand, is a predicate which is such that one may determine whether a given work falls under it *without* knowing whether, and if so which, other books copy some given copy of it.

The distinction between variant and constant predicates is closely connected to what one might have thought was the *point* of our speaking of works of literature. There is much which can be said about books: That they are dirty, that they have some mass, that they contain some word, and so

---

[34] The point of the terminology is that the (generally applicable) means for determining the truth-value of a sentence 'V(work a)', which contains a variant predicate, will vary significantly as 'a' is replaced by names of other books; the (most obvious) means for determining the truth-value of a sentence 'C(work a)', which contains a constant predicate, will, on the other hand, remain constant as 'a' is replaced by names of other books.

forth. But we distinguish what we can discuss, verify, investigate, and question without concerning ourselves with which copy of a given work happens to be to hand at a given time. Our use of names of works would be largely without point if there we not certain predicates of books such as those which play this special role in our discourse about works: Namely, those whose satisfaction by some book implies (in some sense) its satisfaction by any other book which copies it, and whose satisfaction, by a given book, may be determined in the absence of a knowledge of which, if any, other books copy it.

I shall, borrowing the term from the philosophy of science, say that a predicate *Projects* over the R-equivalence classes, if the satisfaction of the predicate by *x* implies that, for each *y* such that *x*R*y*, *y* also satisfies the predicate.[35] (Explaining the sense in which the word 'implies' is used here is a large part of the goal of our discussion.)

On this analysis, the intuition that 'Red-haired(t)' does not express a *property* of day-persons derives from the fact that 'Red-haired(t)' is a *constant* predicate of day-persons: For, in order to know whether a sentence containing 'Red-haired(t)' is true, one must know whether, and if so

[35] I shall capitalize "Projects" to remind the reader of the fact that I am not necessarily, nor am I claiming to be, using the term in the standard way. Plainly, there are similarities which give the use of this term here a point.

which, other persons were born on the same day as some given person; to determine whether 'Red-haired(dap(Vanessa Redgrave))' is true, one must know whether, and if so which, other persons were born on the same day as Vanessa Redgrave. The predicate of *persons* 'red-haired(ξ)' does not Project over the relevant equivalence classes.

The distinction between variant and constant predicates has here been explained only for the case of predicates introduced by Contextual Definition. (At best, the sort of explanation just given extends also to predicates as they occur in sentences of the form 'F(fnc a)'.) A full defense of the coherence and importance of this distinction would surely require a more general formulation, an explanation of the distinction as it applies to *any* predicate. Even if such an explanation is not now to hand, it is perhaps worth noting some reasons to be optimistic about the prospects for its provision.

If a copies b, then the terms 'work a' and 'work b' are names of the same work: That is, we may, by using one or the other, refer to the same object. Let us say that to refer to a work by means of such an expression is to refer to it *basically*. To say that a predicate is variant is therefore to say that, to determine whether a given work, referred to basically, satisfies it, one need not know how else one may refer to it basically. With respect to variant predicates, each basic means of reference has a kind of *autonomy*: The ability to determine the truth-value of a sentence containing a variant predicate and a term which refers basically to an object does not depend upon an ability to determine whether it is possible to refer to that same object in any other way.[36] It is for this reason that Leibniz's Law can be used to extend our knowledge: For acquisition of the knowledge that a is F need not depend upon one's knowing how else one may refer to the a.[37]

We may put the point slightly differently. The fact that one can determine whether a is F without knowing how else one may refer to a is what makes Frege's puzzle about the morning star and the evening star possible. For I can know that Hesperus is F without knowing that I may also refer to Hesperus as Phosphorus. To explain the distinction between variant and constant predicates is therefore precisely to explain what makes Frege's puzzle possible and to explain how the application of Leibniz's Law can extend our knowledge. And, conversely, to explain these things, one needs a distinction like that between variant and constant predicates: For if one must know in what ways one can refer

---

[36] Note that this remark is strictly correct only for one-place predicates. Identity is a good example of a two-place predicate for which such a characterization would need to be re-stated.

[37] Compare David Wiggins, who argues that a proper account of the criterion of identity for objects of a given kind should *entail* the truth of Leibniz's Law: See his *Sameness and Substance* (Oxford: Blackwell, 1980), pp. 48-53.

to a given object to know whether it satisfies a given predicate, then the application of Leibniz's Law, in the case of such a predicate, will not ordinarily extend our knowledge.

These remarks are clearly programmatic: Much more needs to be said to formulate a generalized distinction between variant and constant predicates. But I hope that enough has been said to motivate the distinction, to show that it is probably of some importance, and to show that the distinction, as drawn for cases of the sort we are discussing, is plausibly a special case of a more general distinction.

Now, my suggestion is that the intuition that *there do not exist* such objects as day-persons is closely related to the fact that *we do not understand many variant predicates of day-persons*, i.e., that we do not understand many[30] predicates which, intuitively, express 'properties' of day-persons. How might it have been otherwise? Some years ago, there was a fad about what were called 'bio-rhythms': There were supposed to be certain higher-level affective states which each person had--degrees of awareness, laxity, happiness, and so forth--and persons who were born on the same

---

[30] We do understand *some* such predicates, namely, those which are constructed from the very equivalence relation we used to define names of day-persons in the first place. Identity is itself a variant relation, and we can define variant predicates in terms of such predicates of persons as '...was born on 21 March 1939', and so forth. The important difference between such predicates and the variant predicates relevant to the question of existence should become clear during our discussion.

day were supposed to have the same such higher-level affective states. Persons were, that is, said to have the same *bio-rhythms* if they were born on the same day: And these bio-rhythms were themselves a measure, so to speak, of certain of one's affective states.

