

PHILOSOPHICAL LECTURE

UNDERSTANDING LOGICAL CONSTANTS: A
REALIST'S ACCOUNT¹

By CHRISTOPHER PEACOCKE

Read 13 May 1987

1. *Aims*

OSCAR Wilde has one of his characters say that he can resist anything except temptation (Wilde, 1966, p. 388). This paper is primarily for those who find themselves in a similar position *vis-à-vis* the temptation to suppose that there is some value in tracing out the consequences of a correct account of what it is to possess a given concept. The particular form of temptation to which I will be succumbing is that in which the concept in question is that of a logical constant.

I aim to give an account of the understanding of logical constants, and of the justification and validity of principles containing them, which is available to a realist. By a realist, I mean merely a theorist who allows that a sentence or content can be true though unverifiable by us. Some of the styles of argument I will be suggesting can in fact be transposed to a verificationist context. But one of the main tasks a realist faces in this area is a result of the impossibility of certain transpositions in the opposite direction: I will later be arguing that some of the most frequently cited justifications and constraints on logical principles are ones which are not properly available to the realist. The realist has to identify constraints and justifications of his own.

This single paper is not, and could not be, a general defence of realism. At some points, particularly in connection with the general semantical framework, I take for granted some appa-

¹ A version of this material was presented to a seminar given jointly with Ian McFetridge in London University in the autumn of 1986. I owe him special thanks for his extensive constructive criticisms. I have also received valuable comments from John Campbell, Martin Davies, Graeme Forbes, Daniel Isaacson, David Over, Andrew Rein and David Wiggins.

tus for which a realist must elsewhere provide arguments. But until he has a good theory of the sense of logical constants, any realist's views will be incomplete.

Many of the questions which must be addressed by a good philosophical account of the logical constants are ones which arise for every kind of content. Part of the interest of pursuing these issues about the logical constants is that they provide a source of hypotheses about these more general issues. The more general issues include the relations between conceptual-role and truth-conditional theories of content; the role of normative factors in the individuation of content; the relations between knowing something and the individuation of what is known; the idea that some sentences are automatically true given their meanings; and the pattern of relations between the impression that a certain content holds and its really doing so. The logical constants provide a relatively well-surveyed testing ground for hypotheses on all these matters.

The other questions I will be addressing are specific to the logical constants. These questions include the issue of how logical principles are to be justified, and the possibility of reconciling the utility of deduction with its justifiability. But even in the case of questions apparently specific to the logical constants, there is no prospect of justifying one's answers without appealing to general considerations drawn from the theory of content.

2. *A simple case: claims and observations*

We can lead into these topics by considering first a relatively unproblematic constant, conjunction, with its introduction and elimination rules:

(&I)	(&E ₁)	(&E ₂)
$\frac{A \quad B}{A \& B}$	$\frac{A \& B}{A}$	$\frac{A \& B}{B}$

Anyone who understands '&' finds instances of these principles compelling. That is, we can fix upon particular contents A, B and say: it strikes one who understands '&' as obvious that given A together with B, then A&B. The impression of obviousness is primitive in the sense that it is not consequential upon his acceptance of some more primitive principle; nor upon iterated application of any single principle; nor upon any other belief not

already presupposed in grasp of the component contents A and B.² We can encapsulate this in Claim 1:

(Claim 1) For this logical constant, there are principles containing it whose instances are found primitively obvious by someone who understands it.³

It is plausible that a slightly stronger claim is true, viz., Claim 2:

(Claim 2) For this logical constant, finding instances of these principles primitively obvious is at least partially constitutive of understanding it (of 'grasping the sense it expresses', in the classical terminology).

To say that it is so partially constitutive is to make the following conjunctive claim: that no one understands '&' unless he finds these instances primitively obvious, and that the explanation of this fact is to be traced to what is involved in understanding conjunction.

The reason for accepting Claim 2 is not some generalization to the effect that understanding *any* logical constant always involves finding primitive principles containing it primitively obvious. On the contrary, I shall soon be suggesting that such a generalization is false. However, in the cases in which someone understands a logical constant but does not find an underived principle containing it primitively obvious, there is always some property the understander has, in relation to the constant in question, and which makes it right to say he understands it. (Examples will follow.) But no such intermediate case seems possible for '&'. It seems impossible to think of anything which could make it right to say a thinker does understand '&', but which falls short of requiring him to find instances of its introduction and elimination rules primitively obvious.⁴

² Throughout this paper, when writing of inference and reasoning, I mean only inference and reasoning at the personal level, in Dennett's (1969) sense. Subpersonal inference and reasoning may well be rife in the production of an immediate impression of obviousness in my sense.

³ Strictly, not all instances will be found primitively obvious: long or confusing instances will not. There are two ways one might take account of this. One is to restrict Claim 1 to instances which are obviously of one of the forms in the introduction or elimination rules. The other is to revise it to say that all instances are either found primitively obvious, or are of a form which, when instantiated in the cases found primitively obvious, is the form in virtue of which such instances are found primitively obvious. Henceforth I omit the qualification that not all instances are found primitively obvious.

⁴ Strictly, given the arguments of section 3 below, it is necessary only that either the introduction rules or the elimination rules be found primitively obvious for the classical semantic value to be determined.

Does Claim 2 illegitimately elevate what is just a psychological generalization about subjects who understand conjunction to a purported necessity? No: empirical psychological claims are not at issue here. Claim 2 is a claim about a constant *with a given sense*. If it is true at all, Claim 2 is necessary because it is a claim about what contributes to the individuation of a given sense. It follows from conditions governing the notion of sense that principles which must be found primitively obvious by anyone who grasps a sense also contribute to the individuation of that sense. The argument is as follows. For the logical constants, as for every other category of expression, Fregean sense is individuated by considerations of informativeness. So consider two logical constants $*$ and \S . Suppose anyone understanding $*$ must find a certain principle with the conclusion $A*B$ primitively obvious, whereas the same is not true of \S and the corresponding principle with \S uniformly replacing $*$. It follows that the senses of $*$ and \S are distinct. For the thinker may believe premisses which it is primitively obvious imply $A*B$, while it is not primitively obvious that they imply $A\S B$. Since such a thinker may rationally judge $A*B$ without judging $A\S B$, it follows from the criterion of informativeness that the senses of $*$ and \S differ.

For an example of this state of affairs, we can take $*$ to be conjunction, and \S to be a hypothetical primitive connective **C** with the following properties: one who understands it finds $\sim A$ primitively incompatible with ACB , finds $\sim B$ primitively incompatible with ACB , and does not find anything else essentially involving this connective **C** to be primitively obvious. Now consider a thinker who already accepts A and accepts B . He will accept $A\&B$ if the question arises. But he need not accept ACB ; it would take some reasoning to realize that it can be inferred from the premisses A and B . On the present conception, there is nothing problematic about two constants, like $\&$ and **C**, picking out the same truth function, whilst having different senses.⁵

⁵ Three comments on this example: (a) I have used the pair $A\&B$, ACB to make the point, rather than (say) the pair $A\&B$, $\sim(\sim A \vee \sim B)$. The latter pair does not provide a compelling example to illustrate the possibility of difference of sense with sameness of reference—since it can be captured by using not senses, but tree structures of referents corresponding to the syntactic structure of $\sim(\sim A \vee \sim B)$. This response is unavailable for the unstructured **C**. (b) It is arguable that ACB has a different sense from that of $\sim(\sim A \vee \sim B)$; it is not primitively obvious, but takes some simple reasoning, to realize that $\sim A$ is incompatible with the latter. (c) For enthusiasts concerned that, for all I have said, **C** may have as its semantical value something stronger than the classical truth function for conjunction: an application of the tactics of section 3 below rules out this possibility.

The third Claim I wish to make is this:

(Claim 3) What makes a particular function the semantic value of '&' is that it is that function which, applied to the semantic values of the expressions on which the conjunction operates, ensures that the principles instances of which are found primitively obvious are indeed genuinely truth-preserving.

If Claim 3 is right, then we can argue as follows. If the rule (&I) is to preserve truth, then $A \& B$ must be true when A is true and B is true. Similarly, $A \& B$ must be false when either A is false or B is false; if it were not, then either (&E1) or (&E2) would fail to preserve truth. This determines the classical semantic value for conjunction. The argument is of a form occurring in Hacking (1979) and Peacocke (1976) and in the former is labelled 'Do-It-Yourself Semantics'.

I will be defending Claims 1–3, and suitable generalizations thereof, to harder cases. I will be defending these generalizations not just as intuitively plausible, but by appeal to the explanatory power of a theory which assumes them. It will be obvious to anyone who has thought about this area that, in initially considering just conjunction, we abstract from many crucial problems. But some of the welcome consequences of the more general claims which are true of all logical constants are present in a sharp, uncomplicated form in this simple special case; so let us trace out three of these consequences first, before proceeding.

A first observation is that if Claims 1–3 are right, then the principles which are found primitively obvious are indeed truth-preserving. Claims 1–3 ensure that the primitive impressions of validity involving '&' are veridical.

A second observation is that we have here a confirming instance of the Conjecture (labelled '(C)' in *Thoughts* (Peacocke, 1986)) that normative acceptance conditions determine truth conditions. If Claims 1–3 are true, the principles whose instances are primitively compelling are, as valid, genuinely correct norms for the truth-conditional contents expressed by the sentences in question. If Claims 1–3 are right, these norms determine the distinctive contribution made by conjunction, both at the level of sense and at the level of semantic value. They determine the contribution in the sense that there cannot be two connectives governed by exactly the same normative rules as those for '&' but which differ in respect of the contributions they make to the truth-conditions of sentences containing them.

A third, historical observation is that the respect in which instances of the introduction and elimination rules for '&' are

primitively compelling for one who understands ‘&’ is a respect emphasized by Wittgenstein. He took the respect in question as a mark of possession of a given concept. ‘For the word *must* surely expresses our inability to depart from *this* concept ...’ (1978, IV.30); ‘... The mathematical *Must* is only another expression of the fact that mathematics forms concepts’ (1978, VII.67). When someone understands a word, logical constants included, there are situations in which it will seem to him that one rather than another application of the word is required, if he is to be faithful to its meaning.

The cluster of interrelated phenomena which characterize other cases of successful rule-following are present for logical constants too. I emphasized that it strikes someone who understands ‘&’ that A follows from A&B, and that this impression is not the result of any inference. This is the state of affairs Wittgenstein describes when he says that the rule-follower ‘can give no reason’ for the correctness of his application (1978, VI.24). As we would expect, he emphasizes that stopping at the point at which interpretation comes to an end can nevertheless still be *justified* in the logical case—‘Logical inference is a transition that is justified if it follows a particular paradigm and its rightness is not dependent on anything else’ (1978, VII.66). I will also turn to the question of justification and knowledge some way below. I doubt that Wittgenstein would find everything in the rest of this paper congenial; but on the importance of these points about the primitively compelling impressions of an understander, the two accounts are at one.

3. *Undersived, unobvious but justifiable*

Now we can turn to some more problematic constants, where more interesting issues arise. It is not, for instance, possible to give for existential quantification a treatment which simply parallels that we gave for conjunction. Instances of the natural deduction rule of existential elimination (EE)

$$\begin{array}{c}
 \exists x Fx \quad [Ft] \\
 \vdots \\
 \text{(subject to the usual restrictions on } t) \\
 \hline
 C \\
 \hline
 C
 \end{array}$$

are not primitively obvious to all who understand the existential quantifier. To someone who already understands existential

quantification, the fact that an instance of (EE) is valid can be as informative as the claim that a certain derived transition—such as one of de Morgan's laws—is valid. In being answerable to the individuation of sense by considerations of informativeness, we should be as sensitive to this fact about (EE) as we would to the fact about de Morgan's laws.

