

ANALYTICITY AND IMPLICIT DEFINITION

Kathrin GLÜER
Uppsala University

Summary

Paul Boghossian advocates a version of the analytic theory of a priori knowledge. His defense of an “epistemic” notion of analyticity is based on an implicit definition account of the meaning of the logical constants. Boghossian underestimates the power of the classical Quinean criticisms, however; the challenge to substantiate the distinction between empirical and non-empirical sentences, as forcefully presented in *Two Dogmas*, still stands, and the regress from *Truth by Convention* still needs to be avoided. Here, Quine also showed that there are no implicit definers for the logical constants. Moreover, even if they existed, their epistemic analyticity would, on Boghossian’s own account, be doubtful.

1. *Introduction*

The concepts of analyticity and the a priori are certainly experiencing a renaissance these days. One of the most interesting suggestions defending both analyticity and the a priori has been made by Paul Boghossian. In his papers on analyticity, he advocates a version of what he calls the “analytic theory of a priori knowledge”. This theory, dear to the logical positivists, can be separated, Boghossian claims, from another theory they also held, known as the “linguistic doctrine of necessity”. Crucial to this is a distinction between two different notions of analyticity: the distinction between *metaphysical analyticity* and *epistemic analyticity*. Quine, Boghossian argues, was right about metaphysical analyticity; it cannot be defended, and the same holds for the linguistic doctrine of necessity based upon it. Epistemic analyticity, however, can be defended, and so can an analytic theory of a priori knowledge taking it as its basis.

Boghossian’s version of the analytic theory of a priori knowledge

employs Carnap's idea of implicit definition. He holds that at least some linguistic expressions, the logical constants, get their meaning by implicit definition. Implicit definers, that is, sentences implicitly defining an expression, Boghossian argues, are analytic in the epistemic sense: Belief in their truth is justified by knowing their meaning. The knowledge thus gained, the argument continues, is a priori—nothing but knowledge of meaning is required for having such knowledge. However, according to Boghossian, and contrary to standard logical positivist doctrine, such knowledge is factual knowledge, knowledge about the world. Boghossian thus defends what might be called a *moderate rationalism*—with Carnap against Carnap. But above all, of course, against Quine.

And it is Quine whom I am going to defend. Now, to be a bit more precise, what Boghossian argues for is that Quine's criticisms of analyticity are ineffective as long as they do not *depend* on the thesis of indeterminacy (cf. 1996, 361). In what follows, however, I shall mainly focus on two of the Quinean arguments from the classical anti-analytic campaign, arguments to be found in *Truth by Convention* and *Two Dogmas*. They not only ante-date the indeterminacy thesis, but clearly would not seem to depend on it, either. For most of what follows, I can therefore safely ignore this complication; there will be some occasion to come back to it, though. What I am going to do is roughly this: After setting out Boghossian's account in more detail (sections 2 and 3), I shall investigate how immune this semantic explanation of a priori knowledge really is with respect to the classical Quinean criticisms. I shall concentrate on two of them: the argument from epistemic holism from *Two Dogmas* (section 4) and the regress charge from *Truth by Convention* (section 5). In the last part of my paper, however, I shall argue that an implicit definition account, even if it could be defended against Quine, would still not provide an explanation of a priori knowledge.

2. *Two concepts of analyticity*

It has become standard to point out that Quine, in *Two Dogmas*, employed more than one concept of analyticity and to complain that he did not sufficiently distinguish between these. Tyler Burge, for instance, distinguishes between the notion of truth by virtue of meaning alone, on the one hand, and the idea of reducibility, by substitution of synonyms

for synonyms, to the truths of logic, on the other (cf. Burge 1992, 7 ff.). Boghossian calls the latter kind of analyticity “Frege analyticity”. Truth by virtue of meaning alone, in turn, can be read in two ways, namely as truth independently of matters of fact and as truth independently of experience. Now, all of these notions might in fact coincide according to standard logical positivist doctrine. From a more neutral point of view, however, their interrelations are far from clear. On Quine’s own understanding of the truths of logic, as suggested in the later sections of *Two Dogmas*, truth by virtue of meaning and “Frege analyticity” would, for instance, come apart.

I mainly remind you of all this in order to contrast these distinctions with another distinction, namely that between a “metaphysical” and an “epistemic” notion of analyticity as suggested by Boghossian. He, too, charges Quine with not sufficiently distinguishing between these notions, but distinguishing between metaphysical and epistemic analyticity is not just bringing yet another two notions of analyticity into play. Rather, this distinction seems to cut across the others; at least the idea of truth in virtue of meaning clearly might come in a metaphysical as well as in an epistemic version.

What exactly, then, is the difference between metaphysical and epistemic analyticity? Boghossian takes analyticity in both senses to be a *semantic* notion and starts from the formula that analytic statements are *true in virtue of their meaning alone*. A statement is analytic in the *metaphysical* sense, he explains, if it “owes its truth value completely to its meaning, and not at all to the facts” (1996, 363). It is analytic in the *epistemic* sense “provided that grasp of its meaning alone suffices for justified belief in its truth” (ibid.).