Idealizing, let us suppose that there is a specific theory, Bio-rhythm Theory, making more precise and enlarging upon this idea. This Theory might have been true. Had it been true, then there would have been a great many predicates of persons which Projected over the classes of persons born on the same day. That is to say, there should have been a great many predicates of day-persons which were *variant* predicates. Had Bio-rhythm Theory been true, then there would have been such objects as day-persons: What I am calling day-persons would just have been bio-rhythms. For a bio-rhythm would be an abstract object, a kind of structure of a person's affective states, which would be *shared* by persons born on the same day, just as parallel lines share a direction, as tokens which copy one another share a type.

But, supposing that Bio-rhythm Theory is *not* true, it seems to me that there are no bio-rhythms and that this judgment is here in accord with common sense. The sorts of objects to which speakers would have referred had Bio-rhythm Theory been true do not, since Bio-rhythm Theory is not true, exist. Platonism need not commit itself to the existence of such strange and wonderful objects: It is, in my opinion, better off without them.

One formulation of my proposal would be as follows: A given Contextual Definition successfully introduces names which refer to abstract objects of a given sort if, and only if, we understand a wide variety of variant predicates of those objects. But that can not be correct. The reason is that such a view would make the *existence* of objects of a given sort depend upon our *understanding* of predicates of a given kind, and Platonism need not gratuitously commit itself to Idealism. This formulation does not entail that, if we do not understand such predicates, then there are no such objects: For the claim is that the names introduced by the Contextual Definition will *refer* if, and only if, we understand a variety of variant predicates of those objects. The objects may, for all this view says, *exist* even if we can not refer to them, due to lack of the appropriate understanding of sentences containing what might otherwise be names of them.

The problem, rather, is that we *do*, or at least we seem to, understand a variety of variant predicates of bio-rhythms: Less strongly, some people do (or did) understand such predicates, though they are (or were) not able to refer to bio-rhythms. Our description of how the world might have been, if Bio-rhythm Theory were true, is *eo ipso* a description of how those who believe that Bio-rhythm Theory *is* true believe the world *is*. A description of how we should use certain predicates, of what aspects of that use would correspond to their being variant predicates of day-persons, if Bio-rhythm Theory were true, is also a description of how those who believe that Bio-rhythm Theory *is* true *do* use those predicates.

That is: Those who believe that Bio-rhythm Theory is true, understand (or, less strongly, use) a variety of predicates of day-persons as variant predicates of day-persons. Similarly, I may presumably come to *understand* what they say about day-persons, about Bio-rhythms, even though I do not believe Bio-rhythm Theory to be true: Hence, I can come to understand certain predicates *as* variant predicates of day-persons, though I do not believe myself to refer to bio-rhythms and though I do not refer to bio-rhythms, there being no such objects. How else, one might ask, can I intelligibly debate or investigate whether there are any such objects? How else, for that matter, can I intelligibly *deny* that there are any bio-rhythms?[39]

---

[39] Of course, one might suggest that *none* of us understand the predicates we take ourselves to understand, that, as Evans suggested that one does not understand a proper name unless it refers, one can understand neither variant predicates nor names of day-persons unless day-persons exist. I expect that Tom Kuhn would want to caution against similar remarks about certain kinds of examples, though *not* necessarily about *this* example.

This sort of dispute is not directly relevant here, however, since my view about existence is, so far as I can tell, compatible with such views. In the Evans-style case, we need to give some account of the sense of negative existentials, anyway; and, in the Kuhn-style case, we need an

We must, therefore, distinguish the question whether one understands a predicate as a variant predicate of day-persons from the question whether, *in fact*, the predicate of persons in terms of which it is defined Projects over the classes of persons born on the same day.[40] Those who believe that Bio-rhythm Theory is true *believe*, say, that the predicate "ξ is lethargic" Projects over classes of persons born on the same day, and it is in their use of the corresponding predicate of day-persons in accord with this belief that their understanding of the predicate as a variant predicate is manifested. That is: The understanding of the predicate (of day-persons) 'Lethargic(ξ)' *as* a variant predicate of day-persons partly consists in the knowledge that a sentence of the form 'Lethargic(dap a)' is true *only if* the predicate of persons 'ξ is lethargic' Projects over classes of persons born on the same day.

---

account of what it would be to understand such a language (or theory) and a demonstration that such understanding is, in a certain sense, incompatible with our understanding of our own language. As I understand his current work, appeal to something like a class of variant predicates is essential to this project.

[40] I think that, properly, we should say that we must distinguish the question whether we understand the predicate, as a variant predicate, from the question whether the predicate *refers*. That is: One might deny that the variant predicates of day-persons we do understand refer to any Concept (in the Fregean sense) at all. However, there is much confusion about Fregean Concepts and reference to them; hence, I shall avoid such language here.

Conversely, even if 'lethargic(ξ)' does project over classes of persons born on the same day and even if **a** is lethargic, the sentence 'Lethargic(dap **a**)' may not be true. For the truth of 'Lethargic(dap **a**)' requires that dap **a** *exist* and therefore that day-persons, or bio-rhythms, exist. Hence, the truth of a sentence such as 'Lethargic(dap **a**)' requires not only that a certain predicate of persons, in this case, 'lethargic(ξ)', project but also that there be a *variety* of predicates of persons which project: For only if a variety of predicates of persons project, only if (something like) Bio-rhythm Theory is true, are there any bio-rhythms at all.