Can it help us that, even if (EE) is not primitively obvious, then at least other rules with equal powers in other systems *are* primitively obvious? Perhaps, for instance, it is true that the rule of existential introduction in the antecedent in a Gentzen-style sequent system will be primitively obvious to those who understand its notation.⁶ But many who understand existential quantification have never encountered such a sequent system. If their quantifier has the same sense as one for which certain principles will be immediately obvious to a user of the sequent system, then that quantifier must possess that sense in virtue of actual facts about actual thinkers' use of it. Our task is to say what those facts are.

Is a state of affairs in which a thinker finds an undervived principle not primitively obvious possible only when the thinker's understanding of the existential quantifier is merely partial? The pretheoretical notion of degree of understanding is certainly sufficiently accommodating to allow one to say that the more valid principles for the quantifier someone knows, the better he understands it. (It is sufficiently elastic also to allow one to say truly that the more you know about uranium, the better you understand it.) The important point, however, is that there is a real distinction between (i) a thinker whose understanding is partial in the sense that given the way he understands it, it is not true (or not determined whether) (EE) is a valid principle, and (ii) a thinker who also does not find (EE) primitively obvious, but for whom nevertheless (EE) is a valid principle, and can be shown to be so, for what he means by the existential quantifier. We have to explain how this is possible, consistently with the status of (EE) as an undervived rule and the framework offered so far.

I am discussing grasp of sense in an already existing language.

⁶ For Gentzen's sequent systems see Kleene (1971), chapter XV. The rule of existential introduction in the antecedent is the rule

$$\frac{F(t), \Sigma \Rightarrow \Delta}{\exists x F(x), \Sigma \Rightarrow \Delta}$$

(again subject to the usual restrictions on t).

If our aim were rather to introduce some new vocabulary into a language, then, subject to some requirements to be discussed later, we would be able to *stipulate* certain logical principles as holding for the new vocabulary; and versions of both existential introduction and existential elimination could be amongst those laws. What this shows is that what we may legitimately stipulate does not coincide with what is uninformative. Though there are important connections between them, the project of investigating what may be legitimately stipulated and the project of investigating grasp of sense in an already existing language are distinct. We have here another kind of illustration of the point urged by Burge (1979, 1986) that understanding an expression need not require primitive recognition of the truth of normative statements involving the expression—not even statements which have a meaning-giving status. In Burge's examples, the possibility of understanding without such recognition is present because a thinker may be ignorant of, or have false theories about, the kinds whose environmental relations to him help to individuate the intentional content of his attitudes. In the present case, we have a further form of the phenomenon even for contents which do not refer to object, events, or stuffs in the thinker's environment.

In both the Burgean and the logical cases, we are under an obligation to say why the thinker nevertheless possesses the concept in question. In the logical case, we can discharge this obligation by drawing on resources already available to us. Let us take existential quantification. Here is one way of explaining the possibility in question. In judging an existential quantification, a thinker is judging something to which he is committed by any one of its singular instances. These instances may be said to be canonical grounds for judging the quantification, in that their status as grounds is not dependent upon any collateral information the thinker may possess. The rule of existential introduction is validated by an existential quantification's possession of these canonical grounds, and is primitively obvious. We need, though, to distinguish between a content's possessing certain canonical grounds, and those grounds *exhausting* its range of canonical grounds. A pair of contents may have overlapping but distinct families of canonical grounds. It is one thing to be sensitive in one's judgements to a content's possession of a certain range of canonical grounds. It is another, further, thing to come to realize by reflection that those are all the canonical grounds for that content, that there are no others. When a thinker reaches this

point, he may well be in a position to endorse as valid principles which he could not, rationally, endorse as valid prior to that reflection. (EE) is an example of such a principle. Of course to endorse (EE) rationally, the thinker has not only to realize that a certain family exhausts a judgement's canonical grounds; he has also to realize that (EE) relies only on this point.

It is not necessary on this account that anyone soundly reasoning to the validity of (EE) has thoughts involving the notion of a canonical ground. For instance, a thinker may instead (and more likely) use the notion of truth, and reason from the premise: 'Nothing besides the truth of Fx for a particular object x is sufficient for the thought that something is F to be true.' This premiss is equivalent to the premiss that the singular instances Fx , one for each object x , exhaust the canonical grounds of the thought that something is F . Any other equivalent would serve equally well.

We can begin to generalize this, and to remove some psychological accretions. First we need an auxiliary notion. Consider a logical constant \S , which operates on sentences $A_1 \dots A_n$ to form a sentence $\S(A_1 \dots A_n)$. Let us say that one semantical assignment to the operator \S is *stronger* than a second if anything of the form $\S(A_1 \dots A_n)$ under the first assignment entails the same sentence under the second assignment, while the converse is false (i.e. the sentence under the second assignment does not entail the sentence under the first assignment). So the classical truth function for conjunction is a stronger semantical assignment for a binary sentential connective than the classical truth function for the material conditional.

Now suppose we have a set I of introduction rules for a logical constant, and that a thinker finds instances of these rules primitively obvious. Suppose too that these are *all* the principles essentially containing that constant whose instances he finds primitively obvious. Then we define a *limiting principle* for I as a principle not derivable from I and which is validated by the maximally strong semantical assignment to the logical constant in question which validates all the principles in I . Only the classical model-theoretic value for the existential quantifier is such an assignment in the context of a classical realistic semantic theory; and (EE) is a limiting principle for existential introduction (EI). A *maximal* set of limiting principles for I is a set of limiting principles for I from which any limiting principle for I is derivable. (EE), or strictly its unit set, is a maximal limiting principle for (EI).

In the case of existential quantification, we pick the strongest semantical assignment which validates (EI) not for the reason that it will also validate (EE). That would destroy any chance of justifying (EE). The ground for picking the strongest assignment is rather that nothing else besides an instance Fx is a canonical ground for judging $\exists xFx$. The underlying reason here is that if a weaker semantical assignment were correct, there would have to be further primitively obvious relations involving the existentially quantified form—which there are not.

Of course there is no reason in general why it should be introduction rules which are primitively obvious, rather than elimination rules. For any introduction rules which are not primitively obvious, a dual application of the techniques of the preceding paragraphs takes care of the fact. We similarly introduce the idea of a set of limiting principles for a set of elimination rules E by systematically replacing 'I' by 'E' in the preceding, and 'stronger than' by 'weaker than', where these relations are the converses of one another.

In Prawitz's elegant writings on deduction, one finds a general formulation of the strongest elimination rule corresponding to a general schematic introduction rule (1978, p. 37). Prawitz takes the introduction rules as constitutive of meaning, and the elimination rules as consequential. We have already made the distinction between introducing new constants and describing an already existent understanding. In the case in which we are introducing new vocabulary, we could introduce primitive elimination rules and determine the corresponding introduction rules; or conversely. For the case of an already existent understanding, we will give some examples in the next section in which neither the standard introduction rules nor the standard elimination rules are primitively obvious. What I have in effect been doing in these recent paragraphs is arguing that the conception of an elimination or an introduction rule determined by other rules is not something which has a life only in pure proof-theoretical investigations, but corresponds to cognitive phenomena involving real thinkers.

How do our earlier Claims and observations stand when we bring the quantifiers within our scope? Claims 1 and 2 stand as before; but Claim 3 needs generalization to cover examples in which we have to appeal to limiting principles. The generalization, Claim 3Gen, runs:

(Claim 3Gen) For each constant so far considered, what makes a particular function its semantic value is this: the function is that function of the semantic values of the expressions on which the constant operates which ensures both

- (a) that the principles containing it which a thinker finds primitively obvious are truth-preserving, and
- (b) that any maximal set of limiting principles for the constant is also truth-preserving.

Our earlier consequential observation was that we have determination of the contribution to truth conditions by norms for the contents in question; this continues to hold for the quantifiers too, provided that the norms now include not only what is primitively obvious, but the limiting principles too.

4. *Classical negation*

We can also use limiting principles in treating negation. What is primitively obvious to anyone who understands negation is just that $\sim A$ is incompatible with A . Maybe it is going too far to say that the ordinary user of negation has the concept of incompatibility; but it would not be going too far to say that unless he appreciates that A and $\sim A$ cannot both be true, then he does not understand \sim .

It takes further reflection to realize that $\sim A$ is also the weakest condition incompatible with A . That it is the weakest does not follow just from the incompatibility of $\sim A$ with A . There are many contents stronger than $\sim A$ which are also incompatible with A . There are several essentially equivalent ways in which a thinker may reach the conclusion that $\sim A$ is the weakest such condition. One is by starting from the realization that $\sim A$ is true in any case in which A is not. Another would be the realization that if $\sim A$ were not the weakest such condition, then there would be a consistent content whose truth requires neither that of A nor that of this supposedly stronger negation of A .

It is precisely because $\sim A$ is the weakest condition incompatible with A that $(\sim I)$ is a valid rule (here I follow Prawitz's (1965) notation):

$$\begin{array}{c}
 (\sim I) \quad [A] \quad [A] \\
 \cdot \qquad \cdot \\
 \cdot \qquad \cdot \\
 \hline
 B \quad \sim B \\
 \hline
 \sim A
 \end{array}$$

If $\sim A$ were not the weakest condition incompatible with A , then from the fact that from A (and other assumptions) one can derive incompatible propositions, one would not be able to conclude that $\sim A$. It would be a *non sequitur*, since the premisses would not have excluded the possibility that something holds which is incompatible with A , but which does not imply this stronger 'negation'.

Once a thinker has worked out that $\sim A$ is quite generally the weakest condition incompatible with A , for arbitrary A , he is in a position to infer that double negation elimination is valid. Anything incompatible with $\sim A$, i.e. something which entails $\sim \sim A$, must also entail A too—on pain of $\sim A$ not being the weakest condition incompatible with A .

When in these recent cases we talk of a thinker's reaching a principle by reflection, this cannot be taken as reaching it by deduction from primitively obvious principles. The principle reached by reflection does not follow by reasoning available at this stage from the primitively obvious principles. At some point or other in his reflection, the thinker must use something which goes beyond the primitively obvious incompatibility of $\sim A$ with A . In the thinking suggested above for him, the thinker went beyond the primitively obvious in accepting the principle that in any case in which A is not true, $\sim A$ is true.⁷

On the present treatment, an account of what it is to possess the concept of classical negation does not require primitive unreflective acceptance of either of the classical introduction and elimination rules for negation. These rules are rather justified in a way parallel to that in which limiting principles were justified above. This time, though, instead of a derivability relation, the constant negation is introduced over an incompatibility relation. Suppose then that we have a set S of rules for a constant, rules framed in terms of an incompatibility relation. We suppose too that the thinker finds instances of these rules primitively obvious, and that these are all the principles essentially containing that constant whose instances he finds primitively obvious. Again, a limiting principle for S is a principle not derivable from S and which is validated by the weakest semantical assignment to the

⁷ This raises the question of how the thinker knows such principles. Perhaps an account should be developed on which he knows them because they are true given the way he is using \sim , and his reflections are influenced by (and answerable to) his unreflective practice in using \sim . His unreflective practice with \sim in particular cases can show that no more is required for the truth of $\sim A$ than the non-truth of A . These brief remarks do not do justice to the issues which arise here; the issues deserve extended attention. On the question of how knowledge of underived primitively obvious principles is possible see section 6 below.

constant which validates the primitively obvious incompatibility relations.