Regarding the very idea of metaphysical analyticity, however, Boghossian wonders:

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**? (1996, 364).

The proponent of metaphysical analyticity need not deny this “meaning-truth truism” (1996, 364), that is, he need not deny that for any true statement there is a corresponding fact. Accepting the truism, however,

will drive him into some sort of *creationism*, Boghossian thinks; “he will want to say (...) that, in some appropriate sense, our meaning *p* by *S* makes it the case that *p*” (1996, 364f.). But that Boghossian regards as sheer mysticism. Consider the sentence

(S) Either snow is white or it isn’t.

It should be obvious that it was the case that either snow was white or it wasn’t *before* (S) came to mean that (cf. 1996, 365).¹

Notice, however, that there is a much simpler non-epistemic way of understanding the idea that a statement is true in virtue of its meaning alone: We might think of such statements as true statements whose meaning suffices to *determine* their truth value. The existence of such statements is perfectly compatible with any reading of the meaning-truth truism. For any true statement *S*, there can be a corresponding fact, no matter how this fact is created, no matter whether it is *metaphysically dependent* on our meaning *p* by *S* in any sense. The idea simply is that, for some statements, this is a *necessary* fact, a fact that, so to speak, obtains in every possible world. Therefore, the facts (or the world with respect to which the sentence is evaluated) cannot possibly make a *difference* to the truth value of such a statement. Given its meaning, its truth value is determined.² It seems to me that to claim the existence of such statements is to make an interesting metaphysical claim and that

1. There is a difference, however, between holding that our meaning *p* by *S* makes it the case that *p* and holding that our meaning *p* by *S* makes it the case that *S* is true. For one might very well think that its being the case that *p* is completely independent from our meaning *p* by *S* while at the same time holding that the primary truth-value bearers are *interpreted sentences*. Consequently, there would be no *truths* unless there are interpreted sentences, but the same would not necessarily hold for *facts*.

Boghossian, however, seems to think that avoiding creationism about facts commits us to a notion of *what a sentence expresses* as the primary truth-value bearer (cf. 1996, 365; 380). Given the observation just made, that would at least require further argument; moreover, such a view does not go very well together with his wanting to work with a picture that is “as hospitable as possible to Quine’s basic outlook” (1996, 361) and, therefore, agreeing to consider the *objects of belief* to be interpreted sentences. For, surely, the objects of belief should be the same as the primary truth-value bearers, shouldn’t they?

2. Such a determination relation between the meaning of an analytic statement and its truth value is of course presupposed by Boghossian’s own (epistemic) conception of analyticity. According to him, an analyticity “can only be false by meaning something other than what it means” (1996, 387). That is, after all, *why* knowing its meaning suffices for justified belief in its truth (cf. *ibid.*).

Boghossian's rejection of any metaphysical notion of analyticity thus might turn out to be a tiny bit hasty.

Moreover, it seems worth noting that this idea fits remarkably well with early logical positivist doctrine. The logical positivists would have called a sentence like (S) a "tautology" (Carnap 1930/1, 142) and declared it devoid of empirical content exactly because no possible state of affairs would *make a difference* to its truth. Thus Carnap explains:

If a compound sentence is communicated to us, e.g., 'It is raining here and now or it is snowing,' we learn something about reality. This is so because the sentence excludes certain of the relevant states-of-affairs and leaves the remaining ones open. In our example, there are four possibilities: 1. It is raining and snowing, 2. It is raining and not snowing, 3. It is not raining but it is snowing, 4. It is not raining and not snowing. The sentence excludes the fourth possibility and leaves the first three open. If, on the other hand, we are told a tautology, *no possibility is excluded but they all remain open* (Carnap 1930/1, 143, *emph. mine*).

According to Boghossian, the combination of these ideas with the attempted reduction of the notion of necessity to the conventions of language led the positivists to their version of creationism, i.e. linguistic conventionalism: "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" (1996, 365). I am uncertain as to the historical accuracy of these claims, but this is not the place to discuss this. Notice, however, that conventionalism about truth does *not* follow from conventionalism about necessary truth; what is generated by linguistic meaning, according to conventionalism, might be the *necessity* of a statement, not its *truth*. Truth might be given independently, that is, and necessity added.

Be that as it may, Boghossian holds that a statement's truth always depends both on its meaning and on the facts.³ The facts concerned might of course be very general facts about the world, as for instance

3. And insofar as this means nothing more than that he holds on to a non-creationist reading of the meaning-truth truism, there is of course no substantive conflict with the points just made.

the fact that everything is self-identical, but that does not make their statement any less factual. Boghossian thus holds a factualist view of analytic statements; according to him, they “are capable of genuine truth and falsity” (1996, 379), that is, they do have truth conditions and they do make claims about the world. But that, or so he claims, does not mean that they could not be known a priori, quite the contrary. For epistemologically speaking, analytic statements are statements where knowing the meaning alone is sufficient for justifiedly assigning truth values. Their *truth* is not independent of facts, but their *justification* is. Epistemically analytic statements, one might say, may not be *true* by virtue of meaning alone, but they are *justified by virtue of meaning alone*.