4. The Notion of an Ideology

The view being developed here is best explained in terms of the notion of an *Ideology*. An Ideology about objects of a certain kind is not a specific theory about those objects: Rather, to understand the Ideology associated with objects of a given kind is to understand what sorts of properties such objects typically have, to understand certain predicates of those objects as variant predicates.[41] In the case of names of day-persons, the asso-

---

[41] My use of the term 'Ideology' should echo the distinction between ideology and ontology which one finds, for example, in Quine. The ideology of a theory is, for Quine, a matter of what the admissible or primitive predicates of the theory are. This is plainly related to my use of the term, but is just as plainly different. See W.V. Quine, "Ontological Reduction and the Theory of Numbers", in *The Ways of*

ciated Ideology is not any specific theory of bio-rhythms; it is, rather, a *part* of such a theory, the part which states that certain, or some kinds of, predicates of persons project over classes of persons born on the same day.

My view is thus, first, that an *understanding* of names of abstract objects of a given sort depends upon an understanding of the associated Ideology. I shall not argue for this view in any detail here:[42] A full argument for it would require us to show, first, that to understand any given name it is necessary to understand a *criterion of identity* for that name; and, secondly, that, to grasp the criterion of identity for the name, one must understand some associated Ideology. The discussion in the next few sections bears directly upon this problem: For the discussion concerns what fixes the *kind* of object to which a term refers. Unfortunately, however, that discussion shows only that an appeal to the notion of an Ideology can solve this problem, not that appeal to the notion is required.

It is *prima facie* plausible that one who does not understand the associated Ideology does not understand names

---

Paradox (New York: Random House, 1966), pp. 199-207, at p. 202, and "The Scope and Language of Science", in the same volume, pp. 215-32, at p. 232.

[42] For a defense of a view with a great deal of similarity to this view, see David Wiggins's *Sameness and Substance*. The similarity comes out in such principles as his D(v): "f is a substance concept only if f determines either a principle of *activity*, a principle of *functioning* or a principle of *operation* for members of its extension".

of day-persons. One who knows only that 'Lethargic(dap a)' was 'defined' in terms of the sentence '∀x(x was born on the same day as a → lethargic(x)', who does not know that the former sentence is true only if Bio-rhythm Theory is true and if 'lethargic(ξ)' Projects will be quite unable to understand 'Lethargic(dap a)'. For such a person will take the accidental lethargy of all persons born on the same day as a to establish the truth of this sentence, though a justification of that sort would be rejected by those who speak as I have supposed them to speak.

It is also worth noting a further explanatory consequence of the view that the understanding of names depends upon an understanding of the associated Ideology. It is no accident that, upon first encountering the Contextual Definition of names of day-persons, one may have the sense that one has not the slightest idea what sorts of objects these are meant to be (if not just equivalence classes): But, upon explanation of the associated Ideology, one immediately has a much better idea what sorts of objects are in question. And, indeed, we can imagine quite different Ideologies which might be associated with names introduced by Contextual Definition otherwise just like that of names of day-persons: Some people might believe that other states of mind--say, certain sorts of beliefs--do not vary among persons born on the same day; or, that certain physiological properties do not vary; or that certain gross anatomical features do not vary. Our conception of what sort of object a day-person is varies as we vary the associated Ideology.

It is easy to overlook the presence of the Ideology entirely. Consider, for example, Frege's definition of names of what he calls 'orientations'. The orientation of a plane **a** is the same as the orientation of a plane **b** if, and only if, the planes are parallel.[43] It is an interesting fact that, immediately upon encountering this definition, one immediately has the sense that one knows precisely what Frege means to be talking about when he speaks of orientations. Why does it seem so obvious what sort of object an orientation is? as contrasted with our utter failure to discern a conception of 'bio-rhythms' in the mere Contextual Definition of names of day-persons? The reason is that, in the former case, it is obvious what the Ideology is intended to be. Orientations are *geometrical* objects; the associated Ideology, as is clear from the context of Frege's discussion, is geometrical: The theory of orientations is to be a geometrical theory.[44]

The second component of my view is that the existence of abstract objects of a given sort depends upon the *truth*

---

[43] See Frege, *Foundations*, §64. The remarks made here were suggested to me by discussions with George Boolos.

[44] The importance of this Ideology can, again, be seen by imagining variations. One would have been surprised, at the very least, if Frege had gone on to explain that the study of orientations is the responsibility of physics, certain distributions of physical particles being invariant among parallel planes.

of some Theory which incorporates the associated Ideology. Part of the reason our understanding of the associated Ideology is of importance to our understanding of names of abstract objects is that it is essential to our understanding of in what the existence of such objects consists. One way to put this point is as follows:[45] Among the sentences of which we, as theorists, want to give an account are certain negative existential statements, such as "There are no bio-rhythms". Surely such a statement is intelligible and (at least my) intuition tells me it is true (given that Bio-rhythm Theory is not true). If the sentence is to be capable of being true, there can be no guarantee that, given that names of abstract objects are introduced by Contextual Definition, there are such objects; and, moreover, we need some account of the senses of statements asserting and denying the existence of such objects.

I expect that many will have wanted to object that there is no need to look to *semantics* for a resolution of problems of this kind: Pragmatics might resolve them. If Bio-rhythm Theory is not true, then, while it would be true to say that dap a exists, it might be misleading to say so (perhaps because this assertion implicates the truth of Bio-rhythm Theory or some variant of it). Moreover, if the predicate 'lethargic(ξ)' does not project, then, while the

[45] This way of putting the point was suggested to me by Bob Stalnaker.

assertion 'Lethargic(dap a)' might be true, to say so would again be misleading (perhaps because it implicates the Projectibility of 'lethargic(ξ)'). Such a view may, indeed, be favored, not by one who hopes to defend the view that day-persons exist and are *sui generis*, but by one who, instead, maintains that day-persons exist, but are mere equivalence classes.

