

*Analyticity Reconsidered*<sup>1</sup>

PAUL ARTIN BOGHOSSIAN  
New York University

## I

This is what many philosophers believe today about the analytic/synthetic distinction: In his classic early writings on analyticity—in particular, in “Truth by Convention,” “Two Dogmas of Empiricism,” and “Carnap and Logical Truth”—Quine showed that there can be no distinction between sentences that are true purely by virtue of their meaning and those that are not. In so doing, Quine devastated the philosophical programs that depend upon a notion of analyticity—specifically, the linguistic theory of necessary truth, and the analytic theory of a priori knowledge.

Quine himself, so the story continues, went on to espouse far more radical views about meaning, including such theses as meaning-indeterminacy and meaning-skepticism. However, it is not necessary, and certainly not appealing, to follow him on this trajectory. As realists about meaning, we may treat Quine’s self-contained discussion in the early papers as the basis for a profound *insight* into the nature of meaning facts, rather than any sort of rejection of them. We may discard the notions of the analytic and the a priori without thereby buying in on any sort of unpalatable skepticism about meaning.

Now, I don’t know precisely how many philosophers believe all of the above, but I think it would be fair to say that it is the prevailing view. Philosophers with radically differing commitments—including radically differing commitments about the nature of meaning itself—subscribe to it: whatever precisely the correct construal of meaning, so they seem to think, Quine has shown that it will not sustain a distinction between the analytic and the synthetic. Listen, for example, to Bill Lycan:

It has been nearly forty years since the publication of “Two Dogmas of Empiricism.” Despite some vigorous rebuttals during that period, Quine’s rejection

© 1996 Blackwell Publishers Inc., 238 Main Street, Cambridge, MA 02142, USA, and 108 Cowley Road, Oxford OX4 1JF, UK.

of analyticity still prevails—in that philosophers en masse have either joined Quine in repudiating the “analytic/synthetic” distinction or remained (however mutinously) silent and made no claims of analyticity.

This comprehensive capitulation is somewhat surprising, in light of the radical nature of Quine’s views on linguistic meaning generally. In particular, I doubt that many philosophers accept his doctrine of the indeterminacy of translation. . . .

Lycan goes on to promise that in his paper, he is going to

make a Quinean case against analyticity, without relying on the indeterminacy doctrine. For I join the majority in denying both analyticity and indeterminacy. . . .<sup>2</sup>

Now, my disagreement with the prevailing view is not total. There is a notion of ‘truth by virtue of meaning’—what I shall call the metaphysical notion—that is undermined by a set of indeterminacy-independent considerations. Since this notion is presupposed by the linguistic theory of necessity, that project fails and must be abandoned.

However, I disagree with the prevailing view’s assumption that those very same considerations also undermine the analytic explanation of the a priori. For it seems to me that an entirely distinct notion of analyticity underlies that explanation, a notion that is epistemic in character. And in contrast with the metaphysical notion, the epistemic notion can be defended, I think, provided that even a minimal realism about meaning is true. I’m inclined to hold, therefore, that there can be no effective Quinean critique of the a priori that does not ultimately depend on Quine’s radical thesis of the indeterminacy of meaning, a thesis that, as I’ve stressed, many philosophers continue to reject.

All of this is what I propose to argue in this paper. I should emphasize right at the outset, however, that I am not a historian and my interest here is not historical. Think of me rather as asking, on behalf of all those who continue to reject Quine’s later skepticism about meaning: Can something like the analytic explanation of the a priori be salvaged from the wreckage of the linguistic theory of necessity?

### *Belief, Apriority and Indeterminacy*

We need to begin with some understanding—however brief and informal—of what it is to believe something and of what it is for a belief to count as a priori knowledge.

Let’s work with a picture of belief that is as hospitable as possible to Quine’s basic outlook. According to this ‘linguistic’ picture, the objects of belief are not propositions, but rather interpreted sentences: for a person **T**

to believe that *p* is for *T* to hold true a sentence *S* which means that *p* in *T*'s idiolect.<sup>3</sup>

Against this rough and ready background, we may say that for *T* to know that *p* is for *T* to justifiably hold *S* true, with a strength sufficient for knowledge, and for *S* to be true. And to say that *T* knows *p* a priori is to say that *T*'s warrant for holding *S* true is independent of outer, sensory experience.<sup>4</sup> The interesting question in the analysis of the concept of apriority concerns this notion of warrant: what is it for a belief to be justified independently of outer sensory experience?

On a minimalist reading, to say that the warrant for a given belief is a priori is just to say that it is justified, with a strength sufficient for knowledge, without appeal to empirical evidence.<sup>5</sup> On a stronger reading, it is to say that *and* that the justification in question is not defeasible by any future empirical evidence.<sup>6</sup> Which of these two notions is at issue in the present debate?

My own view is that the minimal notion forms the core of the idea of apriority and, hence, that it would be achievement enough to demonstrate its possibility. However, in this paper I will aim to provide the materials with which to substantiate the claim that, under the appropriate circumstances, the notion of analyticity can help explain how we might have a priori knowledge even in the strong sense. A defense of the strong notion is particularly relevant in the present context, for Quine seems to have been most skeptical of the idea of empirical indefeasibility.

Before proceeding, we should also touch briefly on the notion of meaning-indeterminacy. In Chapter Two of *Word and Object*, Quine argued that, for any language, it is possible to find two incompatible translation manuals that nevertheless perfectly conform to the totality of the evidence that constrains translation. This is the famous doctrine of the indeterminacy of translation. Since Quine was furthermore prepared to assume that there could not be facts about meaning that are not captured in the constraints on best translation, he concluded that meaning facts themselves are indeterminate — that there is, strictly speaking, no determinate fact of the matter as to what a given expression in a language means. This is the doctrine that I have called the thesis of the indeterminacy of meaning.

An *acceptance* of meaning-indeterminacy can lead to a variety of *other* views about meaning. For instance, it might lead to an outright eliminativism about meaning. Or it might be taken as a reason to base the theory of meaning on the notion of likeness of meaning, rather than on that of sameness of meaning.<sup>7</sup> In this paper, I am not concerned with the question what moral should be drawn from the indeterminacy thesis, on the assumption that it is true; nor am I concerned with whether the indeterminacy thesis is true. I am only concerned to show that a skepticism about

epistemic analyticity cannot stop short of the indeterminacy thesis, a thesis that, as I have stressed, most philosophers agree in rejecting.

*Analyticity: Metaphysical or Epistemological?*

Traditionally, three classes of statement have been thought to be the objects of a priori knowledge: logical statements, mathematical statements and such ‘conceptual truths’ as, for example, that all squares are four-sided. The problem has always been to explain what could justify us in holding such statements true on a priori grounds.

The history of philosophy has known a number of answers to this problem, among which the following has had considerable influence: We are equipped with a special evidence-gathering faculty of *intuition*, distinct from the standard five senses; by exercising this faculty, we are able to know a priori such truths as those of mathematics and logic.

The central impetus behind the *analytic* explanation of the a priori is a desire to explain the possibility of a priori knowledge without having to postulate such a special faculty, one that has never been described in satisfactory terms. The question is: How could a factual statement **S** be known a priori by **T**, without the help of a special evidence-gathering faculty?

Here, it would seem, is one way: *If mere grasp of S’s meaning by T sufficed for T’s being justified in holding S true.* If **S** were analytic in this sense, then, clearly, its apriority would be explainable without appeal to a special faculty of intuition: mere grasp of its meaning by **T** would suffice for explaining **T**’s justification for holding **S** true. On this understanding, then, ‘analyticity’ is an overtly *epistemological* notion: a statement is ‘true by virtue of its meaning’ provided that grasp of its meaning alone suffices for justified belief in its truth.

Another, far more *metaphysical* reading of the phrase ‘true by virtue of meaning’ is also available, however, according to which a statement is analytic provided that, in some appropriate sense, it *owes its truth value completely to its meaning*, and not at all to ‘the facts.’

Which of these two possible notions has been at stake in the dispute over analyticity? There has been a serious unclarity on the matter. Quine himself tends to label the doctrine of analyticity an epistemological doctrine, as for example in the following passage from “Carnap and Logical Truth”:

the linguistic doctrine of logical truth, which is an epistemological doctrine, goes on to say that logical truths are true purely by virtue of the intended meanings, or intended usage, of the logical words.<sup>8</sup>

However, his most biting criticisms seem often to be directed at what I have called the metaphysical notion. Consider, for example, the object of disap-

proval in the following famous passage, a passage that concludes the official discussion of analyticity in “Two Dogmas”:

It is obvious that truth in general depends on both language and extralinguistic fact. The statement ‘Brutus killed Caesar’ would be false if the world had been different in certain ways, but it would also be false if the word ‘killed’ happened rather to have the sense of ‘begat’. Thus one is tempted to suppose in general that the truth of a statement is somehow analyzable into a linguistic component and a factual component. Given this supposition it next seems reasonable that in some statements the factual component should be null; and these are the analytic statements. But for all its a priori reasonableness, a boundary between analytic and synthetic statements simply has not been drawn. That there is such a distinction to be drawn at all is an unempirical dogma of empiricists, a metaphysical article of faith.<sup>9</sup>

Now, I think that there is no doubt that many of the proponents of the analytic theory of the a priori, among them especially its positivist proponents, intended the notion of analyticity to be understood in this metaphysical sense; very shortly I shall look at why.