On the basis of an account of what it is to know the meaning of a statement, Boghossian then sets out to substantiate the claim that “the notion of analyticity can help explain how we might have a priori knowledge” (1996, 362). And he is very clear about that what he is after is a priori knowledge in a very *strong* sense. To have such knowledge, the following conditions must be fulfilled. First, the justification for the statements in question must be of an a priori nature in this strong sense. That means that holding such statements true must not only be “justified (...) without appeal to empirical evidence”, but also “that the justification in question is not defeasible by any future empirical evidence” (1996, 362). Second, the justification must of course be strong enough for knowledge. And third, the statements in question have to be in fact true; if all these conditions are fulfilled, holding a statement true amounts to a priori knowledge in the strong sense Boghossian is defending.

3. *Implicit definition, analyticity and the a priori*

The question, then, of course is how any sentence could be epistemically analytic. Boghossian: “Clearly, the answer to his 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” (1996, 366). It is here that implicit definition comes into the picture. Boghossian’s strategy is to start with the truths of logic and show for them how an implicit definition account of their meaning would make them epistemically analytic and knowable a priori.

Regarding the meaning of the logical constants, Boghossian claims, there is no alternative to some kind of conceptual role semantics (cf. 1996, 383). And this should take the form of what he calls *implicit definition*:

- (I) 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 (1996, 376).

Boghossian calls this an “implicit definition account of grasp” (ibid.), so principle (I) presumably is not only to account for how meanings are determined but also for what it is to “grasp” them. We could preliminarily formulate this idea as follows: it is by holding certain sentences true—I shall not bother with always mentioning the inferential alternative—that a speaker manages to attach a certain meaning to his logical constants and this, too, is what it is for him to “grasp” these meanings.⁴

How does an implicit definition work? Suppose we want to define the term F . An implicit definition proceeds by stipulating a sentence (or a set of sentences) S containing F to be true. Let’s call such a sentence (or set of sentences) $S(F)$ an “implicit definer”. Except F , all expressions in the implicit definer are already interpreted. The idea then is that given the interpretation of the other terms, the truth of the implicit definer suffices to determine the meaning of F . In the case of the logical constants, however, what seems to determine their meanings is that sentences of a certain form are true, regardless of the specific meaning of any other expression they contain. It therefore seems *prima facie* plausible to suggest that their implicit definitions would proceed via stipulation of their introduction and, maybe⁵, elimination rules.

4. Boghossian does not argue for the claim that there is no alternative to some kind of conceptual role semantics for the logical constants. What he seems to have in mind is that patterns of holding sentences true/inferences valid are the only *evidence* for interpreting expressions as logical constants; that, however, we could agree to while holding that the semantics of choice for the logical constants is truth-functional.

5. It is standardly held that elimination rules can be justified by introduction rules. Therefore, for purposes of implicit definition, introduction rules would seem sufficient—they “fix” or determine the meaning in question. Of course, if someone *used* an expression according to, for instance, the introduction rule for conjunction but not in accordance with its elimina-

There are some more or less classical problems connected with implicit definitions. The question is: when is an attempt at implicit definition successful, that is, when does it succeed in attaching a meaning to the term that is to be defined? The two best known problems here are those of *existence* and *uniqueness* of the meanings to be determined. We might fail in the attempt at implicit definition in at least two ways: either no meaning gets determined or more than one. In neither case has the term in question been defined. A clear case of failure through non-existence is Prior's famous tonk-operator which is supposed to be implicitly defined by the introduction rule for *or* and the elimination rule for *and*, thus allowing for derivation of absolutely everything from anything. It is in order to allow for implicit definitions failing in this way that Boghossian formulates (I) the way he does: "a particular constant means that logical object, *if any*, which makes valid a specified set of sentences and/or inferences involving it" (1996, 376, *emph. mine*). Boghossian does not seem to worry about uniqueness, however, and neither shall we. What is important here is that, according to Boghossian, implicit definition determines meaning only if there is an interpretation of the implicit definer that makes it true. Otherwise, no meaning is conferred by an attempt at implicit definition.

Now, Boghossian is not interested in how certain expressions actually were introduced into existing languages. This is not a historical model of what happened, for instance, when 'and' was introduced into English. Most probably, there was nothing that could be called an act of implicit definition. Rather, implicit definition provides a model for how certain expressions *could* be introduced and given a meaning. Implicit definers thus are sentences by means of which an expression *F* (that actually already has a certain meaning) *could* be given that very meaning by means of implicit definition. At the same time and more importantly, however, implicit definition is to be a model for what it is to *have* certain meanings, that is, it is to actually provide a semantics for *F*. To understand this, we have to remember that according to Boghossian, implicit definition is to be a model for a conceptual role semantics for the logical constants. And elsewhere he explains: "If expressions

tion rule, we would not have much idea how to interpret these utterances. What this points to, I think, is a certain tension between the idea of meaning "*fixation*" or *determination* by implicit definition and the conditions for *ascribing* certain meanings to expressions on the basis of how speakers use them.

mean what they do by virtue of the inferences they participate in, then some inferences are *constitutive* of an expressions's meaning what it does, and others aren't" (1994, 110). Implicit definition does provide a semantics for the logical constants if implicit definers are taken to be *constitutive of their meaning*. For F to have a determinate meaning there has to be a determinate set of sentences constitutive of its meaning. Now, if such a set $S(F)$ constitutes F 's meaning, then grasping F 's meaning consists in holding $S(F)$ true;⁶ holding its implicit definer to be false *ipso facto* means attaching a different meaning to F . And, as we saw above, in order for F to have any meaning at all, its implicit definer not only needs to be held true, it also has to be in fact true.