I am not going to pursue such a proposal in any detail. It is worth noting, however, that any such proposal is either revisionary of ordinary linguistic practices. The crucial sentences are again the negative existentials. It seems natural and true to say that there are no bio-rhythms: And, if we are to pursue the pragmatic course, this assertion is simply false. (This kind of revisionism is rather more amenable to the sort of Reductionist who would identify day-persons as sets than to a Platonist.)

Naive Platonism may take two forms here. I have so far been speaking of it as embodying the claim that the objects about which we would speak, were Bio-rhythm Theory true, are just the same objects about which we in fact speak, though it is not. It is this view against which I have so far argued: Such a view fails to take seriously the connection between existence and Ideology. However, there is an alternative view, that day-persons are *not* bio-rhythms, that bio-rhythms do not exist, but that day-persons do and are *sui generis*. On this view, were Bio-rhythm Theory true, we should be able to refer *both* to bio-rhythms and to day-persons. Surely there is a parallel between the case of bio-rhythms and day-persons and, for example, the case of word-types: That is, there must also be, so to speak, 'purely abstract' word-types as well as ordinary ones. It is difficult to believe that such duplication of objects can serve any purpose: And moreover, it would seem that the only abstract objects in which we have any interest are precisely *not* the purely abstract ones. As soon as we have anything interesting to say about objects of a certain kind, this constitutes our possession of an Ideology about them; and, at that point, we are no longer talking about the purely abstract objects.

Matters stand quite differently if we are concerned with a reply offered by a Reductionist. But I am not going to rest my case against Reductionism upon the nature of the distinction between semantics and pragmatics: Rather, we need now to turn to a more direct argument that bio-rhythms are neither sets nor persons, but must be construed, given the associated Ideology, as *sui generis*.

5. Trans-sortal Identification

The question which we left unanswered earlier is: What constrains the sort of object to which we may take the names in a given class to refer? What, for example, constrains the sort of object to which we may take names of persons to refer? It is intuitively obvious that names of persons do not refer to celestial bodies: It is not an intelligible possibility that the name "George Bush" (as, of course, we now use it) refers to a planet. Nor, for that matter, is it an intelligible possibility that the name "Julius Caesar" refers to a natural number.[46] But why not?

For the moment, let us focus our attention on functional expressions, such as "the father of ξ", whose range consists of concrete objects. Now, it seems obvious enough that the father of John is a person: And one might suppose that the fact that John and Jane have the same father if, and only if, the same male human is immediately causally implicated in their creation is what determines that 'the father of John' refers to a person.[47] Surely, it is of great importance that 'the father of **a** = the father of **b**' is true if, and only if, the same male human is immediately causally implicated in the creation of **a** and of **a**. But this does not

---

[46] There is a peculiar difference here between the statements "Julius Caesar is (or, might have been) the number 0" and "The number 0 is (or, might have been) Julius Caesar". The former seems, at least to me, to say that Caesar is (might have been) an abstract object--in particular, a number. The latter, on the other hand, seems to say that zero is (might have been) a concrete object, a person.

I have no idea what the significance of this point should be taken to be.

[47] I am abstracting here from the fact that 'the father of ξ' is used with names of other sorts of animals: The discussion could be rephrased in such terms.

entail that 'the father of John' refers to a person. For consider the following expressions:[40]

the set of all persons who have the same father as John
the oldest paternal half-sibling of John
the singleton of the father of John
the (current) location of the oldest paternal half-sibling of John

Each of these expressions has the same *weak identity-conditions* as 'the father of John': That is, the reference of any one of these expressions will remain unchanged if, and only if, the same male human is immediately causally implicated in the creation of John and any person whose name is substituted for his. But not all of these expressions refer to objects of the same sort, and those which do do not refer to the same object.

A similar point applies to names of directions. Consider the following expressions:

the direction of λ
the line through the Origin parallel to λ
the set of lines parallel to λ
the angle at which λ intersects the x-axis

---

[40] A point not unlike this one was, I am told, made in a lecture by Michael Dummett. Dummett remarked that if (what I am calling) the weak identity-conditions determine the sort of object to which a name refers, then it is philosophically confused to think that the eccentricity of an ellipse is a real number.

Warren Goldfarb gave another good example: Let us say that the architect of x = the architect of y if, and only if, x and y are buildings designed by the same person. Surely, we may take 'the architect of the John Hancock Tower' to refer to I.M. Pei--a person. But, of course, there is a set of buildings designed by Pei, a first building designed by Pei, and so forth; and we *might* think, rightly or wrongly, that there are *distinguishing features* which each building has, which it shares with all and only those buildings designed by the same person.

Again, substitution, for λ, of the name of any line parallel to λ will leave the referent of each of these expressions unchanged; and *only* the substitution of names of lines parallel to λ will do so. Nonetheless, not all of these expressions refer to objects of the same sort: One refers to a line; one, to a set; one, to a real number.

This point can be stated quite precisely. Let '$\phi(\xi)$' be a function from objects of sort S to objects of any sort T (not necessarily different from S). Then '$\phi(\xi)$' induces an equivalence relation '$\xi\Phi\eta$' on objects of sort S, which we define as follows:

$$\forall x \forall y (x \Phi y \leftrightarrow \phi(x) = \phi(y))$$

Distinct functions from S to T induce the same equivalence relation, and various functions from S to sorts T′ (distinct from T) also induce the same equivalence relation. There are thus many distinct functions whose domain is objects of sort S and which have the same weak identity-conditions.

Now, it might be said that what fixes the reference of an expression like 'the father of John', or 'the set of lines parallel to λ', is the presence of the relevant *sortal* concept, be it 'father' or 'set'.[47] That is, presumably, right: But it is not an answer to our question, for our question is *how* sortals fix the referents of expressions in

[47] A sortal concept is one whose understanding requires an understanding of "a notion of *identity* for the things which fall under it", as Wright says: p. 2. See also Wiggins, pp. 58ff.

which they occur. Our question is what *kind* of sortal, say, 'direction' is. Is 'direction' just another way of saying 'set of such-and-such a kind' or 'line which passes through the origin'? *Why* is 'father' a sortal under which only persons fall? Indeed, why is 'person' not a sortal under which only sets fall?