As before, the only semantical assignment, in the context of a classical theory, which validates the primitively obvious incompatibility relations in which a negated proposition stands, together with the limiting principles, is the classical truth function for it. Just as we previously extended Claims 1, 2 and 3 to take into account the limiting principles for existential quantification, we should make a parallel move in the case of negation. What makes something the correct semantic value of negation is that it validates not only the primitively obvious incompatibility principles, but also the limiting principles. Henceforth we will read Claim 3Gen as covering the case of negation too.

It may be objected that an intuitionist will equally agree that it is primitively obvious that $\sim A$ is incompatible with A . Indeed he may also agree that there is nothing weaker than $\sim A$ which is incompatible with A . But if this is so, how can these two points determine classical rather than intuitionistic negation?

The answer is that it is not the same points which hold for the intuitionist as for the classicist. When the intuitionist agrees that it is primitively obvious that $\sim A$ is incompatible with A , what he means by 'incompatible' is not what the classicist means. What the intuitionist means by the incompatibility of A with B is that the supposition that A and B are both verified leads to absurdity. In the intuitionist's sense, A is incompatible with 'It is not verified that A '. That is congenial to him, but there is no incompatibility on the classicist's realistic notion. Since different notions of incompatibility are being used, there is no sound objection to the claim that the semantic value of classical negation is determined. (This response to an *ad hoc* objection does not imply that intuitionistic negation is itself unproblematic).

5. *Justification, the semantic constraint and conservative extension*

If Claims 1, 2 and 3Gen are right, then it involves a false dichotomy to suppose that justifications of primitive logical principles must be exclusively either proof-theoretic or semantic. Underlying acceptance of that dichotomy may be the idea that either primitive acceptance of primitive logical principles must be written into an account of what it is to understand the logical constants, or else their rational acceptability flows exclusively from their validation by a semantic theory which owes nothing to proof-theoretic considerations. Thus Prawitz, who favours a proof-theoretic account of the justification of intuitionistic logic, contrasts his own view with 'a second view ... [according to

which] one has first to clarify the notions of meaning and truth independently of the notion of proof, and once this is done, the notion of proof is easily tackled' (1978, p. 25).

On the present theory, however, elements of the views of both parties to this discussion are correct. Learning certain of the primitive laws *is* part of coming to understand some of the constants; but those laws also receive a justification from a non-proof-theoretic semantics. On the other hand, what makes the assignments given in the semantics correct is their validation of certain primitive and limiting principles. Without some such principles, we would have no source for the semantics; without the semantics, we would have, it seems, no justification for the principles.

That last point is likely to trigger a fundamental objection. Why isn't the discussion of the determination of the semantic value of a logical constant in the account so far simply an unnecessary detour? Why can't the principles a thinker has to find primitively obvious be regarded as self-justifying? To say that they are self-justifying is not to say that the question of justification does not arise. It is to allow that the question does arise, but is answered in a special way. The special way is that the meaning of the constant is given by its conformity to these primitively obvious principles; so another constant cannot have the sense of the given constant without conforming to those principles. We know from Prior's famous note 'The Runabout Inference Ticket' (1960) about the spurious connective *tonk* that not just any principles can be introduced as primitive for a constant. But we also know from Belnap's equally famous reply (1961) that a conservative extension requirement over a deducibility relation provides the fundamental constraint to which such an objector can appeal. The limits of the legitimacy of such a conception of principles as self-justifying can, then, apparently be formulated without appeal to semantic notions. So how is this objection to be met?

It seems to me that this objection is well taken against a verificationist, but not against a realist. For the verificationist, to specify how a certain form of conclusion may legitimately be inferred from certain premisses *is* to specify how a means of verifying the premisses may be transformed into a means of verifying the conclusion. Simply appending an inference of the given form to a means of verifying the premisses yields a means of verifying the conclusion. The verificationist's conception is that the meaning of a sentence is to be specified by giving a means for establishing it. So for the verificationist, the possibility

of establishing by means of a certain inference the sentence which forms its conclusion can be a legitimate partial determination of the sense of that sentence. When the inference does partially determine the sense in that way, the inference is automatically validated.

On a theory on which a sentence's sense is not given by some (canonical) means of establishing it, things are very different. The realist cannot adapt the verificationist's means to his ends here. A realist cannot consistently accept as a sufficient condition for the validity of a principle of inference that any means of establishing its premisses can be transformed into a means of establishing its conclusion. For again, this would for the realist incorrectly validate the principle that from A can be inferred 'It's verifiable that A'. The only general condition of validity the realist can accept is (with refinements) the standard one that necessarily if premisses of the given form are true, so is the corresponding conclusion.

This means that in showing the legitimacy of introducing a constant as conforming to certain laws, the fundamental task for a realist must be to demonstrate that *there is* a semantical value for it which makes those laws necessarily truth-preserving. I will refer to the requirement that there exists such a value as 'the Semantic Constraint'. This Semantic Constraint may in some circumstances amount to certain proof-theoretic conditions upon the introduced laws. But the importance of any proof-theoretic requirements is, for the realist, derivative from their relations to the Semantic Constraint.

For the realist, then, the fundamental objection to Prior's runabout inference ticket *tonk* is semantical. *Tonk*, to refresh our memories, Prior introduced as conforming to the inferential principles $A/(A \text{ tonk } B)$ and $(A \text{ tonk } B)/B$. (I will follow the notational convention that $\{A_1 \dots A_n\}_{[S]} B$ means that B is derivable from the set $\{A_1 \dots A_n\}$ in the system S, with omissions of parts of the notation when there is no significant ambiguity.) The semantical objection to *tonk* is that there is no binary function on truth values which validates both its introduction and its elimination rules. Its introduction rule, if it is to be truth-preserving, requires that ' $A \text{ tonk } B$ ' is true when A is true and B is false. Its elimination rule requires that in the same case ' $A \text{ tonk } B$ ' is false; otherwise the rule will lead from truth to falsity.⁸ There is no coherent 'Do-It-Yourself' semantics for *tonk*.

⁸ I am thus in the end in agreement with Stevenson's (1960) reply to Prior. Stevenson, though, does not consider the possibility of Belnap's type of approach, and his reasons may not be the same as mine.

None of this is to deny that verificationists have some access to a notion of truth, or to deny that they can say that valid inferences must be truth-preserving. They do have such access, and should insist on truth-preservation. The point is rather that the notion of truth can, for them, be wholly elucidated in terms of verifiability; and this permits proof-theoretic relations to enter directly a specification of sense-giving verification conditions for sentences containing a new logical constant. For a realist there seems to be no comparable general explanation of truth using a notion which can similarly accommodate proof-theoretic relations in a specification of sense for a new kind of sentence. The realist cannot rephrase away the notion of truth in his statement of the constraint on legitimate stipulation.

What then *is* the relation between the Semantic Constraint and proof-theoretic requirements on the introduction of logical constants, and in particular the requirement of Conservative Extension? There are two cases to consider, according as there is or is not a semantically sound and complete formalization LC of the logic for the unextended L before a new constant is introduced.

Take first the case in which there is such a sound and complete logic LC. Suppose a new logical constant \S is introduced as conforming to a certain set of logical principles $P(\S)$. In this first case, we can argue that if the Semantic Constraint is met, then Conservative Extension is fulfilled. That is, we can argue that if there is a semantic assignment to \S which makes all the principles in $P(\S)$ truth-preserving under all assignments to its atomic components, then $P(\S)$ conservatively extends LC over L. Suppose then that there is some semantic assignment to \S which validates all the rules in $P(\S)$. Suppose too that there are formulae $A_1 \dots A_n, B$ of the unextended language L such that

$$\{A_1 \dots A_n\}_{/[LC+P(\S)]} B.$$

Since LC is sound and so are the rules $P(\S)$ under the given assignment to \S , in any model in which all of $A_1 \dots A_n$ are true, B is also true. By the completeness of LC,

$$\{A_1 \dots A_n\}_{/[LC]} B.$$

So Conservative Extension holds in this first case.⁹

In the second case, there is no sound formalization of the logic

⁹ Note that this argument rests on the tacit assumption that the semantical apparatus used in giving a semantical value to \S does not extend the kind of semantical assignment, or the particular semantical assignments of that kind,

for the unextended language L which is also semantically complete. The argument just given for the first case is then clearly blocked: from the fact that B is true in any model in which $A_1 \dots A_n$ are, we cannot conclude that

$$\{A_1 \dots A_n\}_{[LC]} B.$$

In this second case, however, it would also not be *rational* to demand Conservative Extension as a requirement on the legitimate introduction of a logical constant as conforming to certain principles. There ought to be no objection to the introduction of a new constant extending derivability on sentences of the original language if, by a semantics which overlaps with that of the original language, the newly endorsed derivations are sound.

A demand for Conservative Extension is sometimes defended by saying that the addition of new logical constants should not change the meaning of sentences in the original, unextended language. But if after the addition of a new logical constant and rules for it, all the sentences of the original language which become newly derivable are valid on the semantics for the original language, this kind of violation of Conservative Extension does not involve a change of meaning.

We should draw the conclusion that in the cases in which Conservative Extension is a rational requirement on the introduction of a logical constant, it follows from the Semantic Constraint that Conservative Extension will be fulfilled. It seems to follow that Conservative Extension has no independent status for a realist. It may be an obligatory requirement on conceptions under which the meanings of logical constants are elucidated in purely proof-theoretic terms; but I have argued that such conceptions should not be endorsed by a realist.

The type of incompleteness which makes Conservative Extension an irrational requirement for the realist is a possibility which is realized in the case of second-order logic (see for instance Boolos and Jeffrey (1974, pp. 204–5)). I certainly do not want to claim that our understanding of a notion of second-order validity which goes beyond provability in any recursively axiomatized system does not merit further philosophical investigation. It surely raises many fascinating issues. But anyone who is inclined to be sceptical that we have any such understanding will have to

made to \S -free sentences. If this assumption were not to hold, the soundness and completeness of LC on the semantics used for the unextended language tells us nothing about its soundness and completeness when \S is added and a new semantics is used.

account for the parallel, and connected, phenomenon of our apparent possession of a concept of arithmetical truth which outstrips provability in any recursively axiomatized system.¹⁰

6. *Validity: impressions, truth, knowledge*

Any philosophical theory of the validity of logical principles must face the question of whether it gives a satisfactory account of the relations between a thinker's impression that a principle is valid and the principle's really being valid. It is natural to write of recognizing implications.¹¹ Talk of recognition is appropriate, however, only when there are two distinct things, validity and the recognition of it, however intimate the relations between the two distinct things may be. So the questions which arise for any account are these: Does it endorse this intuitive distinction, or not? If so, what sort of objectivity does the property of validity possess? And are there any kinds of mistake about it which it is impossible to make?

Two indisputable principles, with which the positive account I have offered is consistent, are these:

Principle (1) The impression that a word is correctly applied is not what *makes* an application of it correct;

Principle (2) The impression that a word is correctly applied is not sufficient for an application of it to be correct.

These Wittgensteinian principles should be endorsed by anyone who agrees that a thinker can be under an illusion that a word has some meaning on his lips, when in fact it has none (Wittgenstein, 1958). There can be no correct applications of such a word; but the impressions of correctness will be present for a thinker suffering from the illusion. The general principles (1) and (2) apply to the words 'follows from' just as they do to any others.

Things are different when we consider the corresponding principles concerned with concepts. Maybe

Principle (3) The impression that a particular thing falls under a certain concept is not what makes an application of that concept to that thing correct

¹⁰ For the connection between the phenomena see Boolos and Jeffrey (1974, pp. 204–5 again).