Now it looks as if the connection between implicit definition and a priori knowledge were fairly straightforward. If implicit definition is the right account of an expression's meaning, the following kind of justification might be given for its implicit definer:

1. If F means what it does, then its implicit definer $S(F)$ has to be true, for F means whatever in fact makes $S(F)$ true.
2. F means what it does.

Therefore,

3. $S(F)$ is true (cf. 1996, 376; 386).

Holding $S(F)$ true, it seems, can be justified from premises 1 and 2 above, and knowing them, in turn, does not seem to require anything but knowledge of meaning. Moreover, if F has a meaning, $S(F)$ is in fact true. Thus knowledge of meaning would suffice for knowledge of truth, and such knowledge would be a priori.

This is a wonderfully simple argument. Remember, however, that it rests on a number of highly theoretical assumptions. For instance, on the

6. Cf. 1996, 376: "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*" (first emph. mine). Providing an account of what it is to grasp the logical constants thus would seem one of Boghossian's most important reasons for adopting the implicit definition account of their meaning. Cf. also 1996, 383, where Boghossian explicitly speaks of an implicit definer S as a sentence "that I *must* hold true if S is to mean what it does" (emph. mine).

assumption that a conceptual role semantics is the semantics of choice for the logical constants. And on the assumption that a conceptual role semantics necessarily operates with meaning constitutive inferences. These assumptions are surely controvertible. What I shall do next, however, is investigate whether the classical Quinean criticisms really are ineffective against this semantic account of the a priori. More precisely, I would like to look at the argument from holism from *Two Dogmas* and the regress charge from *Truth by Convention*. In the course of these considerations, I shall have occasion to at least touch upon the second of the above assumptions, though.

4. *Holism*

The first of the Quinean arguments I would like to look at is what could be called the argument from holism. Maybe it is exaggerating a bit to call it an argument, but in the last sections of *Two Dogmas* we find a very suggestive metaphor for both the inter-relations between our beliefs and the relations between our beliefs and experience or the world. The metaphor is, of course, that of a “fabric” or a “web” impinging on experience only at its fringes. And it seems that if a Quinean “web-of-belief” model is correct, then nothing is a priori. It is the whole web, or at least large parts of it, that face experience together, and, or so Quine suggests, there are no reasons in principle to exclude any particular sentence or set of sentences therein from possible defeat by empirical evidence. Or, to put this the other way round, there are no principles by which to distinguish sentences with empirical content from sentences without. Especially, there are no principles to distinguish highly obvious empirical sentences from non-empirical ones. Though empirical content does come in degrees on such a model, there is no reason to say that there are any sentences completely without.

I would like to argue for two things in this section: First, if we think, as most people do these days, that epistemological holism is very plausible for observational contexts, a certain tension arises. For it will be very difficult to prevent such holism from spreading. Thus, the distinction most important to Boghossian’s account, that between meaning constitutive and other sentences becomes threatened. As long as it cannot be substantiated, however, the Quinean challenge stands unanswered. Second, however, I would like to ask how compelling the

development of the idea of implicit definition into a general theory of meaning determination really is.

Boghossian does not discuss epistemological holism in the analyticity papers. He does endorse some form of it in other papers, though. There, he uses an argument from what he calls “the holistic character of belief fixation” (1991, 78, see also 1989, 539 f.) against information-theoretic semantics. “Under normal circumstances”, Boghossian explains, “belief fixation is mediated by background theory—what contents a thinker is prepared to judge depends upon what other contents he is prepared to judge. And this dependence is again typically, arbitrarily robust: just about any stimulus can cause just about any belief, given a suitably mediating set of background assumptions” (1991, 78). Applied to observation contexts, holism of belief fixation means the following: Take an observational belief that p , for instance, the belief that there is a cow in front of you. There is no situation in which this belief could not be caused, given a set of sufficiently deviant background beliefs. Thus, there is equally no situation in which it could not be prevented from being formed by sufficiently deviant background beliefs.

In these contexts, Boghossian takes care to note that this is psychological holism, not the disputed doctrine of semantic holism. It seems pretty obvious, though, that the holism of belief fixation nevertheless creates a certain tension with the view that there are meaning constitutive inferences or sentences. Or at least it would, if such holism were endorsed in an unrestricted form. Consider, for instance, the belief that the creature in front of you is a cow. Assume, too, that in order for this to be a cow-belief, you also have to believe that the creature in front of you is an animal. An unrestricted belief holism would amount to the claim that there are situations in which you can believe that the creature in front of you is a cow without forming the belief that it is an animal—provided you hold a sufficiently deviant background theory. But this clashes with the assumption that the cow-animal connection is meaning constitutive. More generally, if in just about any situation, including other beliefs, a belief that p can be formed, and, conversely, nothing, including other beliefs, necessarily prevents a belief that p from being formed, then there are no particular inferences or sentences such that a speaker has to hold them true in order to have a belief that p . Thus, an unrestricted holism of belief fixation would not seem to go together with the existence of meaning constitutive sentences or inferences.