Before doing that, however, I want to register my wholehearted agreement with Quine that the metaphysical notion is of dubious explanatory value and possibly also of dubious coherence. Fortunately for the analytic theory of the a priori, it can be shown that it need have nothing to do with this discredited idea.

### *The Metaphysical Concept*

What could it possibly mean to say that the truth of a statement is fixed exclusively by its meaning and not by the facts? Isn’t it in general true — indeed, isn’t it in general a truism — that for any statement **S**,

**S** is true iff for some **p**, **S** means that **p** and **p**?

How could the *mere* fact that **S** means that **p** make it the case that **S** is true? Doesn’t it also have to be the case that **p**? As Harman has usefully put it (he is discussing the sentence ‘Copper is copper’):

what is to prevent us from saying that the truth expressed by “Copper is copper” depends in part on a general feature of the way the world is, namely that everything is self-identical.<sup>10</sup>

The proponent of the metaphysical notion does have a comeback, one that has perhaps not been sufficiently addressed. If he is wise, he won’t want to deny the meaning-truth truism. What he will want to say instead

is that, in some appropriate sense, our meaning  $p$  by  $S$  makes it the case that  $p$ .

But this line is itself fraught with difficulty. For how can we make sense of the idea that something is made true by our meaning something by a sentence?

Consider a sentence of the form 'Either  $p$  or not  $p$ '. It is easy, of course, to understand how the fact that we mean what we do by the ingredient terms fixes what is expressed by the sentence as a whole; and it is easy to understand, in consequence, how the fact that we mean what we do by the sentence determines whether the sentence expresses something true or false. But as Quine points out, that is just the normal dependence of truth on meaning. What is far more mysterious is the claim that the *truth of what the sentence expresses* depends on the fact that it is expressed by that sentence, so that we can say that what is expressed wouldn't have been true at all had it not been for the fact that it is expressed by that sentence. Are we really to suppose that, prior to our stipulating a meaning for the sentence

Either snow is white or it isn't

*it wasn't the case that either snow was white or it wasn't?* Isn't it overwhelmingly obvious that this claim was true *before* such an act of meaning, and that it would have been true even if no one had thought about it, or chosen it to be expressed by one of our sentences?

Why, if this idea is as problematic as I have claimed it to be, did it figure so prominently in positivist thinking about analyticity?

Much of the answer derives from the fact that the positivists didn't merely want to provide a theory of a priori knowledge; they also wanted to provide a reductive theory of necessity. The motivation was not purely epistemological, but metaphysical as well. Guided by the fear that objective, language-independent necessary connections would be both metaphysically and epistemologically odd, they attempted to show that all necessities could be understood to consist in linguistic necessities, in the shadows cast by conventional decisions concerning the meanings of words. Conventional linguistic meaning, by itself, was supposed to generate necessary truth; a fortiori, conventional linguistic meaning, by itself, was supposed to generate truth. Hence the play with the metaphysical concept of analyticity.

But this is, I believe, a futile project. In general, I have no idea what would constitute a better answer to the question: What is responsible for generating the truth of a given class of statements? than something bland like 'the world' or 'the facts'; and, for reasons that I have just been outlining, I cannot see how a good answer might be framed in terms of meaning, or convention, in particular.

So I have no sympathy with the linguistic theory of necessity or with its attendant Conventionalism. Unfortunately, the impression appears to be widespread that there is no way to disentangle that view from the analytic theory of the a priori; or, at a minimum, that there is no way to embrace the epistemic concept of analyticity without also embracing its metaphysical counterpart. I don't know whether Gil Harman believes something of the sort; he certainly gives the impression of doing so in his frequent suggestions that anyone deploying the notion of analyticity would have to be deploying both of its available readings simultaneously:

It turned out that someone could be taught to make the analytic-synthetic distinction only by being taught a rather substantial theory, a theory including such principles as that meaning can make something true and that knowledge of meaning can give knowledge of truth.<sup>11</sup>

One of the main points of the present paper is that these two notions of analyticity are distinct and that the analytic theory of the a priori needs only the epistemological notion and has no use whatsoever for the metaphysical one. We can have an analytic theory of the a priori without in any way subscribing to a Conventionalism about anything. It is with the extended defense of this claim that much of the present essay is concerned.

### *The Epistemological Concept*

Turning, then, to the epistemic notion of analyticity, we immediately confront a serious puzzle: How could any sentence be analytic in this sense? How could mere grasp of a sentence's meaning justify someone in holding it true?

Clearly, the answer to this question has to be *semantical*: something about the sentence's meaning, or about the way that meaning is fixed, must explain how its truth is knowable in this special way. What could this explanation be?

In the history of the subject, two different sorts of explanation have been especially important. Although these, too, have often been conflated, it is crucial to distinguish between them.

One idea was first formulated in full generality by Gottlob Frege. According to Frege, a statement's analyticity (in my epistemological sense) is to be explained by the fact that it is *transformable into a logical truth by the substitution of synonyms for synonyms*. When a statement satisfies this semantical condition, I shall say that it is 'Frege-analytic'.<sup>12</sup>

Now, it should be obvious that Frege-analyticity is at best an *incomplete* explanation of a statement's epistemic analyticity and, hence, of its apriority. For suppose that a given sentence S is Frege-analytic. How might this fact explain its analyticity? Clearly, two further assumptions are

needed. First, that facts about synonymy are knowable a priori; and second, that the truths of logic are. Under the terms of these further assumptions, a satisfying explanation goes through. Given its Frege-analyticity, *S* is transformable into a logical truth by the substitution of synonyms for synonyms. Facts about synonymy are a priori, so it's a priori that *S* is so transformable. Furthermore, the sentence into which it is transformable is one whose truth is itself knowable a priori. Hence, *S*'s truth is knowable a priori.

Frege tended not to worry about these further assumptions, for two reasons. First, Frege thought it obviously constitutive of the idea of meaning, that meaning is transparent—that any competent user of two words would have to be able to know a priori whether or not they meant the same. Second, Frege also thought it obvious that there could be no substantive epistemology for logic—a fortiori, not one that could explain its apriority. As a consequence, he was happy to take logic's apriority for granted. For both of these reasons, he didn't worry about the fact that an explanation of apriority in terms of Frege-analyticity simply leaned on these further assumptions without explaining them.

I think the jury is still out on whether Frege was right to take these further assumptions for granted. There is certainly a very strong case to be made for the transparency of meaning.<sup>13</sup> And there are well-known difficulties providing a substantive epistemology for something as basic as logic, difficulties we shall have occasion to further review below. Nevertheless, because we cannot simply assume that Frege was right, we have to ask how a complete theory of the a priori would go about filling in the gaps left by the concept of Frege-analyticity.

I shall have very little to say about the first gap. The question whether facts about the sameness and difference of meaning are a priori cannot be discussed independently of the question what meaning is, and that is not an issue that I want to prejudge in the present context. On some views of meaning—for example, on certain conceptual role views—the apriority of synonymy is simply a by-product of the very nature of meaning facts, so that no substantive epistemology for synonymy is necessary or, indeed, possible. On other views—for example, on most externalist views of meaning—synonymy is not a priori, so there is no question of a sentence's Frege-analyticity fully explaining its epistemic analyticity.

Since this issue about the apriority of synonymy turns on questions that are currently unresolved, I propose to leave it for now. As we shall see, none of the analyticity-skeptical considerations we shall consider exploit it in any way. (Quine never argues that the trouble with Frege-analyticity is that synonymies are a posteriori.)

Putting aside, then, skepticism about the apriority of synonymy, and, for the moment anyway, skepticism about the very existence of Frege-analytic

sentences, let us ask quite generally: What class of a priori statement would an account based on the notion of Frege-analyticity *fail* to explain?

Two classes come to mind. On the one hand, a priori statements that are not transformable into logical truths by the substitution of synonyms for synonyms; and, on the other hand, a priori statements that are trivially so transformable.

Taking the first class first, there do appear to be a significant number of a priori statements that are not Frege-analytic. For example:

Whatever is red all over is not blue.

Whatever is colored is extended.

If x is warmer than y, then y is not warmer than x.

These statements appear not to be transformable into logical truths by the appropriate substitutions: the ingredient descriptive terms seem not to be decomposable in the appropriate way.

The second class of recalcitrant statements consists precisely of the truths of logic. The truths of logic satisfy, of course, the conditions on Frege-analyticity. But they satisfy them trivially. And it seems obvious that we can't hope to explain our warrant for belief in the truths of logic by appealing to their analyticity in this sense: knowledge of Frege-analyticity presupposes knowledge of logical truth and so can't explain it.

How, then, is the epistemic analyticity of these recalcitrant truths to be explained? As we shall see below, the solution proposed by Carnap and the middle Wittgenstein turns on the suggestion that they are to be viewed as *implicit definitions* of their ingredient terms. When a statement satisfies this semantical condition, I shall sometimes say that it is 'Carnap-analytic'. However, before proceeding to a discussion of Carnap-analyticity, I want to re-examine Quine's famous rejection of the much weaker concept of Frege-analyticity.<sup>14</sup>

## II

### *"Two Dogmas" and the Rejection of Frege-analyticity*

For all its apparent limitations, the concept of Frege-analyticity is not without interest. Even though Quine made it fashionable to claim otherwise, the sentence "All bachelors are male," *does* seem to be transformable into a logical truth by the substitution of synonyms for synonyms and that fact *does* seem to have something important to do with its apriority. If, then, appearances are not misleading here, and a significant range of a priori statements are Frege-analytic, then the problem of their apriority is *reduced* to that of the apriority of logic and synonymy and, in this way, a significant economy in explanatory burden is achieved.