¹¹ As Harman does (1986a, p. 131) in an account I consider in an Appendix to this paper.

is also true. But what of Principle (4), the concept analogue of (2)?

Principle (4) The impression that a particular thing falls under a certain concept is not sufficient for the correctness of an application of that concept to that thing.

Principle (4) is questionable. What it says is that for any concept C, the impression that a thing falls under C is not sufficient for C really to apply to that thing. (4) will be rejected by anyone who holds that if it seems to someone that he is in pain, then that subject is in pain. May there not be a similar principle, roughly along the lines that the primitive impression that $A \vee B$ follows from A is sufficient for it to do so?

Presumably there are substantive constraints on the existence of genuine concepts. To deny (4) is not to deny the existence of such constraints. We have already distinguished the apparently false (4) from the true Principles (1) and (2); insistence on Principles (1) and (2) does not carry us as far as (4). In fact, (4) can be denied, as I am denying it, only in cases in which there *is* a genuine concept which provides a counterinstance to its implicit universal quantification. An alleged binary extensional sentential connective, for instance, has to determine a truth function if it is to mean something. Prior's *tonk* fails this condition; as a consequence, it provides no way of getting a falsifying instance of (4).

In every case in which an impression of the applicability of a concept is sufficient for its really applying, we ought to give an explanation of *why* it is sufficient. Evans remarked (1982, p. 229) that in some cases, a judgement's having a certain content about a state 'can be regarded as constituted by its being a response to that state'. A sympathetic way to gloss this and keep it applicable to the content 'I'm in pain' is this: it is partially constitutive of a thinker's possessing the concept *pain* that he is disposed to judge the first-person, present-tense content that he is in pain precisely when (and for the reason that) he is in pain. There is a constitutive fact about concept-mastery which plays a similar role in the logical case. For each logical concept which displays the phenomenon, there are principles containing it such that it is partially constitutive of possessing that concept that the thinker has the primitive impression that those principles hold. Semantic values are assigned to the constants in such a way as to validate these principles. It follows that in these logical cases, the impression that such underived principles hold is sufficient for them to hold.

Still, it may be objected that we have no interesting analogy with *pain* here. For we have in effect taken as a counterexample to (4) cases in which the objects in question are a sentence *B* and a set of sentences $\{A_1 \dots A_n\}$ and *C* is the concept *follows from*. But once we have fixed on particular sentences, *B* either follows from this set or it does not. If it does, what is the point of considering a conditional 'If a thinker has the impression that *B* follows from $\{A_1 \dots A_n\}$, then it does'? For if it does follow, then the antecedent is redundant; and if it does not, the truth of the antecedent will not be sufficient for it to follow.

Of course these last points are true. But they are consistent with the main point I want to make. The main point is that we can *reason from* the presence in a thinker of a certain type of impression that a principle is valid, together with further information, to the conclusion that it is valid. This is parallel in the relevant respects with the case of pain. In the logical case, the further information consists of the premise that the logical constants in *B* and in $A_1 \dots A_n$ meet the Semantic Constraint, together with the outlined theory of what gives a constant its semantic value. In the pain case, the further information consists of a theory of what gives a judgement the first-person content 'I'm in pain'. If our imagined objector still insists that we offer a significant universally quantified conditional which reflects the situation I claim to exist, then I offer him this:

For any $A_1 \dots A_n$, *B*, if the thinker finds it primitively obvious that *B* follows from $A_1 \dots A_n$, and if the logical constants in *B*, $A_1 \dots A_n$ all meet the Semantic Constraint, then *B* does follow from $A_1 \dots A_n$.

Neither one of the antecedents of *this* universally quantified conditional is redundant.

The fact that it strikes a subject that he is in pain is not what *makes* him in pain. The impression that one is in pain requires possession of the concept of pain, and is a fact at the level of thought, which the fact of being in pain is not. Similarly the primitive impression that a certain primitive principle is valid is not by itself what *makes* it valid, not even when the Semantic Constraint is met. What makes it valid is that it is necessarily truth-preserving under the proper assignment of semantic values to its logical constants; though of course what makes an assignment proper depends in part on fact about impressions, if the preceding is right. Finally in this cluster of points, nothing in the present account entails that a thinker is ever necessarily in a

uniquely favourable position to assess whether one form of content follows from another. This is the analogue of the point that acceptance of infallibility about whether one is in pain does not commit a theorist to being a private linguist in Wittgenstein's sense. Others can know that, as he intends the expressions, one sentence follows from a second in a particular individual's language.

Some of these points can be applied to a discussion in Crispin Wright's book on Wittgenstein's philosophy of mathematics (1980, pp. 353–8). There Wright attacks the idea that there can be such a thing as 'reflecting on the content of a sentence and *thereby* coming to know that it cannot but be true' (1980, p. 353). In his discussion, Wright is at some points concerned with coming to see that the sentence in question is necessary; at others he mentions the weaker claim that the sentence can be seen to be true by such reflection, given the meanings of its parts.¹² In any case, I will concentrate on the latter formulation.

Since they present rather different challenges, it will also be helpful to discuss separately the situation of a philosophically reflective subject, who reasons from a philosophical account like that in sections 2–5 above, from the situation of the unreflective language user. We can call these two representative individuals the sophisticate and the ordinary citizen, respectively.

Suppose the account in sections 2–5 is broadly correct. Is there anything in Wright's discussion which prevents the sophisticate from applying that account to himself? Is there anything there which prevents him reasoning from his finding certain principles primitively obvious, via that account, to their being truth-preserving? Wright's central point is that

whatever sincere applications I make of a particular expression, when I have paid due heed to the situation, will seem to me to conform with my understanding of it. There is no scope for a distinction here between the fact of an application's seeming to me to conform with the way in which I understand it [i.e. the expression—CP] and the fact of its really doing so. (1980, p. 355)

¹² At some points, the two ideas occur in the same sentence, and it may be Wright's view that on the intuitive conception he is concerned to attack the two are linked. He writes, summarizing his discussion: 'we have been concerned with the intuitive view that the phenomenon of necessity reduces to that of the occurrence of analytic sentences, the analyticity of a sentence consisting in the circumstance that it expresses a truth just in virtue of the senses of its constituent expressions and its syntax.' (1980, p. 357)

This point establishes that the impression that a word applies cannot be what makes it apply, or even be sufficient for its correct application. What is established by it is just the conjunction of Principles (1) and (2). It does not establish that there is no concept such that the impression that something falls under it is sufficient for that thing to fall under it; that is, it does not establish (4). An argument which tells only in favour of (1) and (2) and not in favour of (4) is not in conflict with the possibility of someone's reasoning from the account of sections 2–5 to the validity of certain principles.¹³

The case of the sophisticate, however, cannot be fundamental. The fundamental case is that of the ordinary citizen. This is so because any principles the sophisticate relies upon in the meta-language in trying to apply the account of sections 2–5 to himself or others must be ratified as knowledge in some other way. Knowledge of *them* needs to be treated in the same way as that of the ordinary citizen; so let us turn to this more basic case.

The most acute problem about the ordinary citizen arises in the case of primitive logical principles, acceptance of which is partially constitutive of understanding some logical constant in them. We want to say the ordinary citizen knows these principles. But he has not inferred them from anything, and he does not know any philosophical account from which he might infer their truth. Even if he did, as we just noted, he would have to rely on some logical principles which could not be ratified as known in that way, on pain of infinite regress. The challenge, then, is to say, why the ordinary citizen still knows (instances of) these primitive logical principles.

In earlier work, I considered what I called a Model of Virtual Inference. According to the Model, a belief is knowledge if there is, for the theorist, a knowledge-yielding abduction from an explicit statement of the subject's reasons for the belief, together with other information available to the subject, to the truth of the belief (Peacocke, 1986, p. 163). The model was considered as a possible account which would ratify a subject's beliefs as

¹³ Wright is of course aware of the difference between (1) and (2) on the one hand, and (4) on the other. He writes that 'it is not to be insisted that objectivity requires an ineliminable possibility of a misapprehension. What it requires is that the facts do not *consist* in one's having a certain impression of them' (1980, p. 356). But he says of accounts which try to explain, consistently with the Wittgensteinian constraints, why in certain cases a misapprehension is impossible, that 'an account of this kind is not going to be forthcoming' (p. 357). I have been trying to offer such an account.

knowledge in certain cases even if the subject does not possess the conceptual resources to frame the knowledge-yielding arguments for the truth of the belief. Can this model help in the present case?

It is hard to see how it can. Presumably the explicit statement of the subject's reasons for accepting certain primitive logical principles would be their primitive obviousness. Someone trying to apply the Model of Virtual Inference might then aim to give a knowledge-yielding abduction: from these factors and a philosophical theory of the determination of the semantic values of the logical constants, to the validity of the principles in question. But the problem with the approach is that the explicit statement of the factors rationally influencing the subject seems redundant. The subject finds it primitively obvious that (say) a content of the form A or B follows from A. But if the theorist is using 'or', he will already find certain principles involving it obvious—he will certainly not need information about another person's impressions to support them. This is very different from the way in which the Model of Virtual Inference applies elsewhere. The premisses about the factors influencing a subject when he makes a judgement based on his perceptual experiences cannot be omitted by the theorist in an argument to the truth of the belief.

One could consider variants on the Model of Virtual Inference, but applied to the logical case, they begin to seem like epicycles, and to threaten the uniformity of the Model. The general problem presented by the ordinary citizen's knowledge of primitive logical principles is that of providing a sufficiently close connection between three things: (i) a sufficient condition for knowledge which is general in the sense of having a rationale derived from general epistemological considerations, rather than working simply *ad hoc* for the logical principles; (ii) the philosophical arguments of the kind in sections 2–5 that in some cases the impression that a principle is truth-preserving is sufficient for its really being so; and (iii) what actually goes on in the mind of an ordinary citizen when he accepts a primitive logical principle.

An alternative approach to this problem is suggested by the partial parallel which has helped us before, the parallel with the concept *pain*. It is plausible that when the ordinary citizen judges that he is in pain, and for the reason that he is experiencing pain, he knows he is in pain. I suggest that the belief is knowledge because an account of mastery of the concept *pain* would require him to be disposed to judge that he is in pain in these circumstances. Can we generalize this connection between an account of

mastery and correct attributions of knowledge in such a way that it also captures the ordinary citizen's knowledge of (instances of) logical principles?

In the context of the claims I have been defending, a connection between reasons mentioned in an account of mastery of the concepts in a given content and knowledge of that content would not be at all *ad hoc*. It would have the following general rationale. In *Thoughts*, I conjectured that the normative acceptance conditions for a content determine its truth conditions. Suppose for the moment this conjecture is true. Now suppose too that certain reasons for judging a content are, according to an account of mastery of the concepts in a given content, conclusively sufficient for judging it. If the conjecture of *Thoughts* is true, then one would expect the content to be true when a thinker judges it for such conclusive reasons. It seems that the content could fail to be true in such a case only if either the conjecture is false or the reasons are not conclusively sufficient. But if the reasons are sufficient in this way to guarantee the truth of the content in question, it will not be surprising if they also yield knowledge of the content accepted. They are reasons of the best kind.

In order to give a general formulation of the link with knowledge, it will help if we draw a preliminary distinction. We can distinguish two forms which may be taken by an account which purports to say what individuates a given concept. ('Concept' here means mode of presentation of any category.) The two forms are intimately related to one another, but it is just one of the two which we need to employ in discussing the link with knowledge. The first form, with some simplification, can be called the *attributional*, and is this:

concept \emptyset is that concept C for which a thinker must meet condition A(C) in order to possess \emptyset .