Now, Boghossian nowhere subscribes to an unrestricted holism of

belief. What should be clear, though, is that he needs to restrict his holism by providing a distinction between empirical and non-empirical sentences. Without such a distinction, that between meaning constitutive sentences and others is lost as well. And one might think that, though they do not coincide, distinguishing meaning constitutive sentences from others could be a first step towards a viable distinction between the empirical and the non-empirical.⁷ Be that as it may, to say that belief holism needs to be restricted in order to accommodate meaning constitutive sentences is just to repeat the Quinean challenge: How are non-empirical sentences to be distinguished from others? What is the substantive content of such a restriction upon belief holism? Boghossian admits that this is a very good question, and also that he does not really have a good answer to it (cf. 1996, 383). Fair enough. But in the absence of a good reply, we should, also in all fairness, consider the Quinean challenge as unanswered.

Boghossian does not agree. As noted above, he only aims at meeting the Quinean challenge insofar as it does not depend on the indeterminacy thesis (cf. 1996, 361). But, or so he tries to convince us, “anyone who rejects radical indeterminacy of meaning must believe that a distinction between the meaning constituting and non-meaning-constituting can be drawn” (1996, 383). Why? Boghossian’s argument runs like this: A conceptual role semantics is the only plausible semantics for the logical constants. According to a conceptual role semantics, an expression has a determinate meaning only if there is a determinate set of meaning-constitutive sentences or inferences. If Quine is right and, in Boghossian’s words, “there is no fact of the matter as to which of the various inferences involving a constant are meaning-constituting” (ibid.), then the constants do not have determinate meanings. “And that, again, is just the dreaded indeterminacy of meaning on which the critique of analyticity was supposed not to depend” (ibid.).

7. This idea is, in fact, suggested by the overall architecture of Boghossian’s rationalistic project: Starting with the implicit definers for the logical constants, he tries to work his way upwards to account for the a priori nature of all logical truth and, by means of the notion of Frege-analyticity, all conceptual truth (cf. 1996, 366 ff.). The way he presents the project here, implicit definition seems to be supposed not only to take care of all logical truth, but also the epistemic analyticity of sentences like ‘Whatever is red all over is not blue’. It is, however, not spelled out how the account is supposed to extend beyond implicit definers for the logical constants.

Quite clearly, what Boghossian seems to have in mind when he says that the critique of analyticity is supposed not to ‘depend on indeterminacy’, is not just that it is not supposed to use *Quine’s* thesis of the indeterminacy of translation⁸ as a *premise*. Rather, *no* kind of ‘indeterminacy of meaning’ is supposed to *either be its premise or its conclusion*. Now, in the above argument ‘indeterminacy of meaning’ follows only given certain assumptions Boghossian subscribes to: That the semantics of choice for the logical constants is a conceptual role semantics and that such a semantics requires a distinction between meaning constitutive and other sentences or inferences. Both assumptions, as remarked above, are controvertible; in fact, I believe that both of them are false. Not being able to argue this in any detail in this paper, I shall confine myself to the following brief comments, however.

Grant, for the sake of argument, a conceptual role semantics. Does this require a distinction between meaning-constitutive and other sentences or inferences? The reason Boghossian thinks it does, presumably, is “the old hat” (Block 1993, 1) familiar from the recent wars on semantic holism:⁹ Our basic semantic idea is that what a speaker means by an expression is determined by the inferences it participates in. This determination cannot be holistic, the argument goes, for if it were no two speakers could ever disagree about anything. Ergo, it is only certain inferences that determine meaning, not all of them; those that do are the meaning-constitutive ones. This argument relies on hidden premises about the nature of the determination relation supposed to hold between inferences and meaning, however. It goes through only on the assumption that this relation is a *one-one relation*, but can be blocked by allowing holistic determination to be *many-one*.¹⁰ If inferences can determine meaning holistically, however, a conceptual role semantics does not necessarily commit us to meaning-constitutive inferences.¹¹ Even if we grant Boghossian a conceptual role semantics for the logical constants, it therefore does not immediately follow that “anyone who rejects radical indeterminacy of meaning [in Boghossian’s sense]

8. As argued for in Chapter 2 of *Word and Object* (1960). According to *Word and Object*, the meaning of the logical constants does *not* fall within the realm of the indeterminate. Nor does Quine adopt a conceptual role semantics for them.

9. As documented, in part, in Fodor/Lepore 1992; 1993; cf. also Boghossian 1993.

10. This point is argued in detail in Pagin 1997.

11. In Glüer 2002, I argue that this, in fact, is an advantage a holistic semantics has over non-holistic ones exactly because it relieves us of the Quinean challenge discussed above.

must believe that a distinction between the meaning constituting and non-meaning-constituting can be drawn” (1996, 383). Pending further argument, it is fair, after all, to consider the Quinean challenge unanswered.