Moreover, there must be some difference between to what the use of the sortal concepts 'person' and 'set' commit us when we use them in given functional expressions: That is, there must be some difference between the kind of use we make of an expression of the form 'the male person who is immediately causally implicated in the creation of ξ' and 'the set of all persons who have the same father as ξ'. Presumably, there is no reason we *could not* use expressions containing the former functional expression as names of sets, use them, so to speak, idiomatically. And that is to say that we may ask how we must use the expression if we are *not* to use it idiomatically, in what sort of way we must use it if we are to use it consistently with the plain intention that it is to be used to refer to persons.

The simplest answer to this question is that the sort of object to which a name refers is determined by the 'criterion of identity' for the names in question: No names which have distinct criteria of identity refer to the same object.[60] By 'criterion of identity', here, I mean no more than the condition for the truth of identity-statements containing such names, what is common to the truth-conditions of statements of the form 'the direction of λ = the direction of μ' or 'the father of John = the father of Jane'. It is important here that this notion of a criterion of identity is *intensional*, not in the sense that reference is made to intensional entities in a specification of the criterion of identity; for no such reference is made. Rather, the notion is intensional in the sense that substitution of a co-extensive relation for any relation mentioned in such a specification need not preserve its status as a correct specification of the criterion of identity.

This claim would immediately entail that, say, directions are neither lines, nor sets, but are *sui generis*. For identity-statements of the form 'the direction of λ = the direction of μ' are true if, and only if, λ is parallel to μ: The identity of sets, however, is determined by co-extensiveness; the identity of lines, by something else still. This simple answer is, however, incorrect. For consider the Contextual Definitions:

$$\text{dir } a = \text{dir } b \equiv_{df} a \parallel b$$
$$\text{dor } a = \text{dor } b \equiv_{df} \exists x(a \perp x \ \& \ b \perp x)$$
$$\text{dur } a = \text{dur } b \equiv_{df} \exists x(\text{angle}(a,x) = \text{angle}(b,x))$$

[60] This view is very similar to that offered by Bob Hale: See p. 215.

It seems to me that 'dorections' and 'durections' might well be our old friends directions. Perhaps they are distinct, but any principle which immediately entails that we can not identify them is too strong.

It is possible to weaken this view: Wright's condition N^{P} is such a weakening. Suppose that *Fx* is a sortal concept, names of objects falling under which are explained by means of a Contextual Definition given in terms of some equivalence relation '$\xi R \eta$' which holds between objects of sort S. Then, Wright's view is that[81]

> *Gx* is a sortal under which instances of *Fx* fall if and only if there are, or could be, terms, '***a***' and '***b***', which recognisably purport to denote instances of *Gx*, such that the sense of the identity statement, '***a***=***b***', can be adequately explained by fixing its truth-conditi tions to be the same as those of a statement which asserts that the given equivalence relation ['$\xi R \eta$'] holds between a pair of objects [of sort S].

We may understand this condition as follows: Objects of a sort F may be identified with objects of another sort G if, and only if, identity-statements concerning (some) Gs may be

---

[81] Wright, p. 114. It is tempting to read this passage as entailing the following: *Gx* is a sortal under which instances of *Fx* fall if there is, or could be, a class of terms recognizably purporting to denote instances of *Gx* which could themselves be intelligibly introduced (or explained) by means of a Contextual Definition otherwise identical to that by means of which names of objects falling under *Fx* were introduced. I do not know, however, whether Wright would accept this reading.

explained in the same way that identity-statements concerning Fs are explained.[52]

This view resolves the problem of directions and dorections. For, plausibly, identity-statements of the form 'the dorection of λ = the dorection of μ' *can* be explained in terms of the parallelism of λ and μ; plausibly, identity-statements of the form 'the direction of λ = the direction of μ' can be explained in terms of there being some line perpendicular to both λ and μ. The difficulty with Wright's proposal, however, is that it is not clear how to apply it in general. *Can* identity-statements of the form 'the work of which ξ is a copy = the work of which μ is a copy' be *explained* in terms of the co-extensiveness of the relevant equivalence classes? Conversely, we may presumably take the relevant names which 'recognizably purport' to refer to sets to be those of the form 'the set of all books which copy ξ'. Is it then possible to explain the senses of statements of

---

[52] There seems no reason not to suppose that the converse must also be true: If some Fs are Gs, then some Gs are Fs, so we must, presumably, also be able to explain identity-statements concerning (some) Fs in the same way we explain identity-statements concerning (some) Gs. Of course, since the condition, as formulated, applies only to sortals *Fx* names of instances of which are introduced by Contextual Definition, we will not be able to apply the condition as formulated unless *Gx* is also such a sortal.

This view would seem to inherit further plausibility from the fact that, if such explanations are possible, then it will be possible to explain the truth-conditions of ***mixed*** identity-statements--such as 'the direction of λ = the dorection of λ'--both in terms of the criterion of identity for directions and in terms of the criterion of identity for dorections.

the form 'the set of all books which copy **a** = the set of all books which copy **b**' in terms of 'ξ copies η'?

Perhaps not: But we need, at least, to be told more about what is packed into the notion of explanation here. For this reason, I shall propose a different sort of answer to our opening question and leave open the question whether it is compatible with Wright's view.