The accounts of understanding the logical constants in the earlier sections of this paper were formulated as attributional accounts. An attributional account can mention reasons for judging or taking other attitudes to contents containing C; but these reasons will be those of the thinker in question. In general, an attributional account may presuppose that the thinker already possesses certain other concepts.¹⁴

¹⁴ An attributional account may also deal with the individuation of several concepts simultaneously. In cases in which there are local holisms, it would have to do so if it were to have a chance of success. So the more general form would be:

concepts $\emptyset_1 \dots \emptyset_n$ are those concepts $C_1 \dots C_n$ for which a thinker must meet condition A(C_1, \dots, C_n) in order to possess $\emptyset_1 \dots \emptyset_n$.

By contrast, a *direct* account has the form:

concept \emptyset is that concept C which rationally requires such-and-such attitudes to so-and-so contents containing C in specified circumstances.

The whole condition following 'rationally requires' can be schematically abbreviated 'D(C)'. Any reasons referred to in a direct account will be reasons not for some thinker referred to in a third-personal say, but reasons for *us* if the account is correct and we are to use the concept it treats. A direct account aims to state rational requirements whose status as such depends only on which concept it is that is in question; it aims to capture all such requirements; and it aims to capture only rational requirements to which a thinker must be sensitive if he is to be counted as possessing the concept. As that last point suggests, there need be no competition between direct and attributional accounts of what individuates a given concept. Ideally, a theorist who wishes to use both should say how to pass from a correct attributional account to a direct account, and *vice versa*. In the above treatment of the logical constants, such a transition would be straightforward. Principles which the attributional account requires to be found primitively obvious are ones whose premisses give reasons mentioned in a direct account for judging their conclusions (or reasons for rejecting them in cases where impressions of incompatibility are mentioned). As in the attributional case, direct accounts may take for granted that we already possess certain other concepts; they may also be given for a family of concepts simultaneously.

We need to concentrate for present purposes on direct accounts, because it is they which identify certain reasons for judging a content as rational. For a given concept \emptyset , then, consider a good direct account of its individuation which supplies a condition D(C) of the sort mentioned in the previous paragraph. This direct account may state that the holding of certain conditions gives a conclusive reason for judging some content P(\emptyset) containing the given concept \emptyset . Where \emptyset is the concept *pain* and P(\emptyset) is the content 'I'm in pain', the subject's being in pain plausibly gives such a conclusive reason.

We can now relabel the conjecture of *Thoughts* that normative acceptance conditions determine truth conditions as the *First Conjecture*. I then suggest the following two-part *Second Conjecture*:

(a) Take any contents $p_1 \dots p_n$ which a correct direct account of \emptyset says are conclusively sufficient for rationally judging that P(\emptyset); then if a thinker judges that P(\emptyset) follows from $p_1 \dots p_n$,

and does so because he finds it primitively compelling that it does, then he knows that $P(\emptyset)$ follows from $p_1 \dots p_n$.

This part (a) ratifies the ordinary citizen's belief that something of the form $A \vee B$ follows from A as knowledge; for a correct direct account of alternation will say that acceptance of A gives a conclusive reason for judging that $A \vee B$, if the question arises. If we hold that what is inferred from known premisses by principles which the thinker knows to be valid is also known, then part (a) of the Second Conjecture also ratifies as knowledge conclusions the ordinary citizen draws from known premisses by primitively obvious inferential principles.

A symmetrical treatment can be given for knowledge of instances of elimination rules. The role of conclusive canonical grounds in (a) will in the elimination case be played by indefeasible canonical commitments.

Part (b) of the Second Conjecture covers the cases like that of the content 'I'm in pain'. Such cases are not captured by part (a), for in these cases the judgement is not inferred from some other contents which are already accepted. The thinker's mental state itself—rather than a content concerning it—gives the thinker a reason for making the judgement:

(b) Take any mental state of the thinker which a correct direct account of \emptyset says is conclusively sufficient for rationally judging a content $P(\emptyset)$: when a thinker judges that $P(\emptyset)$ because he is in that state, then he knows that $P(\emptyset)$.

This Second Conjecture is general, in that its two parts are conjectures about an arbitrary concept \emptyset . We also suggested a general rationale for it which does not turn on properties special to the logical case which originally motivated the discussion. I further suggest it may be a principle we need to use elsewhere in epistemology.¹⁵

¹⁵ Consider, for instance, a subject's belief that a presented object is cylindrical, a belief based on a perception with a fine-grained analogue content (in the sense of my 1987b). This belief can in suitable circumstances be knowledge. I suggest it can be so because the Second Conjecture is true. The Second Conjecture can be applied to a direct account of mastery of *cylindrical* when that account links the mastery to perceptions with certain non-conceptual, analogue representational contents.

7. *Explaining logical beliefs*

Wittgenstein was opposed to the view that logic is some kind of ultraphysics. He wrote (1978, I.8):

But still, I must only infer what really *follows*!—Is this supposed to mean: only what follows, going by the rules of inference; or is it supposed to mean: only what follows, going by *such* rules of inference as somehow agree with some (sort of) reality? Here what is before our minds in a vague way is that this reality is something very abstract, very general, and very rigid. Logic is a kind of ultra-physics, the description of the 'logical structure' of the world, which we perceive through a kind of ultra-experience (with the understanding, e.g.).

Given the tone of this passage, it is no surprise that in later sections Wittgenstein makes clear his opposition to this conception. The conception is far from clear; but even from this brief characterization of it, enough can be extracted for us to give an argument that the account I have offered places me in agreement with Wittgenstein's rejection of logic as ultraphysics.

The fact that something is square can causally explain a subject's experience of it as square. Now take a primitive logical principle, a principle such that to find it primitively obvious is partially constitutive of understanding the expressions it contains. On the present account, one ought not to try to explain causally the fact that a thinker finds it primitively obvious that a certain principle is valid by citing the fact that *is* valid. The principle's being valid consists in its being truth-preserving under all relevant assignments; it is truth-preserving under all relevant assignments in part because of the semantic values given to the logical constants it contains; and in turn these constants receive their semantic values in part because the thinker finds it primitively obvious that the given conclusion follows from the premisses. So in these primitive cases, it would be better to say that the principle's validity consists in part (though of course not wholly) in an impression that it is valid; and this seems incompatible with its validity causally explaining the impression. The failure of the claim that validity causally explains impressions of validity gives a limited sense in which these primitively obvious principles do not hold independently of our impression that they hold. This limited sense is analogous to that in which, on contemporary accounts of secondary qualities, the redness of a perceived object does not hold independently of a normal perceiver's experience of it as red in normal circumstances.

I have in effect declined the option of explaining in a different way how on the present account logic is not a form of ultraphysics. This different way would involve pointing out that on the present theory, a statement of what it is to understand a logical constant needs to make reference to some of the thinker's psychological states, viz. his primitive impressions of what follows from what. This reference is to a psychological state which would not similarly be mentioned in an account of possession of an arbitrary concept; it is rather specific to the logical constants and to some others. On this way of developing the point, it would be agreed that some conclusions can follow from premisses without striking us as following. But this, it would be said on this other way of developing the claim, is analogous to the point that an object can be red and yet not be experienced by us as red (when it is unperceived, or when a subject is misperceiving). These facts about redness are obviously consistent with the view that an account of understanding the word 'red' must make reference to human experiences of objects as red, as again recent (and seventeenth-century) accounts of secondary quality words have emphasized.

The difficulty with developing the contrast with the 'ultraphysics' view along these lines is that it is plausible that an account of understanding the primary quality word 'square' would have to make some reference to experience-types. We should remember the difference between the sense of 'square' and that of '(regular) diamond-shaped'.¹⁶ The concept expressed by 'square' is a mode of presentation of a shape property. This mode of presentation may enter the representational content of experiences in more than one sense modality; it can, for instance, be used in giving the representational content of tactile experience. But it is still to be distinguished, on classical Fregean grounds, from the mode of presentation expressed by '(regular) diamond-shaped'; that the two shapes (considered independently of their orientations) are the same can be informative. Certainly creatures with very different sensory systems from ours could experience things as square. But this shows just that some features of the representational content of experience can be invariant across radically different sensory systems and across other differences in experience-type. An experience of something as square is still to be distinguished from an experience of something as diamond-shaped, even for such a radically different creature. The features that are common to our experience and that of such different

¹⁶ See the discussion in Peacocke (1987b).

creatures would be the ones captured in Thomas Nagel's projected, but as yet somewhat neglected, subject of 'objective phenomenology' (1979). These doubts about this alternative way of grounding the points required for pointing up the agreement between the theory I have outlined and Wittgenstein's remarks about 'ultraphysics' do not call in question the point that an object's being square can, in a suitable context, causally explain an experience of it as square; which is why I have used that way of stating the contrast.

8. *Meaning and truth amongst the logical truths*

I turn now to consider the complex issue of the relation between the meaning of the logical constants and the truth of logical truths containing them. We can consider the issue by addressing these questions, of which answers to the first two will be preliminaries to answering the third:

(i) Does the account given so far entail this: that the truth values of logical truths containing particular logical constants are determined once their meanings are given? Any such claim of determination we can label a *Determination Thesis*. Part of the task in answering the question is to provide a sharp formulation of a Determination Thesis.

(ii) Does any form of Determination Thesis which follows from the account so far give a property which is unique to the logical constants?

(iii) Does the present account entail that logical truths are 'true purely in virtue of their meaning', and thus become vulnerable to Quine's famous attack on Carnap?

Some of those who in the past have wanted to argue for a form of Determination Thesis have done so because they wanted to reconcile the existence of *a priori* knowledge with one or another form of empiricism or positivism. But that motivation is not compulsory. A theorist's concern can be with the relations between the meaning and truth of sentences containing logical constants in their own right. In fact what I will have to say does not eliminate the notion of *a priori* knowledge, but relies on it.

Not all the well-known defenders of a Determination Thesis were obviously empiricists or positivists. But even in the case of those who were not, the arguments they offered for the Thesis often involved elements inessential to the core conception which one would expect to underlie acceptance of the Thesis. This core is that it is a consequence of a correct statement of the meaning

of the logical constants that certain sentences containing them are true.

Perhaps the clearest case of inessential accretions is the view of Wittgenstein in the *Tractatus*. The claims that tautologies say nothing and that all propositions are truth functions of elementary propositions are not obligatory for someone endorsing the core conception (although they may have been unavoidable given the theoretical context endorsed by the early Wittgenstein). The core conception can consistently be held along with a firm sense/reference distinction. It can also be held along with something Wittgenstein did not supply in the *Tractatus*, an account of what makes it the case that a given symbol expresses a particular truth function. But despite the accretions and gaps, Wittgenstein in the *Tractatus* remains an inspiration for a defender of a Determination Thesis, as giving an explicit implementation of the core conception.

At first blush, the claims for which I have been arguing straightforwardly entail a Determination Thesis. The semantic values of the logical constants of a language in use are assigned in virtue of their validating certain principles, in the way described earlier. It follows from that account that these standard axioms are true, and can be known to be so given just the senses and corresponding semantic values of the object language '&', 'v' and '~':

$$\forall A \forall B (\text{True}(A \& B) \equiv (\text{True}(A) \& \text{True}(B))) \quad A_1$$

$$\forall A \forall B (\text{True}(A \vee B) \equiv (\text{True}(A) \vee \text{True}(B))) \quad A_2$$

$$\forall A (\text{True}(\sim A) \equiv \sim \text{True}(A)) \quad A_3$$

With classical logic in the metalanguage, we can prove as a theorem within a truth theory containing axioms A_1 – A_3 such sentences as (L):

$$\forall A (\exists B \exists C (A = \sim (B \& \sim (B \vee C))) \supset \text{True}(A)). \quad (L)$$

On the other hand, using just the semantical axioms, including a denotation axiom for the proper name 'Burma' and a satisfaction axiom for the predicate 'is humid' (and the axioms in the theory about concatenation and sequences), we cannot prove in the truth theory that the object language sentence 'Burma is humid' is true. Is there any obstacle to generalizing this point to all object language logical truths? And can we then take this point as an explication of a Determination Thesis on which that Thesis is a true claim about the meaning of the logical constants?