So far, I have argued that there is a certain tension between belief holism and the doctrine of implicit definition. This tension arises because of the latter’s apparent commitment to meaning-constitutive sentences. In the remainder of this section, I would like to raise even more trouble for this development of the idea of implicit definition into a general theory of meaning determination. The question I would like to raise could be put like this: How plausible is it to hold that an implicit definer is a meaning-constitutive sentence?

Looking at it diachronically, from the perspective of expressions to be introduced into the language, so to speak, this is far from obvious. Quine certainly thought that expressions could be *introduced* by implicit definition. He called that “legislative postulation” (cf. 1954, 117 ff.). But being a postulation, he argued, is “a trait of events and not of sentences” (1954, 119) and concluded: “Legislative postulation contributes truths which become integral to the corpus of truths; the artificiality of their origin does not linger as a localized quality, but suffuses the corpus” (1954, 120). Theoretical terms in scientific theories would seem to provide a case in point. These might certainly be seen as being introduced by implicit definition, but afterwards, their implicit definers face experience along with the rest of the theory’s sentences. That is, such sentences, once completely interpreted, can actually turn out to be empirically *false*. But then, they are clearly not implicit definers anymore. Once interpreted, it thus seems, the implicit definer loses both its claim to having to *be* true and its claim to having to be *held* true (in order for it to have a certain meaning). Once interpreted, it is just a sentence as any other. The possibility of actually introducing expressions by means of implicit definition therefore does not seem to force meaning constitutive sentences upon us.

What about looking at it from the perspective of an expression *F* that already is part of the language and has meaning? Are there sentences that have to be held true in order for *F* to mean what it does? Take the case of the logical constants. What does a speaker have to hold true in order for *F* to mean one of them? Inferences of the form specified by that constant’s introduction and elimination rules, presumably. But what exactly does that mean? The most plausible requirement would

seem to be that the speaker has to hold sentences true in accordance with these rules. For instance, he has to hold 'A and B' true when he holds A true and he holds B true. But saying that the set of sentences or inferences thus specified is constitutive for a certain meaning is to say that *any* of the indefinitely many inferences sharing the specified form is meaning constitutive. To mean one of the logical constants by an expression *F* thus would require absolutely impeccable reasoning with *F*. *Any* mistake automatically would prevent *F* from meaning this constant. Given the kind of creatures we are, this cannot be right, I am afraid; our logical (in)capacities would prevent most of us from ever expressing anything remotely resembling a logical constant.

But wasn't it Quine who taught us that the strongest possible evidence of bad interpretation is seeming to hold a basic logical truth false? Certainly (cf. Quine 1960, 57 ff.). If a speaker seems to deny basic logical truths, we should definitely check our interpretation. But is that the same as saying that in order to mean a specific constant every inference of elimination or introduction form needs to be gotten right? I doubt it. Consider a fairly normal speaker. We think he means conjunction by 'and'. What would count as good evidence for this interpretation's being bad? What would it be to interpret him as denying basic logical truths? Our speaker might not have sentences formulating those truths in his repertoire, and even if confronted with them, he might not be able to assign the right truth value. That by itself would *not* be a reason to doubt that he means conjunction by 'and', however. What he does mean by 'and' would rather be a matter of his holding true sentences according to the right inference patterns. But even here, getting the occasional one wrong would not necessarily be a reason for doubting that he means conjunction. Of course, these are matters of degree, but it seems fairly clear that our speaker would have to make quite a lot of mistakes of a rather inexplicable and destructive nature before we would have good evidence for bad interpretation.

Moreover, it is well known that the reasoning powers even of college students often in fact seem to be amazingly bad. Tests like the so-called selection task show that things as fundamental as reasoning with 'if-then' according to *modus ponens* frequently seem to go wrong (as reported, e.g., in Stich 1990, 4 f.). But then, conceiving of such basic inferences as constitutive of the meaning of the logical constants would commit us to saying that in practice, hardly anyone ever manages to express them. That does not seem right, however; despite these

reasoning failures, it in practice most of the time does seem plausible to interpret normal usage of ‘if-then’ as a conditional fulfilling *modus ponens*.

This is of course not to claim that ‘if-then’, as used in natural language, expresses the material conditional. It does seem plausible to think that it expresses a conditional fulfilling *modus ponens*, however, and, therefore, that its meaning is not determined in accordance with the model of implicit definition. Mistakes regarding inferences of introduction or elimination form are possible without change of meaning. And there is no reason to think that this would not generalize to other implicit definers or, for instance, the speaker who does use ‘if-then’ to express the material conditional. *His* usage certainly would be tellingly different from everyday usage of ‘if-then’, but it need not be impeccably so.

An account working with meaning constitutive inferences thus seems too inflexible to deal with actual speakers in the most plausible way. It assigns *too much weight* to certain types of mistakes, making them, in effect, categorically impossible. Even though simple logical mistakes do seem to be the best evidence for bad interpretation, even the best evidence seems more plausibly construed as less than conclusive. To interpret actual practice in a plausible way, it therefore seems preferable to work with a model according to which (occasional) basic logical mistakes are possible.