Let us return to a question we raised earlier. Consider the functional expressions 'the set of all books which copy ξ' and 'the oldest extant book which copies ξ'. Plausibly, these expressions refer to objects of different sorts: The former, to a set; the latter, to a book. Now, I said earlier that our use of the word 'set' in the former, and ou use of the word 'book' in the latter, expression must commit us to using these expressions in a particular way: That is, there is some sort of way we must use these expressions if we are to use them with the senses they appear to have. One might say, echoing Wittgenstein,[23] that we understand to what sort of object such an expression refers only because "the place for it was already prepared": That is, we already know how, in general, names of sets or names of books are used, and we are being told that, with this expression, we are to form names which are used in *that kind* of way. But in what way?

---

[23] Ludwig Wittgenstein, *Philosophical Investigations*, 3rd ed., tr. by G.E.M. Anscombe (Oxford: Blackwell, 1958), §31.

One would, to be sure, be surprised to hear such remarks as 'the set of books which copy **a** is in Texas', or '...was printed in 1935', or '...contains a torn page'. These are not the sorts of things one says of sets. The problem is not that we could not give an appropriate sense to such statements: We may say that a particular set is in Texas if all (or most) of its members are; we may say that a set contains a torn page if all (or some) of its members do. This point too can be formulated quite generally. For any functional expressions 'fnc1(ξ)' and 'fnc2(ξ)', which share weak identity-conditions, there will be, for any predicate 'F(ξ)' fit to be satisfied by objects to which we refer by means of the former expression, a predicate 'F^(ξ)', fit to be satisfied by objects to which we refer by means of the latter expression, with the following property:

F^(fnc2(a)) ≡ F(fnc1(a))

Anything which can be 'said about', say, the set of all books which copy **a** can be 'said about' the oldest extant copy of **a**, and *vice versa*: For example, the predicate corresponding to *membership* in such a set is just 'ξ copies **a**'.[54]

Thus, neither the weak identity-conditions associated with a given class of terms, nor the class of predicates one

---

[54] This point, of course, is just an 'object-language' re-formulation of the 'meta-language' point made earlier, that, so far as getting the truth-values correct is concerned, it matters not what we take the domain to be.

may intelligibly use with those terms, nor the sort of thing which can be said using those terms can determine the sort of object to which such terms refer.

Intuitively, however, even if there are always *predicates* which will serve to 'say about the set' the same thing we 'say about the book', nonetheless the book and the set have different sorts of *properties*. Having introduced the notions of a variant predicate and of an Ideology, we may now explain this intuition as reflecting the fact that we associate very different sorts of *Ideologies* with names of sets, on the one hand, and with names of books, on the other. What distinguishes a function from which we form names of sets from one from which we form names of books is the associated Ideology. There are certain sorts of predicates which one must understand how to use in conjunction with an expression like 'the set of books which copy **a**' if one is to understand it as a name of a set at all. And, in the same way, there are certain sorts of predicates which one must understand how to use in conjunction with an expression like 'the work of which **a** is a copy' if one is to understand it as a name of a work; with 'the father of b', if one is to understand it as a name of a person; with 'the location of b', if one is to understand it as a name of a place.

Given that we do not have a general account of the notion of a variant predicate and that, therefore, we have no general account of the notion of an Ideology, this proposal, in full generality, necessarily remains somewhat programmatic. But the proposal has a great deal to recommend it. Firstly, it has a compelling intuitive motivation, explained here in terms of the notion of a 'property'. Secondly, it gives us *some* kind of answer to the question what fixes the sort of object to which terms in a given class refer: That is, it gives us an answer to the question what distinguishes terms which share weak identity-conditions but which, intuitively, refer to objects of different sorts.

Thirdly, the account of what fixes the kind of object to which terms in a given class refer coheres with the earlier offered account of in what the *existence* of the referents of such terms consists. I argued that the existence of the referents of, e.g., names of day-persons consists in the truth of some theory which incorporates the relevant Ideology. It is natural to expect that our conception of the *kind* of object to which such terms refer should be closely connected with our conception of what it is for such objects to *exist*. For to say that a term refers to a day-person (if it refers at all) entails that it refers only if day-persons exist.

As I said earlier, however, a general defense of this proposal is beyond us at this time: And, presumably, the coherence of the proposal in any particular case depends upon its defense in the general case. Nonetheless, we do have an explanation of the distinction between variant and constant predicates in the case of names introduced by Contextual Definition: Hence, we have an explanation of the notion of an Ideology as it applies to such cases. It is possible, therefore, to make some further progress evaluating the proposal as it applies to such cases.

Consider the case of works of literature again. To understand the relevant Ideology in this case is, at least, to have a conception of *what is said* in a work and to know that (in the basic cases) what is said in a given work may be determined from any copy of it, that what is said is invariant among the copies. One who does not know such things does not understand our talk about works. Depending upon just what she does know, such a person's use of names of works should be all but indistinguishable from her use (or from others' use) of names of some different kind. For example, if she understood only statements such as 'work **a** = work **b**' or 'Each copy of work **a** has a torn page', and if she understood statements such as 'In work **a**, it is said that p' on the model of 'It is said that p in each copy of work **a**', her use of names of works would be all but indistinguishable from our use of names of *sets* of books.[88]

---

[88] It is worth mentioning here that the Ideology associated with sets (in such cases, equivalence classes) is in a certain sense ***minimal***: For, with a few exceptions, ***every*** statement about an equivalence class contains a constant predicate. That is to say, the determination of the truth or

Why should this matter? Why should it matter that, unless a speaker understands some variant predicates of works, we are tempted to say that she does not understand names of works (as we do, anyway)?--We are in search of a justification for the claim that this fact shows that names of works do not refer to sets. At least a partial justification is available to us: We have parallel intuitions in other cases; this is the point of our consideration of such functions as "the set of all books which copy ξ" and "the oldest extant copy of ξ". If a speaker uses, with expressions such as "the oldest extant copy of the *Grundlagen*", only such predicates as are fit to be used with names of sets, the speaker does not understand this expression as we do, as a name of a *book*. It is important not only that a speaker understand predicates of the relevant sort (say, variant predicates of books), but that she know that *that* class of predicates is the relevant class in a given case.