It was, therefore, an important threat to the analytic theory of the a priori to find Quine arguing, in one of the most celebrated articles of this century, that the apriority of no sentence could be explained by appeal to its Frege-analyticity, because no sentence of a natural language could *be* Frege-analytic.

It has not been sufficiently appreciated, it seems to me, that “Two Dogmas,” is exclusively concerned with this weaker notion of Frege-analyticity, and not at all with the more demanding project of explaining the apriority of logic. But this is made very clear by Quine:

Statements which are analytic by general philosophical acclaim are not, indeed, far to seek. They fall into two classes. Those of the first class, which may be called *logically true*, are typified by:

(1) No unmarried man is married.

The relevant feature of this example is that it is not merely true as it stands, but remains true under any and all reinterpretations of ‘man’ and ‘married’. If we suppose a prior inventory of *logical* particles . . . then in general a logical truth is a statement that remains true under all reinterpretations of its components other than the logical particles.

But there is also a second class of analytic statements, typified by:

(2) No bachelor is married.

The characteristic of such a statement is that it can be turned into a logical truth by putting synonyms for synonyms. (pp. 22–23)

Quine goes on to say very clearly:

Our problem . . . is analyticity; and here the major difficulty lies not in the first class of analytic statements, the logical truths, but rather in the second class, which depends on the notion of synonymy. (p. 24)

Most of the rest of TD is devoted to arguing that no good sense can be made of such analyticities of the ‘second class’.

None of this would make any sense unless Quine were intending in “Two Dogmas” to be restricting himself solely to the notion of Frege-analyticity. Of course, it is the point of two other important papers of his—“Truth by Convention” and “Carnap and Logical Truth”—to argue that there is no non-trivial sense in which *logic* is analytic. We will turn to that issue in due course. Relative to the Fregean notion, however, the logical truths are trivially analytic; and so, given his apparent desire to restrict his attention to that notion in TD, he simply concedes their ‘analyticity’ in the only sense he takes to be under discussion. What he wishes to resist in TD, he insists, is merely the claim that there are any *non-trivial instances of Frege-analyticity*.<sup>15</sup>

*Skeptical Theses about Analyticity*

What form does Quine's resistance take? We may agree that the result being advertised isn't anything modest, of the form: There are fewer analyticities than we had previously thought. Or, there are some analytic truths, but they are not important for the purposes of science. Or anything else of a similar ilk. Rather, as a very large number of Quine's remarks make clear, the sought-after result is something ambitious to the effect that the notion of Frege-analyticity is, somehow or other, not cogent. TD's many admirers have divided on whether to read this as the claim that the notion of Frege-analyticity does not have a well-defined, determinate factual content, or whether to read it merely as claiming that, although it has an intelligible content, it is necessarily uninstantiated. I'll call the first claim a *Non-factualism* about analyticity:

(NF) No coherent, determinate property is expressed by the predicate 'is analytic' (or, since these are correlative terms, the predicate 'is synthetic'); consequently, no coherent factual claim is expressed by sentences of the form 'S is analytic' and 'S is synthetic.'

And the second an *Error Thesis* about analyticity:

(ET) There is a coherent, determinate property expressed by 'is analytic', but it is necessarily uninstantiated; consequently, all sentences of the form 'S is analytic' are necessarily false.<sup>16</sup>

Regardless, however, of how TD's skepticism about Frege-analyticity is understood, I don't see how either thesis can plausibly stop short of a radical indeterminacy about meaning.

*Non-Factualism about Frege-Analyticity*

Let's begin with the non-factualist version. To say that there is no such property as the property of Frege-analyticity is essentially to say that, for *any* sentence, there is no fact of the matter as to whether it is transformable into a logical truth by the substitution of synonyms for synonyms. Presumably, this itself is possible only if, either there is no fact of the matter as to what counts as a logical truth, or no fact of the matter as to when two expressions are synonymous. Since the factuality of logic is not in dispute, the only option is a non-factualism about synonymy.

But, now, how can there fail to be facts about whether any two expressions mean the same—even where these are drawn from within a *single* speaker's idiolect, so that no questions of *interlinguistic* synonymy arise? Wouldn't this have to entail that there are no facts about what each expression means individually? Putting the question the other way: Could there

be a fact of the matter about what each expression means, but no fact of the matter about whether they mean the same?<sup>17</sup>

Let's consider this question first against the background of an unQuinean relational construal of meaning, according to which an expression's meaning something is a relation **M** between it and its meaning, the meaning **C**. Someone who held that a non-factualism about synonymy could co-exist with a determinacy about meaning would have to hold that, although it might be true that some specific word — say, "cow" — bears some specific relation **M** to some specific meaning **C**, there is no fact of the matter about whether some *other* word — some other orthographically identified particular — bears precisely the same relation to precisely the same meaning.

But how could this be? How could it conceivably turn out that it is intelligible and true to say that "cow" bears **M** to **C**, and not merely false but *nonfactual* to say that some other word — "vache" as it may be — also does? What could be so special about the letters "c", "o", "w"?

The answer, of course, is that there is nothing special about them. If it is factual that one word bears **M** to **C**, it is surely factual that some other word does. Especially on a relational construal of meaning, it makes no sense to suppose that a determinacy about meaning could coexist with a non-factualism about synonymy.

The question naturally arises whether this result is forthcoming *only* against the background of a relational construal of meaning. I think it's quite clear that the answer is 'No'. To see why, suppose that instead of construing meaning facts as involving relations to meanings we construe them thus: "cow" means *cow* just in case "cow" has the monadic property **R** — a history of use, a disposition, or whatever your favorite candidate may be. Precisely the same arguments go through: it remains equally difficult to see how, given that "cow" has property **R**, it could fail to be factual whether some other word does.

### *The Error Thesis about Frege-Analyticity*

I think, then, that if a plausible skepticism about Frege-analyticity is to be sustained, it cannot take the form of a non-factualism. Does an Error thesis fare any better? According to this view, although there are determinate facts about which sentences are transformable into logical truths by the appropriate manipulations of synonymy, this property is necessarily uninstantiated: it is nomically impossible for there to be any Frege-analytic sentences. Our question is: Does at least this form of skepticism about Frege-analyticity avoid collapse into the indeterminacy doctrine?

Well, I suppose that if we are being very strict about it, we may have to admit that it is barely *logically possible* to combine a denial of indeterminacy with an error thesis about synonymy, so that we can say that although there are determinate facts about what means what, it is impossible for any

two things to mean the same thing. But is such a view plausible? Do we have any reason for believing it? I think not.

Let's begin with the fact that even Quine has to admit that it is possible for two *tokens of the same orthographic type* to be synonymous, for that much is presupposed by his own account of logical truth.<sup>18</sup>

What about two tokens of different types? Here again, our own argument can proceed from Quine's own admissions. For even Quine concedes that two expressions can mean the same thing, provided that they are explicitly stipulated to mean the same thing.<sup>19</sup> So his skepticism about synonymy has to boil down to the following somewhat peculiar claim: Although there is such a thing as the property of synonymy; and although it can be instantiated by pairs of tokens of the same orthographic type; and although it can be instantiated by pairs of tokens of distinct orthographic types, provided that they are related to each other by way of an explicit stipulation; it is, nevertheless, in principle impossible to generate instances of this property in some other way, via some other mechanism. For example, it is impossible that two expressions that were introduced independently of each other into the language, should have been introduced with exactly the same meanings.

But what conceivable rationale could there be for such a claim? As far as I am able to tell, there is precisely one argument in the literature that is supposed to provide support for it. It may be represented as follows:

Premise: Meaning is radically holistic in the sense that: "What our words mean depends on *everything* we believe, on *all* the assumptions we are making."<sup>20</sup>

Therefore,

Conclusion: It is very unlikely that, in any given language, there will be two words of distinct types that mean exactly the same thing.

I am inclined to agree that this argument (properly spelled out) is valid, and so, that if a radical holism about meaning were true, then synonymies between expressions of different types would be rare.

However, I note that "rare" does not mean the same as "impossible," which is the result we were promised. And, much more importantly, I am completely inclined to disagree that TD provides any sort of cogent argument for meaning holism in the first place.

It's easy to see why, if such a radical meaning holism were true, synonymies might be hard to come by. For although it is not unimaginable, it is unlikely that two words of distinct types will participate in *all* of the same

beliefs and inferences. Presumably there will always be some beliefs that will discriminate between them — beliefs about their respective shapes, for example.

But what reason do we have for believing that *all* of a word's uses are constitutive of its meaning?