The explication certainly needs some refining. We have to

remember again that for some logics, such as second-order logic, there is no complete formalization. As a result, there will always be some logical truths in a second-order object language for which we will not be able to prove a corresponding version of (L) above, if the logic in the metalanguage is recursively axiomatized.

When a logic is not completely formalizable, it is no way out of the present problem to add to the truth theory any *a priori* principle in the metalanguage. There is no guarantee that every second-order formula which is true in all models is knowable at all, let alone knowable *a priori*. When the logic for some language is completely formalizable, there is such a guarantee in principle, since a proof procedure will eventually yield a proof of any valid sentence. Incompleteness precisely removes this guarantee.

On the other hand, suppose we allow the addition to the truth theory of principles which are not *a priori*. Then the resulting explication of the Determination Thesis states a relationship between the meaning of the logical constants and logical truths which equally holds between the meanings of 'Burma' together with 'is humid', and the object language truth 'Burma is humid'. For if we allow the addition to the truth theory of the *a posteriori* truth that Burma is humid as an axiom of the metalanguage, then we can certainly prove in the enlarged theory that the object language sentence 'Burma is humid' is true. With such additions, we could hardly hold that the resulting Determination Thesis tells us anything interesting about the logical constants: it would state a relation which holds between the meaning of any expression and a truth containing it.

A tighter relation would be given thus:

(LT) Take any logical truth, picked out by a structural description *s*, say. Then 'True(*s*)' is a logical consequence, in the model-theoretic sense, of the correct truth-theoretical axioms dealing with the logical constants in the object language sentence described by *s*, together with some *a priori* truths.

This sidesteps the incompleteness problem by using logical consequence instead of a provability relation in characterizing the consequences of the axioms of the truth theory. *One* way one can show that the condition (LT) is met for a particular object language sentence is by deriving its truth from the axioms of the truth theory in some sound formal logic. But, for a given formal logic, there does not have to be such a derivation in it for the condition (LT) to be met.

We will take (LT) as our formulation of a Determination Thesis which follows from the earlier account of understanding logical constants. It will equally follow from any other account of understanding which certifies such standard axioms as A1–A3 as true and as meaning-giving in some suitable sense. The formulation continues to use the notion of the *a priori*.

Our question (ii) is now whether (LT) gives a property which is unique to the logical constants. It does not. The truth of many object language arithmetical truths, for instance, will be logical consequences of the truth-theoretic axioms dealing with arithmetical vocabulary, together with suitable *a priori* principles in the metalanguage. The relation between the meanings of the logical constants and logical truths given by (LT) is at best privileged, rather than unique. It is privileged in the sense that it is not universal. It still does not apply to the relation between the meanings of 'Burma' and 'is humid' and the truth 'Burma is humid'.¹⁷

As one last preliminary to considering the third question, it will be helpful to invoke a distinction drawn by Wittgenstein. He wrote: 'What I always do seems to be—to emphasize a distinction between the determination of a sense and the employment of a sense' (1978, III.37). In Wittgenstein's terminology, the account in sections 2–5 of this paper is an account of how the senses of the logical constants are determined. In the axioms A1–A3, on the other hand, we simply *employ* in the metalanguage an expression with the sense thus determined, in stating the uniform contribution made by particular object language expressions to the truth conditions of sentences containing them. We have to give such a uniform statement if we are to meet the requirement of showing that there is a single sense of 'or' used both in logical principles found primitively obvious and in such empirical truths as 'The plane stopped in Cyprus or Malta'. As usual, we will also want to distinguish, amongst the theorems of a theory of truth, some T-sentences as meaning-giving: these sentences on their right-hand sides say the same as the object language sentences described on their left-hand sides.

If we regard the relation between a substantive theory of what

¹⁷ Even this conclusion relies upon not altering 'logical consequence' in (LT) to something more specific to Burma and humidity when considering its analogue for 'Burma is humid'. Since I will be arguing against the idea that logical truths are true purely by virtue of meaning, I will not press this point further. But to someone who does accept the idea, the point is a serious obstacle (amongst others) to trying to use (LT) to defend it.

individuates the sense of a logical constant and the truth-theoretic axiom for that constant in this way, it becomes clear that the present position ought to be insisting upon one of the points Quine made against Carnap in his formidable essay 'Carnap and Logical Truth'. At one point, Quine formulated the linguistic doctrine of logical truth to which he objected as the doctrine that logical truths are 'true by virtue purely of the intended meanings, or intended usage, of the logical words' (1976, p. 110). Against the doctrine, Quine remarked that logical truths, like any other truths, are true partly in virtue of meaning and partly in virtue of the way the world is; even such a truth as $\forall x(x=x)$ 'depends on an obvious trait, viz. self-identity, of its subject matter, viz. everything' (1976, p. 113). On the position I have defended so far, each closed object language logical truth, like any other (closed) object language sentence, has a meaning-giving T-sentence. A logical truth, also like any other, is true just in case what it says is so; and what it says is given by the right-hand side of its meaning-giving T-sentence. All this supports Quine's point. The point is in no way undermined by a theory which uses a notion of sense and gives a special role to the primitively obvious in determining the sense and semantic value of individual logical constants.

The point is also consistent with the fact that there are ways of proving the truth of object language logical truths from the axioms of a truth theory other than by applying *modus ponens* to one direction of the meaning-giving T-sentence. We can, for example, reproduce in a truth theory the intuitive reasoning for each line of a truth table to establish that an object language tautology is true. The proof need not proceed via a derivation of the tautology's T-sentence. But similarly consider a sentence which is not a logical truth: there are ways of establishing in the theory that it is true other than by applying *modus ponens* to one half of its meaning-giving T-sentence. There are many ways in which we may have, in the metalanguage of the truth theory, principles or premisses which are sufficient for the obtaining of a sentence's truth condition. When we have principles or premisses which are sufficient, the derivation of the sentence's truth may take many different forms.

Does this mean the present position is fully compatible with Quine's in 'Carnap and Logical Truth'? Matters are not so simple, however. In that essay, Quine offers at least three nonequivalent formulations of the linguistic doctrine of logical truth. Besides the formulation already quoted, Quine also describes the doctrine as stating that a logically true sentence 'is a

sentence which, given the language, automatically becomes true' (1976, p. 108). This second formulation seems to me to state a truth if two conditions are met. The first condition is that (LT) is correct. The second condition is that if one sentence is a logical consequence (in the model theoretic sense) of another, then it is legitimate to say that the first is automatically true if the second is. Under these conditions, a logical truth *is* automatically true given the language.¹⁸

Quine's third formulation is that 'the truths of logic have no content over and above the meanings they confer on the logical vocabulary' (1976, p. 109). This by contrast is equally rejected by the present position. The logical truths do have content, fixed by the sense of their constituents and their mode of combination. I have been arguing that a sense-conferring role for certain principles of inference is not at variance with all the sentences which comprise their instances possessing structured senses and truth conditions.

A different objection to any Determination Thesis is that the logical constants of the classical logician and the logical constants of the intuitionist have the same meanings. This view may be held because of the substantial overlap of the principles they accept as valid. This is an argument given by Putnam (1975). It may also be held because some rationale is offered for distinguishing a proper subset of the classical laws as those guaranteed by the meaning of the logical constants. This is the view of Quine in *The Roots of Reference* (1974). His rationale concerned properties of the 'verdict tables' for the logical constants.

It matters that this Claim of Common Meaning for the constants of classical and intuitionistic logic is a thesis about the fundamental way of individuating their meanings. Any theory of the logical constants which uses the notion of meaning at all is likely to be able to characterize some notion of closeness of meaning and of overlap of meaning. A truth-conditional Fregean sense theory may be able to make some sense of a closeness relation which holds between classical constants and others whose valid principles substantially overlap with those which are classically valid. The Claim of Common Meaning is saying more than that. It is saying that on the fundamental way of individuating the sense of a logical constant, whatever it is, the

¹⁸ It does not follow that the truth of the sentence does not require the world to be a certain way. All that follows is that any condition of the world required by the truth of the sentence is one that automatically obtains.

constants of the classicist and the intuitionist have the same sense. This Claim is in conflict with the account I have offered at several points. The account of the sense of particular classical logical constants certainly gives them a different sense from those of the intuitionist; and one who accepts the Claim of Common Meaning must reject any Determination Thesis.

The problem for the defender of the Claim of Common Meaning is in giving an answer to the following question. How are we to demonstrate the invalidity or the correctness of some alleged logical truth when it lies outside the common core whose validity is, according to him, guaranteed by meaning? We would normally appeal to a semantics, together with either a counterexample within that semantics, or a proof which is sound according to that semantics. But there is, for the defender of the Claim of Common Meaning, no semantics which can provide a rationale outside the common core. If a semantics applicable beyond the common core has a good rationale given the meaning of the logical constants it treats, then the Claim of Common Meaning is being rejected. If it does not have such a rationale, what makes a method of assessing validity or invalidity outside the common core correct?

This dilemma should be found compelling by anyone who thinks there can be such a thing as a rationale for accepting or for refusing to accept logical principles outside the common core. Putnam at one point remarks that the claim that the classical and intuitionistic constants have different meanings could at most be 'a point about the philosophy of *linguistics* and not the philosophy of *logic*' (1975, p. 190). But the point just made is one about philosophical logic, about the ratification of validity or invalidity. Far be it from me to debar either linguistics or the philosophy thereof from using model theory! But points about what kind of *rationale* should be available for justifying assertions of validity or of invalidity are broadly the concern of philosophical logic, in the sense that they can be settled only by considerations relating to the nature of meaning for actual and possible logical constants in general.

Before moving on, I want to note some other relations of the present account to that of Quine. There are several areas in which the account I have offered overlaps with Quine's. I have in effect applied a form of the Principle of Charity in assigning semantic values to logical constants; and in *Word and Object* (1960) it was in connection with the logical constants that Quine first endorsed the Principle of Charity. Another area of overlap

acceptance, in some writings, of a form of inseparability thesis, 'the inseparability of the truths of logic from the meanings of the logical vocabulary' (1976, p. 109). But, even after we have factored out differences which result from the fact that I am operating at a level Quine would view with great suspicion, the level of sense, this agreed inseparability thesis still rests on very different grounds in Quine's thought and in the present account.

In the fifties, Quine wrote that 'Deductively irresolvable disagreement as to a logical truth is evidence of deviation in usage (or meanings) of words'. His reason was that 'elementary logic is obvious or can be resolved into obvious steps' (1976, both quotes p. 112). If what I have been saying is right, in that last sentence Quine gives as a reason something which is not in fact true. The point is not just that Quine is insisting on an alleged 'behavioural sense' of obviousness.¹⁹ He does insist on that; but his position would be open to objection even if his account were formulated with a more intuitive notion of obviousness. The objection is rather based on the earlier arguments I gave that some underived valid principles for a given logical constant need not be found primitively obvious by one who understands it. These principles are not 'potentially obvious' in the sense Quine uses this phrase, viz. of being derivable by obvious principles from obvious primitive principles. Not all of classical first-order logic is then 'potentially obvious' in Quine's sense. It is nevertheless still rational to accept (EE), say, after reflection. It is not clear how Quine could give a plausible account of why it is rational. It is not plausible that the rationality of accepting it emerges only upon considering the contribution of (EE) to total theories in which it is embedded.