One might feel like replying that a speaker can ‘follow a rule’ like an introduction or elimination rule without getting every single application right. In a way, that is just my point. For the defender of meaning-constitutive sentences, however, it means that we are back to square one. For the distinction between meaning-constitutive and other sentences or inferences was supposed to deal with exactly this problem: how to be able to countenance differences of application (mistakes) while ascribing the same meaning. Allowing mistakes of introduction or elimination merely postpones the problem; the question now becomes: which of the applications of an introduction or elimination rule are constitutive for its being ‘followed’ and which aren’t?

To sum this discussion up: The challenge from epistemological holism and the web of belief model is to come up with a viable principled distinction between meaning-constitutive and other sentences. In the absence of such a distinction, the Quinean challenge must be considered as unanswered. Moreover, the distinction seems neither forced upon us by the possibility of introducing expressions by implicit defi-

nitions nor does it seem to be required by considering logical mistake the strongest evidence for bad interpretation. On the contrary, accounts dispensing with meaning-constitutive sentences seem to be able to make better sense of actual speakers. The case against the web-of-belief model and for meaning-constitutive sentences thus can, at best, be considered as still open.

5. *Conventionalism*

I would now like to leave this issue and look at another aspect of Boghossian's model, its conventionalism. Independently of any results reached so far, our question is whether this conventionalism can be defended against Quine's attack on a very similar conventionalism in *Truth by Convention* (1935).

According to Boghossian, the logical constants get their meanings by "our conventionally stipulating that certain sentences (or inferences) involving them are to be true" (1996, 380 f.). All this amounts to, is a *conventionalism about meaning*. For remember, interpreted implicit definers are factual sentences, according to Boghossian. What they express are language independent factual truths constitutive of a certain meaning. What is conventional is only *which expression* comes to have that meaning. Nevertheless, this model seems to offer an *explanation* for why we think of certain sentences as a priori: These are sentences that get their interpretations by being conventionally stipulated true. And as we saw before, being a definer on this model is not a passing trait; a sentence continues to have it as long as it continues to have the same meaning. To be held true by convention thus might well seem to be a distinguishing trait of meaning-constitutive sentences.

It is exactly such a conventionalism about the meanings of the logical constants that is under attack in *Truth by Convention*. There, Quine argues that the conventionalist is faced with a dilemma: His view either leads into an infinite regress, or, if modified so as to avoid the regress, loses all its explanatory power, all its power, that is, to make any real distinction.

Let's quickly look at the first horn of this dilemma, the infinite regress. According to Boghossian,

the meaning for ‘and’ is fixed by our stipulating that the following inferences are to be valid:

A and B
A

A and B
B

A, B
A and B (1996, 381).

To be more precise, however, these are not inferences but inference-schemata. What needs to be stipulated is that *all* inferences having one of these forms are valid. Since there is an infinite number of such inferences, they cannot be stipulated true one by one. A general stipulation is needed. Any such general stipulation, however, needs to make use of a sentence containing already interpreted logical expressions—hence, regress.

It is widely thought that this Quinean argument, even though clearly valid against any sort of explicit conventionalism, can be easily circumvented by “going implicit”. That is, the argument is taken to bite only, if the conventions in question have to be *explicitly stated* in order to be adopted. “Implicit conventions”, however, do not have to be stated; they can be in force for and observed in behavior even though no one ever actually formulated them. By means of understanding the conventionalism about the meanings of the logical constants as one employing implicit conventions, then, it might seem that it can be saved from Quine’s regress charge. And this seems to be exactly what Boghossian suggests to do, too. He writes: “surely, 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**” (1996, 381).

We must not forget, however, that Quine anticipates the conventionalist’s “going implicit”. He readily admits that implicit conventionalism would fit actual language acquisition and use much better than explicit conventionalism. However, “going implicit” is not as easy as it might seem. For what exactly does it mean that something is an implicit convention? Or that behavior is following an implicit convention? What would, to speak with Boghossian, the “sort of fact about **x**’s behavior with **e**” exactly be that makes it convention-governed behavior? Of course, implicit definers would be sentences very firmly accepted. But, again, that is a characteristic they share with a lot of sentences that we would not want to classify as held true by convention, especially highly

obvious empirical sentences. These are held true with equal regularity. It's the classical problem for the implicit conventionalist: How to distinguish convention-governed from merely regular behavior. And this is, of course, the second horn of Quine's dilemma: Unless the conventionalist can come up with a relevant behavioral difference, he risks explanatory emptiness. The notion of a linguistic convention reduces to an "idle label" if there is no difference between conventionally holding true and other forms of firm acceptance (cf. Quine 1935, 106).¹² "Going implicit" thus is more difficult than it might *prima facie* appear. It cannot be done by merely vaguely indicating "some sort of fact about (...) behavior" with an expression that would make it convention-governed; that just leaves us exactly where *Truth by Convention* got us, that is, in need of knowing the kind of fact that would do it.¹³

Moreover, this is only part of the story of *Truth by Convention*. That there is more is often overlooked, but this forgotten part of *Truth by Convention* is no less crucial to our investigation than the first. For it seems to show that quite independently of its conventionalism, an implicit definition account of the meanings of the logical constants is impossible; the conventions in question would not only have to be implicit, they would also have to be directed at something that does not exist.