Many Quineans seem to hold that the crucial argument for this intuitively implausible view is to be found in the concluding sections of TD. In those concluding sections, Quine argues powerfully for the epistemological claim that has come to be known as the Quine-Duhem thesis: confirmation is holistic in that the warrant for any given sentence depends on the warrant for every other sentence. In those concluding sections, Quine also assumes a Verificationist theory of meaning, according to which the meaning of a sentence is fixed by its method of confirmation. Putting these two theses together, one can speedily arrive at the view that a word's meaning depends on *all* of its inferential links to other words, and hence at the thesis of meaning-holism.<sup>21</sup>

This, however, is not a very convincing train of thought. First, and not all that importantly, this couldn't have been the argument that *Quine* intended against Frege-analyticity, for this argument for meaning holism is to be found in the very last pages of TD, well after the rejection of Frege-analyticity is taken to have been established.

Second, and more importantly, the argument is not very compelling because it depends crucially on a verificationism about meaning, a view that we have every good reason to reject, and which has in fact been rejected by most contemporary philosophers.

Finally, and perhaps most importantly, any such holism-based argument against the possibility of synonymy would need to be supported by something that no one has ever provided — a reason for believing that yielding such an intuitively implausible result about synonymy isn't itself simply a *reductio* of meaning holism.<sup>22</sup>

### III

#### *The Analyticity of Logic*

If the preceding considerations are correct, then there is no principled objection to the existence of Frege-analyticities, and, hence, no principled objection to the existence of statements that are knowable a priori if logical truth is.<sup>23</sup>

But what about logical truth? Is it knowable a priori? And, if so, how?<sup>24</sup>

In the case of some logical truths, the explanation for how we have come to know them will be clear: we will have deduced them from others. So our question concerns only the most elementary laws of sentential or first-order

logic. How do we know a priori, for example, that all the instances of the law of non-contradiction are true, or that all the instances of modus ponens are valid?

As I noted above, Frege thought it obvious that there could be no substantive answer to such questions; he was inclined, therefore, to take appearances at face value and to simply *assume* the apriority of logic.

What Frege probably had in mind is the following worry. ‘Explaining our knowledge of logic’ presumably involves finding some *other* thing that we know, on the basis of which our knowledge of logic is to be explained. However, regardless of what that other thing is taken to be, it’s hard to see how the use of logic is to be avoided in moving from knowledge of that thing to knowledge of the relevant logical truth. And so it can come to seem as if any account of how we know logic will have to end up being vacuous, presupposing that we have the very capacity that’s to be explained.

Michael Dummett has disputed the existence of a real problem here. As he has pointed out, the sort of circularity that’s at issue isn’t the gross circularity of an argument that consists of including the conclusion that’s to be reached among the premisses. Rather, we have an argument that purports to prove the validity of a given logical law, at least one of whose inferential steps must be taken in accordance with that law. Dummett calls this a “pragmatic” circularity. He goes on to claim that a pragmatic circularity of this sort will be damaging only to a justificatory argument that

is addressed to someone who genuinely doubts whether the law is valid, and is intended to persuade him that it is. . . . If, on the other hand, it is intended to satisfy the philosopher’s perplexity about our entitlement to reason in accordance with such a law, it may well do so.<sup>25</sup>

The question whether Dummett’s distinction fully allays Frege’s worry is a large one, and I can’t possibly hope to settle it here. If something along these general lines can’t be made to work, then *any* explanation of logic’s apriority – or aposteriority, for that matter – is bound to be futile, and the Fregean attitude will have been vindicated.

However, the question that particularly interests me in the present essay is this: Assuming that the very enterprise of explaining our knowledge of logic isn’t shown to be hopeless by Frege’s straightforward argument, is there any *special* reason for doubting an explanation based on the notion of analyticity? Quine’s enormously influential claim was that there is. I shall try to argue that there isn’t – that, in an important sense to be specified later on, our grasp of the meaning of logical claims can explain our a priori warrant for holding them true (provided that the Fregean worry doesn’t defeat all such explanations in the first place).

*The Classical View and Implicit Definition*

It's important to understand, it seems to me, that the analytic theory of the apriority of logic arose indirectly, as a by-product of the attempt to explain in what a grasp of the meaning of the logical constants consists. Alberto Coffa lays this story out very nicely in his recent book.<sup>26</sup>

What account are we to give of our grasp of the logical constants, given that they are not to explicitly definable in terms of *other* concepts? Had they been explicitly definable, of course, we would have been able to say — however plausibly — that we grasp them by grasping their definitions. But as practically anybody who has thought about the matter has recognized, the logical constants are not explicitly definable in terms of other concepts, and so we are barred from giving that account. The question is, what account are we to give?

Historically, many philosophers were content to suggest that the state of grasping these constants was somehow primitive, not subject to further explanation. In particular, such a grasp of the meaning of, say, 'not', was to be thought of as prior to, and independent of, a decision on our part as to which of the various sentences involving 'not' to count as true. We may call this view, following Wittgenstein's lead, the doctrine of

*Flash-Grasping*: We grasp the meaning of, say, 'not' "in a flash" — prior to, and independently of, deciding which of the sentences involving 'not' are true.

On this historically influential picture, Flash-Grasping was combined with the doctrine of Intuition to generate an epistemology for logic:

*Intuition*: This grasp of the concept of, say, negation, along with our intuition of its logical properties, explains and justifies our logical beliefs involving negation — e.g., that 'If not not  $p$ , then  $p$ ' is true.

As Coffa shows, this picture began to come under severe strain with the development of alternative geometries. Naturally enough, an analogous set of views had been used to explain the apriority of geometry. In particular, a flash-grasp of the indefinables of geometry, along with intuitions concerning their necessary properties, was said to explain and justify belief in the axioms of Euclidean geometry.

However, with the development of alternative geometries, such a view faced an unpleasant dilemma. Occupying one horn was the option of saying that Euclidean and non-Euclidean geometries are talking about the *same* geometrical properties, but disagreeing about what is true of them. But this option threatens the thesis of Intuition: If in fact we learn geometrical truths by intuition, how could this faculty have misled us for so long?

Occupying the other horn was the option of saying that Euclidean and non-Euclidean geometries are talking about *different* geometrical properties—attaching different meanings to, say, ‘distance’—and so not disagreeing after all. But this option threatens the doctrine of Flash-Grasping. Suppose we grant that a Euclidean and a non-Euclidean geometer attach different meanings to ‘distance’. In what does the difference in the respective psychological states consist? Officially, of course, the view is that one primitive state constitutes grasp of Euclidean distance, and another that of non-Euclidean distance. But absent some further detail about how to tell such states apart and the criteria that govern their attribution, this would appear to be a hopelessly ad hoc and non-explanatory maneuver.

The important upshot of these considerations was to make plausible the idea that grasp of the indefinables of geometry consists precisely in the adoption of one set of truths involving them, as opposed to another. Applied to the case of logic, it generates the semantical thesis that I shall call

*Implicit definition:* It is by arbitrarily stipulating that certain sentences of logic are to be true, or that certain inferences are to be valid, that we attach a meaning to the logical constants. More specifically, a particular constant means that logical object, if any, which makes valid a specified set of sentences and/or inferences involving it.

Now, the transition from this sort of implicit definition account of grasp, to the analytic theory of the apriority of logic, can seem pretty immediate. For it would seem that the following sort of argument is now in place:

1. If logical constant C is to mean what it does, then argument-form A has to be valid, for C means whatever logical object in fact makes A valid.
2. C means what it does.

Therefore,

3. A is valid.

I will return to various questions regarding this form of justification below.<sup>27</sup> For now I want to worry about the fact that neither Carnap nor Wittgenstein was content merely to replace Flash Grasping with Implicit Definition. Typically, both writers went on to embrace some form of anti-realism about logic. Intuitively, the statements of logic appear to be fully factual statements, expressing objective truths about the world, even if necessary ones, and even if (on occasion) highly obvious ones. Both

Carnap and Wittgenstein, however, seemed inclined to deny such an intuitive realism about logic, affirming in its place either the thesis of logical Non-Factualism or the thesis of logical Conventionalism, or, on occasion, both theses at once.

By *logical Non-Factualism*<sup>28</sup>, I mean the view that the sentences of logic that implicitly define the logical primitives do not express factual claims and, hence, are not capable of genuine truth or falsity. How, on such a view, are we to think of their semantic function? On the most popular version, we are to think of it as prescriptive, as a way of expressing a rule concerning the correct use of logical expressions. By contrast, logical Conventionalism is the view that, although the sentences of logic are factual—although they can express truths—their truth values are not objective, but are rather determined by our conventions.

Despite this important difference between them, there is an interesting sense in which the upshot of both views is the same, a fact that probably explains why they were often used interchangeably and why they often turn up simultaneously in the analytic theory of logic. For what both views imply is that, as between two different sets of decisions regarding which sentences of logic to hold true, there can be no epistemic fact of the matter. In short, both views imply an epistemic relativism about logic. Conventionalism implies this because it says that the truth in logic is up to us, so no substantive disagreement is possible; and Non-Factualism implies this because it says that there are no truths in logic, hence nothing to disagree about.

Nevertheless, for all this affinity of upshot, it should be quite plain that the two views are very different from each other—indeed, incompatible with each other. Conventionalism is a factualist view: it presupposes that the sentences of logic have truth values. It differs from a realist view of logic in its conception of the *source* of those truth values, not on their existence. Therefore, although it is possible, as I have noted, to find texts in which a rule-prescriptivism about logic is combined with Conventionalism, that can only be a confusion.