9. *Informative, justified deduction for realists*

Finally I want to outline a realist's response to the challenge to show simultaneously both the utility and the justifiability of deduction. The response can be built up by contrasting it with Dummett's classic answer on behalf of the verificationist (1978).

Dummett draws on the idea of the *most direct* method of verifying a sentence. The most direct method of verifying a sentence is that which naturally corresponds, step by step, with

¹⁹ 'I must stress that I am using the word "obvious" in an ordinary behavioural sense [*sic*], with no epistemological overtones. When I call " $1+1=2$ " obvious to a community I mean only that everyone, nearly enough, will unhesitatingly assent to it, for whatever reason' (1970, p. 82).

the semantically significant syntactic structure of the sentence. The most direct method of verifying a conjunction is by verifying its two conjuncts; the most direct method of verifying an existential quantification is by verifying some instance. Dummett notes that the most direct method of verifying a sentence need not be the shortest. A logical truth of the propositional calculus may be verified most quickly by a proof, so that there is no need to decide the truth value of its atomic components. This brief characterization of the most direct method leaves much unresolved, but it is all we need at present.

This conception of the most direct method plays three roles in Dummett's thought. First, it helps to meet an obligation to show how in grasping a sense, one is grasping something which determines a referent. In any theory built along Fregean lines to the extent of distinguishing between sense and semantic value, this obligation is incurred at several levels. It is incurred at the level of the senses of atomic expressions, and at the level of semantically significant structured components, up to and including whole sentences. It is clear how, for one of a verificationist cast of mind, the idea of a most direct method can assist in meeting this obligation at the level of whole sentences. A method of verification, when carried out, can determine a truth value. If to grasp the structured sense which a given sentence expresses is to know its most direct method of verification, then in grasping a sense a thinker is grasping something which will, together with the world, determine a truth value. Since the method of verification would presumably proceed via the identification of objects as referents and by investigation of their properties, this conception can equally hope to meet the obligation at the subsentential level.

This first role of the most direct method is not one I will be discussing here in its own right, except to make the following remark. Suppose the First Conjecture, that normative acceptance conditions determine classical truth conditions, is correct. Then there is the possibility of developing a non-verificationist theory on which, at the level of sentences, grasp of sense is grasp of something which is, together with the world, capable of determining a truth value.

It is the closely related second and third roles of the most direct method which I will discuss in slightly more detail. The second role it has in Dummett's thought is that of showing how two logically equivalent sentences can have distinct senses. *That* they may do so is not controversial, given the cognitive character of

Frege's notion of sense. A task for any substantive theory of content is to show that it ensures, in its treatment of the logical constants, that this is so in the cases in which, pretheoretically, it is so.

The second role of the conception of the most direct method depends then on the claim that the senses of two sentences differ if the most direct methods of verifying them differ. It can require reasoning for a subject who judges something of the form $p \& (q \vee r)$ to realize that he can with equal reason also judge that $(p \& q) \vee (p \& r)$. On Dummett's treatment, the senses here are distinct because the method of first discovering whether p and then whether $(q \vee r)$ is distinct from the method of discovering whether one out of $(p \& q)$ and $(p \& r)$ hold.

We need a non-verificationist replacement for this second role of the most direct method. There is a salient realistic alternative when we remember this: that even though verificationist and realistic theories diverge radically in the way they try to meet the obligations of a theory of sense, they can share a common structure. Both can aim to characterize some core notion of understanding a sentence, something which has a life outside the context of the philosophical theory of deduction, and then try to put it to work inside that context. The details of the characterization of the core notion will fix the range of materials which can be drawn upon in constructing a philosophical account of deduction.

Suppose we have endorsed a theory of grasp of two logically equivalent contents with this property: that, according to the theory, a thinker can meet the condition for one of the two to be the content of his judgement without meeting the condition for the other to be so. Then we will have an account of how a thinker can judge a content without judging a second content to which it is logically equivalent. So if the realist has proposed a theory of the grasp of the constituent senses of two logically equivalent contents, he already has the materials for offering a realistic replacement for the second role of the most direct method. As an example, consider two contents of the forms $\exists x(Fx \supset p)$ and $(\forall x Fx) \supset p$. Intuitively, a thinker can judge either one without judging the other. On the account I would offer on behalf of the realist, what makes it the case that the first content is the content of a thinker's judgement is that he is committed to it by an arbitrary singular content of the form $Ft \supset p$. What makes the second content the content of a judgement is, amongst other things, that it is a content from which, together with a premiss

$\forall xFx$, the thinker is willing to infer that p . Each of these conditions can be fulfilled without the other being fulfilled.

The fact that it is possible for a thinker to meet the condition for judging one content without meeting the condition for judging a second does not imply that it is possible to *grasp* the first content without grasping the second. If 'grasping the content' means being capable of making judgements having that content, that will not always be possible when the senses are distinct. It will not be possible when they are made up from the same components, differently ordered. The possibility mentioned in the realist's account of distinctness of sense is a possibility concerning the conditions for actually judging a given content, rather than a possibility concerning the capacity for judgement.

The third role for a verificationist of the notion of the most direct method is one played in the philosophical theory of the justification of deduction. The notion plays a part both in the statement of the problem as it arises for a verificationist, and in the solution offered. The second role of the most direct method obviously leaves room for a solution to the problem: the task is to say how this room is to be filled.

As it arises in Dummett's account, the problem is how there can be a legitimate but indirect means of establishing a sentence; that is, a means which is sound but which is not included in a specification of the sense-determining, most direct, means of establishing the sentence. The solution Dummett offers is that a valid deduction is one in which any means of establishing its premisses can be transformed into a means of establishing its conclusion. This account of validity is one on which it is legitimate to establish a content as true by deduction from true premisses; it is also one on which the content thus established need not be one individuated by reference to that means of establishing it.²⁰

We saw far back that the realist cannot accept the verificatio-

²⁰ The notion of a most direct method plays a role in the solution not only as it occurs in the preceding sentence. It would also need to be employed in a more careful statement of the account of validity than that given in the preceding paragraph. A more careful statement would be that a valid inference is one in which any *direct* means of establishing the premisses can be transformed into a *direct* means of establishing the conclusion. Unless this stricter condition is met we will not have covered all cases in which the problem arises. If any case remains in which something is apparently legitimately deductively established by a means which is not the most direct, and the stricter condition is not met, the proposed solution will not have accounted for all cases of the phenomenon.

nist's substantive account of validity. At this point, the realist's path divides. One of the two paths is taken by a realist who attempts to defend his views by appeal to what may be verified in some ideal circumstances, or by a being with certain ideal powers. This is the direction Dummett assumes the realist will take when he writes:

On any molecular theory of meaning, the individual content of a sentence is determined by its internal structure, and relates, in the first place, to whatever constitutes the most direct means of recognising it as true; on a realist theory, this direct means of recognition of truth will often be inaccessible to us. (1978, p. 314)

If the realist takes this path, then he can take over Dummett's account of the justification of deduction and of validity; it will just be that for this realist, 'establishable' has some ideal reading. For this realist, a deduction from premisses which are true but not verifiable by us can still be valid. Its validity consists in the transformability of any means, not necessarily available to us, of establishing the premisses, into a means, not necessarily available to us, of establishing the conclusion. This realist will have no use for the alternatives we are offering him.

Any realist taking this path will face a host of familiar obstacles to acceptance of theories of content which appeal to verifiability in principle.²¹ It is not clear that such theories can explain what it is to have the conceptions employed in contents which this realist says are verifiable in principle. Nor is it clear that these theories cover all the cases. More generally, when we can conceive of a being who can verify contents which we cannot, there is a question about the order of explanation of this fact. It may be that, in some cases, once we are capable of judging thoughts which may be unverifiably true, we can then, drawing on our grasp of such contents, form the conception of beings who could determine the truth of those contents. That is, in some cases we *can* form the latter conception; but only as the conception of a being who could determine the truth of those contents which we grasp independently of conceiving of such beings. If that is so, the account of grasp of those contents cannot, on pain of circularity, make reference to the abilities of such ideal beings.

But the realist need not be entangled with these problems, since he could have taken the other path of the fork. The realist

²¹ There is a summary of the obstacles in Peacocke (1987a).

on this other path accepts the First Conjecture that an adequate account of what makes it the case that a thinker is judging a particular content determines its classical, possibly verification-transcendent truth condition. The classical model-theoretic definition of logical consequence can be seen as a technical elaboration of the intuitive idea that, in a valid inference, the conclusion is true when the premisses are, independently of the particular assignments to the nonlogical constants. The fundamental principle on which the realist who takes this second path relies in using model theory as a justification for an inference is this:

In every case in which a sentence is true, the sense of each of its components (together with the world) determines a semantic value of the sort assigned to it (or to expressions of its category) by the model theory; and the semantic value of the whole is determined from the semantic values of its components in accord with the compositional principles of the model theory.

If this assumption were false, the possibility would be left open that a sentence could be true in a kind of case not captured in the model theory. If the assumption were false, a sentence which the model theory counts as a logical consequence of a premiss might nevertheless not be true when the premiss is: for the model theory might not have captured every kind of case in which the premiss is true. The task of vindicating the assumption is one for a substantive theory of sense and reference applied to a particular language. This is another point at which a theory of the justification of logic must make contact with the core notions used in a substantive theory of concept possession.

On this second path, then, the realist gives a philosophical account of validity which motivates the approach of classical model theory. With this he has addressed the question of the legitimacy of deduction. Such a realist has already shown, in the alternative he can give to the second role of direct methods in the verificationist's theory, how a thinker can judge a content without being disposed to judge all its logical consequences; and so he also speaks to the question of the utility of deduction. These are, of course, all answers at the most general level of how it is possible that deduction should be both useful and legitimate.²² We are not here aiming at non-circular justifications of deduc-

²² That is, they are answers at the third of the three levels Dummett distinguishes (1978, pp. 308–9).

tion; obviously model theory, like almost any other interesting theory, must make extensive use of deduction itself.

On this second realist's account, then, the answer to some of the philosophical problems raised by deduction does not turn on a distinction between two kinds of method for establishing a content. It turns on the distinction between characterizations of content which are concerned with what else must be true if a given form of content is true, and characterizations sensitive to what makes that content, rather than any other, the one the thinker is judging.

In this material, I have been able to consider only a sampling of the areas with which a philosophical theory of the logical constants must deal. But I do hope to have suggested one general moral: that philosophical accounts of the semantics and the epistemology of the logical constants cannot be developed in isolation from one another. In aiming at an integrated account of these two aspects, I have been trying to provide for the logical constants something which we have eventually to provide for every type of concept: an account on which its referential and psychological properties are fully integrated.