If an understanding of the Ideology associated with such functions as "the set of all books which copy ξ" and "the oldest extant copy of ξ" is required if one is to understand them to refer to objects of the appropriate sort,

---

falsity of a statement about the equivalence class ordinarily depends upon one's knowledge of what members the set has: I.e., which, if any, other objects bear the relevant equivalence relation to some given object. It is for this reason that, in the absence of any indication of a more substantial Ideology, we can not help but take names introduced by Contextual Definition as names of sets. The more substantial the Ideology, the less 'set-like' the objects become.

surely we must say the same about such functions as "the work of which ε is a copy". The only relevant contrast is that, in the former case, we assume, for the sake of argument, that *there are* such objects as books and sets and we are asking how reference to them is to be distinguished: In the latter case, we are arguing that the kind of thing which distinguishes expressions which refer to books from those which refer to sets also distinguishes our use of names of works from our use of expressions of either of these sorts. Since these expressions must refer to something, they must refer to something other than books or sets.

6. Closing

The most plausible reply to such considerations is that, in talking about such aspects of the use of certain functions, we have assumed that we are thereby talking about features of a speaker's *understanding* of names in a given class: The objection is that *semantics* ought not to be expected to concern itself with such matters. From this perspective, there is no deep distinction between predicates of the sort I have called 'variant' and those of the sort I have called 'constant': What is said in a given work of literature is merely what *interests* us most, rather than, say, whether all copies of it have torn pages: This kind of distinction falls within the domain of pragmatics, not within that of semantics.

This view appears to be committed to the claim that there is but one sort of object. Even the view that all abstract objects are sets would not be sufficient here. For consider, again, the contrast between "the set of all copies of ℓ" and "the oldest extant copy of ℓ". Either (terms formed from) these refer to objects of the same sort or they do not: If they do not, then we are owed some account of the distinction, which is therefore of semantical significance. An account which is an alternative to mine might well be offered; whatever it is like, the dispute then concerns the nature of such an account, not its semantical or pragmatic nature. We shall return to this point.

If the objection is to be that *any* such distinction is merely pragmatic, (terms formed from) any two functional expressions must refer to objects of the same sort. But this entails that there is only one sort of object. This follows from the observation that, for each sort of object, there is a functional expression (with the same weak identity-conditions) whose range purportedly consists of objects of that sort. Hence, if all names formed from functional expressions refer to objects of the same sort, there is only one sort of object. This consequence can only be avoided by denying that a name such as "the oldest extant copy of a" refers, as it appears to refer, to a book: On this view, all (names formed from) functional expressions refer to objects of the same sort, a sort which is distinct from any to which names not formed by means of functional expressions refer. But this view does not appear to be coherent. And the alternative view, that there is only one sort of object, faces objections I shall not rehearse here: The fundamental difficulties facing a physicalistic construal of mathematics (other, I should again emphasize, than a Fictionalist account) are conclusively enough presented by Frege.[56]

To develop an alternative to the view presented here, however, it is not necessary to claim that *any* such distinction is of only pragmatic significance. The argument given above was predicated upon the assumption that we shall want to make distinctions among the sorts of objects to which different classes of expressions which refer to *concrete* objects refer. The general form of the argument was: Once we have found a way to distinguish among expressions which refer to different sorts of concrete objects, we shall thereby have found a way to distinguish among different sorts of abstract objects.

If this is the structure of the argument, however, it is clear what sort of view constitutes an alternative. What is required is a way of drawing a distinction between names of concrete and names of abstract objects. The claim would then be that there are essentially *two* sorts of objects, concrete and abstract. The former may plausibly be identified as the concrete mereological atoms and their fusions;

---

[56] See *Grundlagen*, §§7-10, 23-5.

the latter, as sets.[57] One may be skeptical that there is any very definite line between abstract and concrete objects.[58] The choice of the example of day-persons was intended to reinforce such skepticism. Nevertheless, it may well prove possible to draw this distinction in a principled way.[59] And, if it is, a defender of this alternative view could well take over such an account as an account of what distinguishes names which refer to abstract objects from those which refer to concrete ones.

I can not offer any appraisal of this view here. It is, however, clear upon what its appraisal rests. Its tenability rests upon the tenability of the view that we need not distinguish among sorts of concrete objects. Such a view is of far greater generality than any I have been able to consider here; its evaluation similarly depends upon far more general considerations.[60] But it is enough to have reduced the dispute over the intelligibility of reference to abstract objects to one about the necessity for a distinc-

---

[57] And, indeed, sets may be taken to be abstract mereological atoms (singletons) and *their* fusions. If so, then there will likely also be peculiar 'mixed' objects--the fusion of myself with my singleton, for example. But that is beside the point. See David Lewis's *Parts of Classes* (Oxford: Blackwell, 1991).

[58] This lesson might be gleaned from Dummett's discussion of the distinction in Ch. 14 of his *Frege: Philosophy of Language*.

[59] For an attempt, see Hale, Ch. 3.

[60] For a presentation of some of these considerations, see, of course, Wiggins's *Sameness and Substance*.

tion among sorts of concrete objects. Indeed, one might well emphasize, at this point, that the notions of reference to works of literature, to numbers, and so forth, as distinct from the notion of reference to sets, is now no worse off than are the notion of reference to persons, to stars, to statues, and, indeed, to books themselves as distinct from the notion of reference to the matter which constitutes them.[41]

---

[41] I should like to thank George Boolos, Sylvain Bromberger, Bob Hale, Jim Higginbotham, Thomas Kuhn, Bob Stalnaker, and Crispin Wright for their comments upon and criticisms of earlier versions of this material. I should also like to thank those who attended a reading at the Wolfson Colloquium in Oxford for the helpful discussion.