The important question is: Why did the proponents of Implicit Definition feel the need to go beyond it all the way to the far more radical doctrines of Non-Factualism and/or Conventionalism? Whatever problems it may eventually be discovered to harbor, Implicit Definition seems like a plausible candidate for explaining our grasp of the logical constants, especially in view of the difficulties encountered by its classical rival. But there would appear to be little that *prima facie* recommends either Non-factualism or Conventionalism. So why combine these dubious doctrines with what looks to be a plausible theory of meaning?

Apparently, both Carnap and Wittgenstein seem to have thought that the issue was forced, that Implicit Definition entailed one or the other anti-

realist thesis. It seems quite clear that Carnap, for example, believed that Implicit Definition brought Conventionalism immediately in its wake; and Quine seems to have agreed. What separated them was their attitude towards Conventionalism. Carnap embraced it; Quine, by contrast, seems to have been prepared to reject any premise that led to it; hence his assault on the doctrine of Implicit Definition.

But if this is in fact the correct account of Quine's motivations, then they are based, I believe, on a false assumption, for neither form of anti-realism about logic follows from the thesis of Implicit Definition.

I will proceed as follows. First, I will argue that Implicit Definition, properly understood, is completely independent of any form of anti-realism about logic. Second, I will defend the thesis of Implicit Definition against Quine's criticisms. Finally, I will examine the sort of account of the apriority of logic that this doctrine is able to provide.

### *Implicit Definition and Non-Factualism*

Does Implicit Definition entail Non-Factualism? It is certainly very common to come across the claim that it does. Coffa, for instance, writes that from the new perspective afforded by the doctrine of Implicit Definition, the basic claims of logic are

our access to certain meanings, definitions in disguise, devices that allow us to implement an explicit or tacit decision to constitute certain concepts. . . . From this standpoint, necessary claims do not tell us anything that is the case both in the world and in many others, as Leibniz thought, or anything that is the case for *formal* reasons, whatever that might mean, or anything that one is forced to believe due to features of our mind. They do not tell us anything that is the case; so they had better not be called claims or propositions. Since their role is to constitute meanings and since (apparently) we are free to endorse them or not, it is better to abandon the old terminology (a priori "principles", "laws," etc.) that misleadingly suggests a propositional status and to refer to them as "rules." (pp. 265–266)

I have no desire to engage the exegetical issues here; as far as I can tell, the middle Wittgenstein seems very much to have been a non-factualist about the implicit definers of logic, just as Coffa says. What I dispute is that it *follows* from the fact that a given sentence **Q** is being used to implicitly define one of its ingredient terms, that **Q** is not a factual sentence, not a sentence that "tells us anything that is the case." These two claims seem to me to be entirely independent of each other.

To help us think about this, consider Kripke's example of the introduction of the term 'meter'. As Kripke imagines it, someone introduces the term into his vocabulary by stipulating that the following sentence is to be true:

[1] Stick **S** is a meter long at **t**.

Suppose that stick **S** exists and is a certain length at **t**. Then it follows that ‘meter’ names that length and hence that [1] says that stick **S** is that length at **t**, and since it is that length at **t**, [1] is true.

Knowing all this may not be much of an epistemic achievement, but that isn’t the point. The point is that there appears to be no inconsistency whatsoever between claiming that a given sentence serves to implicitly define an ingredient term and claiming that that very sentence expresses something factual.

Similarly, I don’t see that there is any inconsistency between supposing that a given logical principle—for instance, the law of excluded middle—serves to implicitly define an ingredient logical constant, and supposing that that very sentence expresses a factual statement capable of genuine truth and falsity.<sup>29</sup>

### *Implicit Definition and Conventionalism*

So far I have argued that it is consistent with a sentence’s serving as an implicit definer that that very sentence come to express a fully factual claim, capable of genuine truth and falsity. Perhaps, however, when implicit definition is at issue, the truth of the claim that is thereby fixed has to be thought of as conventionally determined? Does at least Conventionalism follow from Implicit Definition?<sup>30</sup>

It is easy to see, I suppose, why these two ideas might have been run together. For according to Implicit Definition, ‘if, then’, for example, comes to mean the conditional precisely by my assigning the truth value True to certain basic sentences involving it, for example, to

If, if **p** then **q**, and **p**, then **q**.

And in an important sense, my assigning this sentence the value True is arbitrary. Prior to my assigning it that truth value, it didn’t have a complete meaning, for one of its ingredient terms didn’t have a meaning at all. The process of assigning it the value True is simply part of what fixes its meaning. Had I assigned it the value False, the sentence would then have had a *different* meaning. So, prior to the assignment there couldn’t have been a substantive question regarding its truth value. And after the assignment there couldn’t be a substantive question as to whether that assignment was correct. In this sense, then, the sentence’s truth value is arbitrary and conventional. Doesn’t it follow, then, that Implicit Definition entails Conventionalism?

Not at all. All that is involved in the thesis of Implicit Definition is the claim that the conventional assignment of truth to a sentence determines

what claim that sentence expresses (if any); such a view is entirely silent about what (if anything) determines the truth of the claim that is thereby expressed – a fortiori, it is silent about whether our conventions determine it.

Think here again of Kripke's meter stick. If the stick exists and has such-and-so length at *t*, then it is conventional that 'meter' names that length and, therefore, conventional that [1] expresses the proposition stick *S* has such-and-so length at *t*. However, that stick *S* has that length at *t* is hardly a fact generated by convention; it presumably had that length prior to the convention, and may continue to have it well after the convention has lapsed.<sup>31</sup>

I anticipate the complaint that the entailment between Implicit Definition and Conventionalism is blocked only through the tacit use of a distinction between a sentence and the proposition it expresses, a distinction that neither Carnap nor Quine would have approved.

Such a complaint would be mistaken, however. The argument I gave relies not so much on a distinction between a sentence and a proposition in the technical sense disapproved of by Quine, as on a distinction between a sentence and *what it expresses*. And it is hard to see how any adequate philosophy of language is to get by without some such distinction.<sup>32</sup> Even on a deflationary view of truth, there is presumably a distinction between the *sentence* 'Snow is white' and that which makes the sentence true, namely, snow's being white. And the essential point for my purposes is that it is one thing to say that 'Snow is white' comes to express the claim that snow is white as a result of being conventionally assigned the truth value True; and quite another to say that snow comes to be white as a result of our conventions. The first claim is Implicit Definition (however implausibly applied in this case); and the other is Conventionalism. Neither one seems to me to entail the other.

#### *Quine against Implicit Definition: Regress*

As I noted above, I am inclined to believe that erroneous opinion on this score has played an enormous role in the history of this subject. I conjecture that had Quine felt more confident that Implicit Definition could be sharply distinguished from Conventionalism, he might not have felt so strongly against it.

In any event, though, whatever the correct explanation of Quine's animus, we are indebted to him for a series of powerful critiques of the thesis of Implicit Definition, critiques that have persuaded many that that thesis, and with it any explanation of the apriority of logic that it might be able to ground, are fundamentally flawed. We must now confront Quine's arguments.

According to Implicit Definition, the logical constants come to have a

particular meaning in our vocabulary by our conventionally stipulating that certain sentences (or inferences) involving them are to be true. For instance, let us assume that the meaning for ‘and’ is fixed by our stipulating that the following inferences involving it are to be valid:

$$[2] \quad \frac{A \text{ and } B}{A} \quad \frac{A \text{ and } B}{B} \quad \frac{A, B}{A \text{ and } B}$$

Now, Quine’s first important criticism of this idea occurs in his early paper ‘Truth by Convention’.<sup>33</sup> As Quine there pointed out, there are an infinite number of instances of schema [2]. Consequently, the inferences of this infinitary collection could not have been conventionally stipulated to be valid singly, one by one. Rather, Quine argued, if there is anything at all to this idea, it must be something along the following lines: We adopt certain general conventions from which it follows that all the sentences of the infinitary collection are assigned the value Valid. Such a general convention would presumably look like this:

Let all results of putting a statement for ‘p’ and a statement for ‘q’ in ‘p and q implies p’ be valid.

However, the trouble is that in order to state such a general convention we have had, unavoidably, to use all sorts of logical terms — ‘every’, ‘and’, and so on. So the claim, essential to the proposal under consideration, that all our logical constants acquire their meaning via the adoption of such explicitly formulated conventional assignments of validity must fail. Logical constants whose meaning is not fixed in this way are presupposed by the model itself.<sup>34</sup>

This argument of Quine’s has been very influential; and I think that there is no doubt that it works against its target as specified. However, it is arguable that its target as specified isn’t the view that needs defeating.

For, surely, it isn’t compulsory to think of someone’s following a rule **R** with respect to an expression **e** as consisting in his *explicitly stating* that rule in so many words in the way that Quine’s argument presupposes. On the contrary, it seems far more plausible to construe x’s following rule **R** with respect to **e** as consisting in some sort of fact about x’s *behavior* with **e**.

In what would such a fact consist? Here there are at least a couple of options. According to a currently popular idea, following rule **R** with respect to **e** may consist in our being disposed to conform to rule **R** in our employment of **e**, under certain circumstances. On this version, the notion of rule-following would have been *reduced* to a certain sort of dispositional fact. Alternatively, one might wish to appeal to the notion of following a given

rule, while resisting the claim that it can be reduced to a set of naturalistically acceptable dispositional facts. On such a non-reductionist version, there would be facts about what rule one is following, even if these are not cashable into facts about one's behavioral dispositions, however optimal.