10. *Appendix: a comparison with Harman's theory*

In his paper 'The Meaning of Logical Constants' Harman (1986a) offers a theory of the meaning of the logical constants, and criticizes truth-conditional specifications of the meaning of logical constants. I am indebted to Harman for a component of the positive account above, in particular for the point that in one way or another, logical constants may be introduced over incompatibility as well as over deducibility relations. Harman's own theory is that the meanings of logical constants are to be given by the relations of immediate implication and/or immediate incompatibility in which propositions containing them stand. Immediate implications and incompatibilities are ones which can be immediately recognized (1986a, pp. 131-2). For each of the traditionally acknowledged logical constants, Harman gives a detailed specification of such relations. He also argues that the meaning of a logical constant cannot be regarded as given by its contribution to truth conditions (*ibid.*, p. 130); nor, he says, need the statements of immediate implication and incompatibility for a constant be parallel to or be readable off

from its truth-theoretic clause (p. 127). He concludes that truth conditions are no more relevant to the meanings of logical constants than they are to the meanings of non-logical predicates (p. 134).

We can compare the account of this paper with Harman's on the following topics:

(a) the adequacy conditions for a theory of the logical constants;

(b) the detailed specifications offered for particular logical constants; and

(c) the question of whether meaning can be given by contribution to truth conditions. I take these in turn.

(a) We ought to aim to give an account of validity itself, as opposed to mere impressions of validity; and we ought to explain the relation of validity to impressions of validity. The relation may be quite close, as it is on the account of this paper, but it must still be a relation between distinct things. An account which mentions only impressions of validity will not be fulfilling this aim. The aim is one we should adopt if we want to legitimize the talk of recognizing validities; or if we want to give an account of the justification of logical principles; or if we want to develop a theory of meaning which says what is wrong with *tonk*. This adequacy condition is very abstract, and could be met by many different realistic or anti-realistic theories; indeed it could be met by theories which do not use any notion of truth at all. Harman does not assert that there is no such adequacy condition; but in practice he develops his theory without trying to meet it.

(b) At the level of detail, we can consider Harman's account of negation. He defines it as that one-place connective \mathcal{N} such that

$\mathcal{N}(p)$ is immediately inconsistent with p and is immediately implied by any set of propositions immediately inconsistent with p ; furthermore, any set of propositions immediately inconsistent with $\mathcal{N}(p)$ immediately imply p . (1986a, p. 132)

We have argued that someone who understands classical negation may still have to reflect to appreciate that $\sim \sim p$ implies p . The reflection might take the form of the intuitive argument that since in general $\sim p$ is true in any case in which p is not true, any case in which $\sim p$ fails to hold, i.e. any case in which $\sim \sim p$ is true, must be one in which p holds too. The reflection might also take other forms. Harman's definition, though, entails that double-negation elimination is immediately obvious. If we sub-

stitute ' $\sim p$ ' for ' p ', we get from the first conjunct of the first clause of Harman's definition that $\sim \sim p$ is immediately inconsistent with $\sim p$. No one should object to that; but then from this and the clause 'any set of propositions immediately inconsistent with $N(p)$ immediately imply p ', we get that (the unit set of) $\sim \sim p$ immediately implies p .

This argument would be blocked if the occurrences of 'immediately inconsistent' and of 'immediately imply' in Harman's final clause were replaced by the simple 'inconsistent' and 'imply' respectively. Indeed in *Change in View* (1986b) Harman gives as his rationale for this treatment of negation that 'What distinguishes $N(p)$ from other contraries of p is that $N(p)$ is the most inclusive contrary of p ; it is implied by any other contrary of p , that is, by anything else excluded by p '. (This is the point we earlier formulated by saying that the negation of a proposition is the weakest proposition contrary to it.) This rationale *does* require one to strip off the two occurrences of 'immediately' in the second clause, for the rationale says nothing about the immediacy of the logical relations it mentions. But moving to the stripped-down clause raises a question. Harman moved to immediate implication because it is a psychological notion which slices finely enough to capture meanings. Implication is not a psychological notion. Suppose we include it in a clause defining negation. Can we still say that the clause makes the issue of whether some conceptual constituent employed by a thinker is negation a matter of its psychological relations in his thought to other sentences or contents? Perhaps we could if implication could be defined as the transitive closure of immediate implication. But to offer that definition brings us back to issue (a) again: for there may be thinkers for whom would-be sentences containing *tonk* stand in relations of immediate implication—but they imply nothing, for they mean nothing. It seems that implication should rather be explained in terms of the notions used in characterizing validity itself, rather than in terms of impressions thereof.

(c) Lastly there is the global issue of whether we should accept Harman's attitude to the relation between such specifications for logical constants of relations of immediate implication and exclusion—however the details go—and the contribution the logical constants make to the truth conditions of sentences containing them.

I have been arguing that conceptual role semantics and truth-theoretic semantics need not be in competition, if 'conceptual

role semantics' is taken widely enough to include normative conditions on judging contents. They will be consistent if normative conceptual role—that is, a specification of conceptual role concerned with the norms governing judgement of a content—actually determines truth conditions. Harman notes that there is a parallelism between the truth-theoretic clause for conjunction, and the introduction and elimination rules for that constant. The introduction rule, he says, 'is just another way of saying' that the truth of both of its constituents is sufficient for the truth of a conjunction; while the elimination rule is just another way of saying that is necessary for the truth of a conjunction (1986a, p. 127). But, he continues, this parallelism between truth-theoretic clauses and natural deduction rules can be extended to other components (\sim , \vee) only by a series of *ad hoc* devices (ibid., p. 130).

However, a determination of truth conditions by normative conceptual role does not require that there be some uniform algorithm which for each constant allows one to read off its truth-theoretic clause from the specification of its conceptual role. On the contrary, on the account I offered the disquotational clauses for negation and for existential quantification in a truth theory will not be as closely related syntactically to the inference rules for those constants as the clause for conjunction is to its introduction and elimination rules. The form of the argument for the determination of truth conditions by normative acceptance conditions also varied as between cases: I used limiting conditions in some cases but not others. All this variety is in order; it does not undermine the general claim of determination of truth conditions by normative conditions relating to acceptance of the contents. The variety stems from the variety of the contents.

Harman also brings a more specific objection to the idea that the meaning of logical constants is to be given by their contribution to truth conditions. He introduces binary connectives C_1 and C_2 operating on p and q to have the same meanings respectively as $p \& q$ and $\sim(\sim p \vee \sim q)$ (1986a, pp. 126, 130). He says that C_1 and C_2 have different meanings, but make the same contributions to truth conditions. His reason is that the results of applying them to given propositions are logically equivalent (ibid.). I agree that C_1 and C_2 have different meanings, but do not agree that they make the same contribution to truth conditions. Throughout, I am using 'truth condition' as answerable to Fregean requirements, that is as individuated by cognitive

significance and as correlative to Frege's notion of a Thought. By an argument parallel to that we gave for * and § back in section 2 of this paper, differences in immediate implication will contribute to differences of Fregean truth condition. The difference would also be captured in a truth theory. The proper truth-theoretic clauses for C_1 and for C_2 would be

$$\text{True}(\text{'pC}_1\text{q'})} \text{ iff } (\text{True}(\text{p}) \ \& \ \text{True}(\text{q}))$$

and

$$\text{True}(\text{'pC}_2\text{q'})} \text{ iff } \sim ((\sim \text{True}(\text{p})) \vee (\sim \text{True}(\text{q}))).$$

The right-hand sides of these two biconditionals do not say the same; and so do not attribute the same contributions to truth conditions to C_1 and to C_2 .

Harman remarks that there is something trivial about the disquotational truth-theoretic clauses

the predicate 'horse' is true of something iff it is a horse

and

a sentence of the form 'p and q' is true iff p is true and q is true.

His reason for saying this is that one can know that these clauses are true simply by knowing that 'horse' is a one-place predicate and that 'and' is a binary truth-functional connective; 'you do not even have to know what these expressions mean' (Harman, 1986a, p. 125). But as Dummett emphasized (1975, pp. 106–7) there is nothing trivial about knowing the propositions which these displayed clauses express, as opposed to knowing that they express truths. One is not in a position to know these propositions simply by knowing that 'horse' is a predicate and that 'and' is truth-functional. It is the propositions which the displayed clauses express which are, on my account, determined by the normative conceptual role of 'and'. It is not merely the fact that the truth-theoretic clause expresses a truth. If the First Conjecture (see section 6) is correct, this places logical notions on a par with non-logical concepts.

References

- Belnap, N. (1961) 'Tonk, Plonk and Plink', *Analysis*, xxii, 130–4
 Boolos, G and Jeffrey, R. (1974) *Computability and Logic* (Cambridge: Cambridge University Press)

- Burge, T. (1979) 'Individualism and the Mental', *Midwest Studies in Philosophy*, iv
- (1986) 'Intellectual Norms and the Foundations of Mind', *Journal of Philosophy*, lxxxiii, 697–720
- Dennett, D. (1969) *Content and Consciousness* (London: Routledge)
- Dummett, M. (1975) 'What is a Theory of Meaning? (I)' in S. Guttenplan (ed.), *Language and Mind* (Oxford: Clarendon Press)
- (1978) 'The Justification of Deduction' repr. in his *Truth and Other Enigmas* (London: Duckworth)
- Evans, G. (1982) *The Varieties of Reference* (Oxford: Clarendon Press)
- Hacking, I. (1979) 'What is Logic?', *Journal of Philosophy*, lxxvi, 285–319
- Harman, G. (1986a) 'The Meaning of Logical Constants' in E. LePore (ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson* (Oxford: Blackwell)
- (1986b) *Change in View* (Cambridge, Mass.: MIT Press)
- Kleene, S. (1971) *Introduction to Metamathematics* (6th reprint, Groningen: Walters-Noordhoff)
- Nagel, T. (1979) 'What is it Like to be a Bat?' repr. in his *Mortal Questions* (Cambridge: Cambridge University Press)
- Peacocke, C. (1976) 'What is a Logical Constant?', *Journal of Philosophy*, lxxiii, 221–40
- (1986) *Thoughts: An Essay on Content* (Oxford: Blackwell)
- (1987a) 'The Limits of Intelligibility: A Post-Verificationist Proposal', forthcoming in *The Philosophical Review* (Scheduled for October 1988)
- (1987b) 'Perceptual Content', forthcoming in J. Almog, J. Perry and H. Wettstein (eds.), *Themes From Kaplan* (Oxford: Clarendon Press)
- Prawitz, D. (1965) *Natural Deduction* (Stockholm: Almqvist and Wicksell)
- (1978) 'Proofs and the Meaning and Completeness of the Logical Constants' in J. Hintikka et al. (eds.), *Essays on Mathematical and Philosophical Logic* (Dordrecht: Reidel)
- Prior, A. (1960) 'The Runabout Inference-Ticket', *Analysis*, xxi, 38–9
- Putnam, H. (1975) 'The Logic of Quantum Mechanics' repr. in his *Philosophical Papers*, Volume I: *Mathematics, Matter and Method* (Cambridge: Cambridge University Press)
- Quine, W. V. (1960) *Word and Object* (Cambridge, Mass.: MIT Press)
- (1970) *Philosophy of Logic* (Englewood Cliffs, New Jersey: Prentice-Hall)
- (1974) *The Roots of Reference* (LaSalle, Illinois: Open Court)
- (1976) 'Carnap and Logical Truth', repr. in *The Ways of Paradox* (2nd edn., Cambridge, Mass.: Harvard University Press)
- Stevenson, J. (1960) 'Roundabout the Runabout Inference-Ticket', *Analysis*, xxi, 124–8
- Wilde, O. (1966) *Complete Works* (London: Collins)
- Wittgenstein, L. (1958) *Philosophical Investigations* (2nd edn., Oxford: Blackwell) tr. G. E. M. Anscombe
- (1978) *Remarks on the Foundations of Mathematics* (3rd edn., Oxford: Blackwell) tr. G. E. M. Anscombe
- Wright, C. (1980) *Wittgenstein on the Foundations of Mathematics* (London: Duckworth)