BIBLIOGRAPHY

Barwise, Jon and Cooper, Robin. "Generalized Quantifiers and Natural Language", *Linguistics and Philosophy* 4 (1981), pp. 159-219.

Boolos, George. "The Consistency of Frege's *Foundations of Arithmetic*", in Thomson, ed., *On Being and Saying*, pp. 3-20.

-----. "The Standard of Equality of Numbers", from Boolos, ed., *Meaning and Method*, pp. 261-77.

Boolos, George, ed. *Meaning and Method: Essays in Honor of Hilary Putnam*. Cambridge: Cambridge University Press, 1990.

Bromberger, Sylvain. "Types and Tokens in Linguistics", in George, ed., *Reflections on Chomsky*, pp. 58-89.

Davidson, Donald. "Reality Without Reference", *Dialectica* 31 (1977), pp. 247-58.

Dummett, Michael. *Truth and Other Enigmas*. Cambridge MA: Harvard University Press, 1978

-----. "Wang's Paradox", in *Truth and Other Enigmas*, pp. 248-68.

-----. "Realism", in *Truth and Other Enigmas*, pp. 145-65.

-----. "The Philosophical Basis of Intuitionistic Logic", in *Truth and Other Enigmas*, pp. 215-47.

-----. "What Is a Theory of Meaning? (II)", in Evans and McDowell, eds., *Truth and Meaning*, pp. 67-138.

-----. *Frege: Philosophy of Language*, 2nd ed. London: Duckworth, 1981.

-----. *The Interpretation of Frege's Philosophy*. London: Duckworth, 1981.

-----. "Reply to Wright", in Taylor, ed., *Essays on the Philosophy of Michael Dummett*, pp. 221-34.

-----. *The Logical Basis of Metaphysics*. Cambridge MA: Harvard University Press, 1991.

Evans, Gareth. "Can There be Vague Objects?", *Analysis* XXXVIII (197f), p. 208.

-----. *Varieties of Reference*, ed. by J. McDowell. Oxford: Oxford University Press, 1982.

Evans, Gareth and McDowell, John, eds. *Truth and Meaning*. Oxford: Oxford University Press, 1976.

Frege, Gottlob. *The Foundations of Arithmetic*, tr. by J. Austin. Evanston IL: Northwestern University Press, 1980.

-----. *Collected Papers*, ed. by Brian McGuinness, tr. by M. Black, *et al*. Oxford: Basil Blackwell, 1984.

Field, Hartry. *Science without Numbers*. Oxford: Basil Blackwell, 1980.

Friedman, Harvey and Sheard, Michael. "An Axiomatic Approach to Self-Referential Truth" (draft).

Garrett, B.J. "Vagueness and Identity", *Analysis* XLVIII (1988), pp. 130-4.

George, Alexander, ed. *Reflections on Chomsky*. Oxford: Basil Blackwell, 1989.

Hale, Bob. *Abstract Objects*. Oxford: Basil Blackwell, 1987.

Hodes, Harold. "Ontological Commitment: Thick and Thin", in Boolos, ed., *Meaning and Method*.

Horwich, Paul. *Truth*. Oxford: Basil Blackwell, 1990.

Lewis, David. "Vague Identity: Evans Misunderstood", *Analysis* XLVIII (1988), pp. 128-30.

-----. *Parts of Classes*. Oxford: Basil Blackwell, 1991.

McGee, Vann. "Applying Kripke's Theory of Truth", *Journal of Philosophy* LXXXVI (1989), pp. 530-9.

Pelletier, Francis Jeffry. "Another Argument Against Vague Objects", *Journal of Philosophy* LXXXVI (1989), pp. 481-92.

Putnam, Hilary. "Models and Reality", in *Realism and Reason*. Cambridge Cambridge University Press, 1983, pp. 1-25.

-----. "Model Theory and the 'Factuality' of Semantics", in George, ed., *Reflections on Chomsky*, pp. 213-32.

Quine, Willard van Orman. *From a Logical Point of View*, 2nd ed. rev. Cambridge MA: Harvard University Press, 1980.

-----. "On What There Is", in *From a Logical Point of View*, pp. 1-19.

-----. *Word and Object*. Cambridge MA: MIT Press, 1960.

-----. *The Ways of Paradox*. New York: Random House, 1966.

-----. "Ontological Reduction and the Theory of Numbers", in *The Ways of Paradox*, pp. 199-207.

-----. "The Scope and Language of Science", in *The Ways of Paradox*, pp. 215-32.

Taylor, Barry, ed. *Essays on the Philosophy of Michael Dummett*. Dordrecht: Maritnus Nijhoff, 1987.

Thomson, Judith Jarvis, ed. *On Being and Saying: Essays for Richard Cartwright*. Cambridge MA: Massachusetts Institute of Technology Press, 1987.

Wiggins, David. *Sameness and Substance*. Oxford: Basil Blackwell, 1980.

Wittgenstein, Ludwig. *Philosophical Investigations*, tr. by G.E.M. Anscombe. Oxford: Basil Blackwell, 1958.

-----. *Remarks on the Foundations of Mathematics*, 3rd ed., tr. by G.E.M. Anscombe. Oxford: Basil Blackwell, 1978.

Wright, Crispin. *Frege's Conception of Numbers as Objects*. Aberdeen: Aberdeen University Press, 1983.

-----. *Realism, Meaning, and Truth*. Oxford: Basil Blackwell, 1986.