For myself, I am inclined to think that the reductionist version won't work, that we will have to employ the notion of following a rule unreduced.<sup>35</sup> But because it is more familiar, and because nothing substantive hangs on it in the present context, I will work with the reductionist version of rule-following. Applied to the case we are considering, it issues in what is widely known in the literature as a "conceptual role semantics."

According to this view, then, the logical constants mean what they do by virtue of figuring in certain inferences and/or sentences involving them and not in others. If some expressions mean what they do by virtue of figuring in certain inferences and sentences, then some inferences and sentences are *constitutive* of an expression's meaning what it does, and others aren't. And any CRS must find a systematic way of saying which are which, of answering the question: What properties must an inference or sentence involving a constant *C* have, if that inference or sentence is to be constitutive of *C*'s meaning?

#### *Quine against Implicit Definition: Constitutive Truth*

Now, Quine's second objection to Implicit Definition can be put by saying that there will be no way of doing what I said any CRS must do — namely, systematically specify the meaning-constituting inferences. Quine formulated this point in a number of places. Here is a version that appears in 'Carnap and Logical Truth':

if we try to warp the linguistic doctrine of logical truth into something like an experimental thesis, perhaps a first approximation will run thus: *Deductively irresolvable disagreement as to a logical truth is evidence of deviation in usage (or meanings) of words.* . . . [However] the obviousness or potential obviousness of elementary logic can be seen to present an insuperable obstacle to our assigning any experimental meaning to the linguistic doctrine of elementary logical truth. . . . For, that theory now seems to imply nothing that is not already implied by the fact that elementary logic is obvious or can be resolved into obvious steps.<sup>36</sup>

Elsewhere, Quine explained his use of the word "obvious" in this connection thus:

In "Carnap and Logical Truth" I claimed that Carnap's arguments for the linguistic doctrine of logical truth boiled down to saying no more than that they were obvious, or potentially obvious — that is, generable from obvieties by obvious steps. I had been at pains to select the word 'obvious' from the vernacular, intend-

ing it as I did in the vernacular sense. A sentence is obvious if (a) it is true and (b) any speaker of the language is prepared, for any reason or none, to assent to it without hesitation, unless put off by being asked so obvious a question.<sup>37</sup>

Quine's important point here is that there will be no substantive way of distinguishing between a highly obvious, non-defining sentence and a sentence that is an implicit definer. Both types of sentence—if in fact both types exist—will have the feature that any speaker of the language will be prepared to assent to instances of them, “for any reason or none.” So in what does the alleged difference between them consist? How is distinctive content to be given to the doctrine of Implicit Definition?<sup>38</sup>

Now, there is no doubt that this is a very good question; and the impression that it has no good answer has contributed greatly to the rejection of the doctrine of Implicit Definition. Jerry Fodor and Ernie Lepore, for example, base the entirety of their recent argument against a conceptual role semantics on their assumption that Quine showed this question to be unanswerable.<sup>39</sup>

If Quine's challenge is allowed to remain unanswered, then the threat to the analytic theory of the a priori is fairly straightforward. For if there is no fact of the matter as to whether *S* is a sentence that I must hold true if *S* is to mean what it does, then there is no basis on which to argue that I am entitled to hold *S* true without evidence.

But that would seem to be the least of our troubles, if Quine's argument is allowed to stand. For what's threatened is not only the apriority of logical truths but, far more extremely, the *determinacy* of what they claim. For as I've already pointed out, and as many philosophers are anyway inclined to believe, a conceptual role semantics seems to be the *only* plausible view about how the meaning of the logical constants is fixed. It follows, therefore, that if there is no fact of the matter as to which of the various inferences involving a constant are meaning-constituting, then there is also no fact of the matter as to what the logical constants themselves mean. And that, again, is just the dreaded indeterminacy of meaning on which the critique of analyticity was supposed not to depend.

The simple point here is that if the only view available about how the logical constants acquire their meaning is in terms of the inferences and/or sentences that they participate in, then any indeterminacy in what those meaning-constituting sentences and inferences are will translate into an indeterminacy about the meanings of the expressions themselves. This realization should give pause to any philosopher who thinks he can buy in on Quine's critique of implicit definition without following him all the way to the far headier doctrine of meaning-indeterminacy.

There has been a curious tendency to miss this relatively simple point. Jerry Fodor seems a particularly puzzling case. For Fodor holds all three of

the following views. (1) He rejects indeterminacy, arguing forcefully against it. (2) He follows Quine in rejecting the notion of a meaning-constituting inference. (3) He holds a conceptual role view of the meanings of the logical constants. As far as I am able to judge, however, this combination of views is not consistent.<sup>40</sup>

Part of the explanation for this curious blindness derives from a tendency to view Quine's argument as issuing not in an indeterminacy about meaning, but, rather, in a *holism* about it. In fact, according to Fodor and Lepore, the master argument for meaning holism in the literature runs as follows:

A. Some of an expression's inferential liaisons are relevant to fixing its meaning.

B. There is no principled distinction between those inferential liaisons that are constitutive and those that aren't. (The Quinean result.)

Therefore,

C. All of an expression's inferential liaisons are relevant to fixing its meaning. (Meaning Holism)

Fearing this argument's validity, and seeing no way to answer Quine's challenge, they spend their whole book trying to undermine the argument's first premise, namely, the very plausible claim that at least *some* of an expression's inferential liaisons are relevant to fixing its meaning.<sup>41</sup>

But they needn't have bothered, for I don't see how the master argument could be valid in the first place. The claim that *all* of an expression's inferential liaisons are constitutive of it cannot cogently follow from the claim that it is *indeterminate* what the constitutive inferences are. If it's *indeterminate* what the constitutive inferences are, then it's genuinely *unsettled* what they are. And that is inconsistent with saying that they are *all* constitutive, and inconsistent with saying that *none* are constitutive and inconsistent with saying that some specified subset are constitutive.

Fodor and Lepore are not alone in not seeing the problem here. Let me cite just one more example. In his comments on an earlier version of the present paper, Gil Harman says:

Can one accept Quine's argument against analyticity without being committed to the indeterminacy of meaning? Yes and no. By the "indeterminacy of meaning" might be meant an indeterminacy as to which of the principles one accepts determine the meanings of one's terms and which simply reflect one's opinions

about the facts. Clearly, Quine's argument against analyticity is committed to that sort of indeterminacy. [However] that by itself does not imply full indeterminacy in the sense of Chapter Two of *Word and Object*.<sup>42</sup>

As Harman correctly says, Quine has to deny that there is a fact of the matter as to which of **T**'s principles determine the meanings of his terms and which simply reflect **T**'s opinions about the facts — that, after all, is just what it is to deny that there are facts about constitutivity. However, Harman insists, this denial in no way leads to the indeterminacy thesis of Chapter Two of *Word and Object*.

But this is very puzzling. Against the background of a conceptual role semantics, according to which the meaning of **T**'s term **C** is determined precisely by a certain subset of the principles involving **C** that **T** accepts, an indeterminacy in what the meaning-determining principles are will automatically lead to an indeterminacy in what the meaning is — in the full sense of Chapter Two of *Word and Object*. If a subset (not necessarily proper) of accepted principles is supposed to determine meaning; and if there is no fact of the matter as to which subset that is; then there is, to that extent, no fact of the matter as to what meaning has been determined.

I think there is really no avoiding the severe conclusion that meaning is indeterminate, if the Quinean challenge to constitutivity is allowed to remain unanswered. I'm inclined to think, therefore, that anyone who rejects radical indeterminacy of meaning must believe that a distinction between the meaning-constituting and the non-meaning-constituting can be drawn. The only question is how.

Well, that is not the task of the present paper. Although there are some good ideas about this, I don't have a fully thought-through proposal to present just now.<sup>43</sup> My main aim here is not to *solve* the fundamental problem for a conceptual role semantics for the logical constants; rather, as I have stressed, it is to show that, against the background of a rejection of indeterminacy, its insolubility cannot be conceded.

Pending the discovery of other problems, then, it seems open to us to suppose that a plausible theory of meaning for the logical constants is given by something like the following:

A logical constant **C** expresses that logical object, if any, that makes valid its meaning-constituting inferences.

#### *Implicit Definition, Justification and Entitlement*

Now, how does any of this help vindicate the analytic theory of the apriority of logic, the idea that logic is epistemically analytic? Let us consider a particular inference form, **A**, in a particular thinker's (**T**) repertoire; and

let's suppose that that inference form is constitutive of the meaning of one of its ingredient constants **C**. How, exactly, might these facts help explain the epistemic analyticity of **A** for **T**?

To say that **A** is epistemically analytic for **T** is to say that **T**'s knowledge of **A**'s meaning alone suffices for **T**'s justification for **A**, so that empirical support is not required. And it does seem that a conceptual role semantics can provide us with a model of how that might be so. For given the relevant facts, we would appear to be able to argue as follows:

1. If **C** is to mean what it does, then **A** has to be valid, for **C** means whatever logical object in fact makes **A** valid.
2. **C** means what it does.

Therefore,

3. **A** is valid.

Now, it is true that this is tantamount to a fairly broad use of the phrase "knowledge of the meaning of **A**," for this knowledge includes not merely knowledge of what **A** means, strictly so-called, but also knowledge of how that meaning is fixed. But this is, of course, both predictable and unavoidable: there was never any real prospect of explaining apriority merely on the basis of a knowledge of propositional content. Even Carnap realized that one needed to know that a given inference or sentence had the status of a 'meaning postulate'.

But isn't it required, if this account is to genuinely explain **T**'s a priori justification for the basic truths of logic, that **T** know the premisses a priori as well? Yet, it hasn't been shown that **T** can know the premisses a priori.

It is quite correct that I have not attempted to show that the relevant facts about meaning cited in the premisses are knowable a priori, although I believe that it is intuitively quite clear that they are. I have purposely avoided discussing all issues relating to knowledge of meaning facts. My brief here has been to defend epistemic analyticity; and this requires showing only that certain sentences are such that, *if* someone knows the relevant facts about their meaning, *then* that person will be in a position to form a justified belief about their truth. It does not require showing that the knowledge of those meaning facts is itself a priori (although, I repeat, it seems quite clear to me that it will be).<sup>44</sup>

Isn't it a problem for the aspirations of the present account that a thinker would have to use modus ponens to get from the premisses to the desired conclusion?

Not if Dummett's distinction between pragmatic and vicious circularity

is credited with opening a space for an epistemology for logic, as discussed above.

Finally, how could such an account possibly hope to explain the man in the street's justification for believing in the truths of logic? For such a person, not only would the relevant meaning facts be quite opaque, he probably wouldn't even be capable of framing them. Yet such a person is obviously quite justified in believing the elementary truths of logic. Thus, so our objector might continue, this sort of account cannot explain our ordinary warrant for believing in logic; at best, it can explain the warrant that sophisticates have.

I think that, strictly speaking, this objection is correct, but only in a sense that strips it of real bite. Philosophers are often in the position of articulating a warrant for an ordinary belief that the man in the street would not understand. If we insist that a person counts as justified only if they are aware of the reason that warrants their belief, then we will simply have to find another term for the kind of warrant that ordinary folk often have and that philosophers seek to articulate. Tyler Burge has called it an "entitlement":

The distinction between justification and entitlement is this. Although both have positive force in rationally supporting a propositional attitude or cognitive practice, and in constituting an epistemic right to it, entitlements are epistemic rights or warrants that need not be understood by or even be accessible to the subject. . . . The unsophisticated are entitled to rely on their perceptual beliefs. Philosophers may articulate these entitlements. But being entitled does not require being able to justify reliance on these resources, or even to conceive such a justification. Justifications, in the narrow sense, involve reasons that people have and have access to.<sup>45</sup>

When someone is entitled, all the facts relevant to the person's justification are already in place, so to say; what's missing is the reflection that would reveal them.

Just so in the case at hand. If a conceptual role semantics is true, and if **A** is indeed constitutive of **C**'s meaning what it does, then those facts by themselves constitute a warrant for **A**; empirical support is not necessary. **A** can only be false by meaning something other than what it means. But these facts need not be known by the ordinary person. They suffice for his entitlement, even if not for his full-blown justification. This full-blown justification can be had only by knowing the relevant facts about meaning.

### *Conclusion*

Quine helped us see the vacuity of the metaphysical concept of analyticity and, with it, the futility of the project it was supposed to underwrite — the

linguistic theory of necessity. But I don't see that those arguments affect the epistemic notion of analyticity that is needed for the purposes of the theory of a priori knowledge. Indeed, it seems to me that epistemic analyticity can be defended quite vigorously, especially against the background of a realism about meaning.

On the assumption that our warrant for believing in elementary logical truths cannot be explained, the outstanding problem is to explain our a priori knowledge of conceptual truths. For this purpose, the crucial semantical notion is that of Frege-analyticity. I have argued that this notion is bound to be in good standing for a meaning realist.

If the project of explaining logic is not ruled hopeless, then I have tried to show how the doctrine that appears to offer the most promising account of how we grasp the meanings of the logical constants—namely, Implicit Definition—can explain the epistemic analyticity of our logical beliefs and, hence, our a priori warrant for believing them. As long as we are not prepared to countenance radical indeterminacy, we should have every confidence that this form of explanation can be made to work.<sup>46</sup>

### Notes

<sup>1</sup>This is a shorter, and somewhat modified, version of a paper entitled "Analyticity," which is to appear in Crispin Wright and Bob Hale (eds.): *A Companion to the Philosophy of Language* (Cambridge: Blackwell, 1996). I am grateful to Blackwell, and to the editors, for permission to use some of that material here.

<sup>2</sup>"Definition in a Quinean World," in J. Fetzer, D. Shatz, and G. Schlesinger (eds.): *Definitions and Definability: Philosophical Perspectives* (Dordrecht: Kluwer, 1991), pp. 111–131.

<sup>3</sup>As I say, I am going to work with this linguistic picture out of deference to my opponents. I would prefer to work with a propositionalist picture of belief, according to which the objects of belief are propositions in the technical sense—mind- and language-independent, abstract objects which have their truth conditions essentially. Most of the crucial notions developed in this paper, and much of the argument involving them, can be translated, with suitable modifications, into this propositionalist framework.

<sup>4</sup>The inclusion of the word "outer" here is partly stipulative. I have always found it natural to regard a priori knowledge as encompassing both knowledge that is based on no experience as well as knowledge that is based purely on *inner* experience.

<sup>5</sup>In the interests of brevity, I shall henceforth take it as understood that "justification" means "justification with a strength sufficient for knowledge."

<sup>6</sup>Even this strong notion is not as demanding as many have supposed. For instance, it is consistent with a belief's being a priori in the strong sense that we should have *pragmatic* reasons for dropping it from our best overall theory. For illuminating discussion of the modesty of the notion of the a priori see Crispin Wright: "Inventing Logical Necessity," in Butterfield (ed.) *Language, Mind and Logic* (Cambridge: Cambridge University Press, 1984) and Bob Hale: *Abstract Objects* (Oxford: Blackwell, 1986), ch. 6.

<sup>7</sup>See Gilbert Harman, *Thought* (Princeton: Princeton University Press, 1973).

<sup>8</sup>*The Ways of Paradox* (Cambridge: Harvard University Press, 1976), p. 103.

<sup>9</sup>*From a Logical Point of View* (Cambridge: Harvard University Press, 1953), pp. 36–7.

<sup>10</sup>"Quine on Meaning and Existence I," *Review of Metaphysics* 21: 124–151, p. 128. I am grateful to Paul Horwich for emphasizing the importance of this point.

<sup>11</sup>"Doubts About Conceptual Analysis," MS, p. 5. See also his "Quine on Meaning and Existence I."

<sup>12</sup>See G. Frege (Austin, trans.): *The Foundations of Arithmetic*, sec 3, Oxford: Blackwell,

1950). (Some may regard the attribution of precisely this notion to Frege controversial. What matters to me is not who came up with the idea, but rather the philosophical role it has played.)

My use of the term ‘analytic’ in connection with Frege’s *semantical* notion as well as with the preceding epistemic and metaphysical concepts may be thought ill-advised. But I do so deliberately, to highlight the fact that the term has been used in the literature in general, and in Quine in particular, to stand for all three different sorts of notion, often without any acknowledgement of that fact. This terminological promiscuity has undoubtedly contributed to the confusion surrounding discussions of this issue.

<sup>13</sup>For some discussion see my “The Transparency of Mental Content,” in *Philosophical Perspectives*, v.8, 1994, pp.33–50.

<sup>14</sup>What follows is a compressed discussion of Frege-analyticity. For a fuller treatment see “Analyticity,” *op. cit.*

<sup>15</sup>Exegetically, this does leave us with a few puzzles. First, TD does contain a brief discussion of the implicit definition idea, under the guise of the notion of a “semantical rule.” Given that, why does Quine insist that he intends only to discuss the notion of Frege-analyticity? Second, the notion of a semantical rule is discussed only in connection with non-logical truths; since, however, the deployment of this idea would be exactly the same in the logical case, why is the analyticity of logic expressly excluded? Third, given that the analyticity of logic is expressly excluded, on what basis does Quine allow himself to draw morals about logic’s revisability towards the end of TD? I think there is no avoiding the conclusion that, on this and other related issues (see below), TD is confused. It would, in fact, have been surprising if these rather tricky problems had all been in clear focus in Quine’s pioneering papers.

<sup>16</sup>In this context, nothing fancy is meant by the use of such expressions as ‘property’ and ‘proposition’. For present purposes they may be understood in a thoroughly deflationary manner.

I have sometimes been asked why I consider just this particular weakening of a non-factualist thesis, one that involves, problematically from Quine’s official point of view, a modal notion? Why not rather attribute to him the following *Very Weak Thesis*:

(VWT) There is a coherent, determinate property expressed by ‘is analytic’, but *as a matter of fact*, it has never been instantiated; consequently, all tokens of the sentence ‘S is analytic’ have been false up to now.

There are two reasons. First, the VWT is not a philosophically interesting thesis and, second, it could not have been argued for on the basis of a *philosophy* paper—i.e., on the sorts of a priori grounds that Quine offers. So although Quine may not be entitled to precisely the ET, I am going to ignore that and not hold it against him.

<sup>17</sup>This question was first asked by Grice and Strawson in their “In Defense of a Dogma,” reprinted in Grice: *Studies in the Way of Words* (Cambridge: Harvard University Press, 1989). Grice and Strawson didn’t sufficiently stress, however, that Quine was committed to a skepticism even about *intra*linguistic synonymy, and not just about *inter*linguistic synonymy, for the theory of apriority doesn’t care about the interlinguistic case.

<sup>18</sup>See Peter Strawson, *Logico-Linguistic Papers*, (London: Methuen, 1971), p.117.

<sup>19</sup>See the discussion of stipulative definitions in TD. For further discussion see “Analyticity.”

<sup>20</sup>Harman: *Thought*, *op. cit.*, p. 14, emphasis in the original.

<sup>21</sup>Recent formulations of this argument may be found in Fodor: *Psychosemantics*, (Cambridge, MA: MIT Press, 1987), pp 62ff; Fodor and Lepore: *Holism: A Shopper’s Guide* (Oxford: Blackwell, 1991), pp. 37ff; Devitt: *Coming to Our Senses*, (New York: Cambridge University Press, 1995), p.17. None of the authors mentioned approve of the argument.

<sup>22</sup>A further TD-based argument for meaning holism, this time invalid, will be considered further below, in connection with the discussion of the thesis of Implicit Definition.

<sup>23</sup>As before, subject to the proviso about the apriority of synonymy.

<sup>24</sup>I am ignoring for now the class of a priori truths that are neither logical nor Frege-analytic. As we shall see, the very same strategy—implicit definition—that can be applied to explain our knowledge of logic can be applied to them as well.

<sup>25</sup>*The Logical Basis of Metaphysics* (Cambridge: Harvard University Press, 1991), p.202.

<sup>26</sup>Coffa, A.: *The Semantic Tradition*, (Cambridge: Cambridge University Press, 1991), ch.

14. In the next three paragraphs, I follow the general contours of the account that Coffa puts forward. However, the formulations are mine and they differ in important respects from Coffa's, as we shall see further on.

<sup>27</sup>Readers who are acquainted with a paper of mine of mine entitled "Inferential Role Semantics and the Analytic/Synthetic Distinction," *Philosophical Studies*, Spring 1994, pp. 109–122, will be aware that I used to worry that Implicit Definition could not generate a priori knowledge because of the falsity of something I called "The Principle." The Principle is the thesis that it follows from a sentence's being an implicit definer that that sentence is true. This is a tangled issue that I cannot fully discuss here. I will have to settle for a few brief remarks. I stand by the letter of what I said in the earlier paper. However, *part* of the problem there highlighted for the theory of the a priori is taken care of here by a reformulation of the thesis of Implicit Definition; another part is taken care of by a reformulation of the relation between Implicit Definition and the a priori; and, finally, a residual problem, not discussed in this paper, is met by the section entitled "A Pragmatic Solution" in "Analyticity," *op. cit.*. Readers for whom this footnote reads darkly may ignore it in its entirety.

<sup>28</sup>Not to be confused with the non-factualism about Frege-analyticity discussed earlier in the paper.

<sup>29</sup>Someone may object that the two cases are not relevantly analogous. For the meter case is supposed to be a case of the *fixation of reference*, but the logical case an instance of the fixation of meaning. Doesn't this difference between them block the argument I gave?

I don't see that it does. First, the two cases really are disanalogous only if there is an important difference between meaning and reference; yet, as is well-known, there are many philosophers of language who are inclined to think that there isn't an important such difference. Second, it seems to me that even if we allowed for a robust distinction between meaning and reference, the point would remain entirely unaffected. Whether we think of an implicit definer as fixing a term's reference directly, or we think of it as first fixing its meaning, which then in turn fixes its reference, seems to me entirely irrelevant to the claim that Implicit Definition does not entail Non-Factualism. As long as both processes are consistent with the fixation of a factual claim for the sentence at issue—as they very much seem to be—the important point stands.

<sup>30</sup>Certainly many philosophers seem to have thought so. Richard Creath, for example, sympathetically expounds Carnap's view that the basic axioms of logic implicitly define the ingredient logical terms by saying that on this view "the postulates (together with the other conventions) create the truths that they, the postulates express." See his "Carnap's Conventionalism," *Synthese* 93: 141–165, p. 147.

<sup>31</sup>This point is also forcefully made by Nathan Salmon in "Analyticity and Apriority," *Philosophical Perspectives*, 1994, and by Stephen Yablo in his review of Sidelle, *Philosophical Review* 1988.

<sup>32</sup>Notice that conventionalists themselves need to make crucial use of such a distinction when they describe their own position, as in the passage cited above from Creath:

the postulates (together with the other conventions) create the truths that they, the postulates, express.

As Hilary Putnam pointed out some time ago, it's hard to see how distinctive content is to be given to Conventionalism without the use of some such distinction. For a conventionalism merely about linguistic expressions is trivial. A real issue is joined only when the view is formulated as a claim about the truths expressed. See Putnam, "The Refutation of Conventionalism," in his *Mind, Language and Reality: Philosophical Papers v.2* (NY: Cambridge University Press, 1975)

<sup>33</sup>Quine's argument here is officially directed against a Conventionalism about logical truth, that is, against the idea that logical truth is determined by our conventions. This idea we have already rejected in our discussion of the metaphysical concept of analyticity. However, Quine attacks Conventionalism by attacking the semantical thesis of Implicit Definition. Hence, the need for the present discussion.

<sup>34</sup>Quine claims that this argument may also be put as follows: The claim that the sentences of logic lack assignment of truth value until they are conventionally assigned such values must

fail. For logic is needed in order to infer from a formulated general convention that the infinitely many instances of a given schema are true. Hence, sentences of logic whose truth value is not fixed as the model requires, are presupposed by the model itself.

It's unclear to me that this is a formulation of precisely the same argument. However, to the extent that it is distinct, it is also addressed by the proposal I put forth below.

<sup>35</sup>For discussion see my "The Rule-Following Considerations," *Mind*, 1989, pp.507–549. 36p. 105, *op. cit.*

<sup>37</sup>"Reply to Hellman," in Schilpp (ed.): "The Philosophy of WVO Quine," (La Salle: Open Court, 1975), p. 206.

<sup>38</sup>For all its influence, it is still possible to find the force of the Quinean point being underestimated by the friends of Implicit Definition. Christopher Peacocke, for example, in a recent, subtle defense of an inferential role semantics claims that what makes the inferences involving the logical constants constitutive is that a thinker finds those inferences "primitively compelling" and does so because they are of those forms. He goes on to explain:

To say that a thinker finds such instances primitively compelling is to say this: (1) he finds them compelling; (2) he does not find them compelling because he has inferred them from other premises and/or principles; and (3) for possession of the concept in question . . . he does not need to take the correctness of the transitions as answerable to anything else.

*A Study of Concepts* (Cambridge: MIT Press, 1992), p. 6. I think it is plain, however, that these conditions are insufficient for answering the Quinean challenge: a non-constitutive, though highly obvious, form of inference may also be found compelling because of its form, and not on the basis of inference from anything else. So these conditions cannot be what distinguish between a constitutive and a non-constitutive inference.

<sup>39</sup>"Why Meaning (Probably) Isn't Conceptual Role," *op. cit.*

<sup>40</sup>For Fodor's views on the mentioned issues, see his *Psychosemantics* (Cambridge: MIT Press, 1989) and *The Elm and the Expert* (Cambridge: MIT Press, 1994).

<sup>41</sup>See Fodor and Lepore: *Holism: A Shopper's Guide* (Oxford: Blackwell, 1993).

<sup>42</sup>Harman, "Comments on Boghossian," APA Symposium on Analytic Truth, Boston, MA, December 1994.

<sup>43</sup>For a good start, see Peacocke: *A Study of Concepts*, *op. cit.*

<sup>44</sup>For a discussion of why the second premiss is a priori see "Analyticity," *op. cit.*

<sup>45</sup>Burge, "Content Preservation," *The Philosophical Review*, October 1993.

<sup>46</sup>I am grateful to a number of audiences—at MIT, CUNY Graduate Center, Michigan State, the University of Chicago, the SOFIA Conference on Tenerife, the Chapel Hill Colloquium, Dartmouth College, London University and Oxford University. An earlier version of this paper was presented at the NEH Institute on the "Nature of Meaning," held at Rutgers University in the summer of 1992. It was there that I first became aware that Christopher Peacocke has been thinking along somewhat similar lines about the a priori—see his "How are A Priori Truths Possible?" presented at the Rutgers conference. Although there are a number of differences between our approaches, and although Peacocke's focus is not on the notion of analyticity, I have benefited from discussing these matters with him. I also benefited from presenting a version of this paper as part of a symposium on Analytic Truth, involving Gil Harman, Burton Dreben and WVO Quine, at the 1994 Eastern Division meetings of the APA. I am especially grateful to Gil Harman, Elizabeth Fricker, Hartry Field, Gary Gates, Bill Lycan, Stephen Schiffer and Barry Loewer for their detailed comments on previous versions of this paper. Special thanks are due to Bob Hale and Crispin Wright for their patience and for their very helpful reactions to several different drafts. For other helpful discussion and commentary, I want to thank Jennifer Church, Jerry Fodor, Albert Casullo, Norma Yunez, Neil Tennant, Peter Unger, Tom Nagel, Paul Horwich, Ned Block, Richard Creath, Allan Gibbard, Stephen Yablo and David Velleman.