Let me explain. The doctrine of implicit definition says: "It is by arbitrarily stipulating that *certain sentences* of logic are to be true (...) that we attach a meaning to the logical constants" (Boghossian 1996, 376, *emph. mine*). What gets stipulated to be true, according to implicit definition, are *sentences*. That is, no matter whether these definitions are adopted explicitly or implicitly, there need to be sentences that we could at least *ex post* identify as such that stipulating them to be true would result in the meaning assignment. However, as Quine argued, *there are no such sentences*. For they would have to be general, that is,

12. In Glüer 2002a, I go through a number of influential proposals for distinguishing implicit norms, rules, or conventions and try to show that they all, ultimately, are stuck between the horns of Quine's dilemma. I also give some reasons why, as long as we are after meaning or content determining norms, rules, or conventions, the dilemma might actually be inescapable.

13. Hinting at dispositions, as Boghossian does (without seeming to endorse his own hint, however; cf. 1996, 381 f.), surely does not help much — for regarding our dispositions to hold them true, there does not seem to be any difference between meaning constitutive and other firmly accepted sentences.

they would have to contain already interpreted expressions of the very kind they are to implicitly define.¹⁴

That this point is easily overlooked may be due to the fact that Quine presents it as one more aggravation for the conventionalist: He has to go for conventions that, at the point of adoption, cannot even be formulated. Otherwise, even implicit conventions do not help with the regress. Quine: “When we first agree to understand ‘Cambridge’ as referring to Cambridge in England, failing a suffix to the contrary, and then discourse accordingly, the role of linguistic conventions is intelligible; but when a convention is *incapable of being communicated until after its adoption*, its role is not so clear” (1935, 106, *emph. mine*). However, for an implicit definition account of the meaning of the logical constants, the point seems simply fatal. Sentences such that stipulating them to be true would result in giving the logical constants their meanings do not exist—neither before nor after the adoption of any convention, be it explicit or implicit, and regardless of our dispositions to use the expressions in question. Of course, in some sense the constants do get their meaning by their use. But *secunda facie* it does *not* seem plausible to use implicit definition to account for how that works.

To sum this discussion up: Again, the Quinean challenge to substantiate the crucial difference between meaning constitutive and other sentences, here in the guise of a distinction between those held true by convention and others, is left unanswered. Moreover, even if the distinction could be substantiated, that would not seem to help. For the forgotten part of the argument from *Truth by Convention* seems to generally rule out an implicit definition account of the meaning of the logical constants.

6. *The return of the facts*

So far, I have made use of classical Quinean arguments to convince you that his case against analytic apriorities should not be underestimated and, in fact, seems far from lost. I would like to conclude

14. And analogously, of course, for stipulating inferences valid: Whether implicitly or explicitly, we need to stipulate an indefinite number of them to be valid. According to the model of implicit definition, this is done via a *specification* of the inferences. The stipulation, that is, requires a linguistic expression representing the inferences in question. And this, again, would need to be general, i.e. to contain expressions of the very kind to be defined.

by raising yet another worry. For there seems to be a question as to whether, even granted everything else, implicit definers would really be analyticities—even if we understand analyticity in the epistemic sense recommended by Boghossian. And that would, of course, destroy the prospects of giving an analytic explanation of the a priori by means of implicit definition.

Recall Boghossian's epistemology for implicit definers:

1. If F means what it does, then its implicit definer $S(F)$ has to be true, for F means whatever in fact makes $S(F)$ true.
2. F means what it does.

Therefore,

3. $S(F)$ is true (cf. 1996, 376; 386).

Recall also, that $S(F)$, according to Boghossian, is a factual sentence. Knowing that $S(F)$ is true thus amounts to knowing a fact about the world. At the same time, $S(F)$ is supposed to be epistemically analytic, justified in virtue of its meaning alone: All we need to know in order to be justified in our belief in $S(F)$'s truth is its meaning. That is, even though $S(F)$ itself is a factual sentence, its justification is *not* to depend on or involve matters of fact.

But how could that be so given the model of implicit definition employed by Boghossian? Look at premises 1 and 2 together. That F means what it does depends on there in fact being something that makes $S(F)$ true, according to premise 1. Only if there is such a fact does F have any meaning. Being justified in believing premise 2, therefore, requires being justified in believing that there is something that makes $S(F)$ true. Moreover, it requires being justified in believing *that* $S(F)$.

To put this in terms of propositions, let P be the proposition that $S(F)$ expresses. On Boghossian's account, it seems to me, being justified in believing that $S(F)$ means P presupposes being justified in believing P . Of course, it then *follows* from being justified in believing that $S(F)$ means P , that I am justified in believing P . However, believing P cannot be *justified* this way; for knowledge of the meaning of the sentence already requires such justification and, therefore, cannot provide it.

This would mean two things. First, implicit definition does not *explain* a priori knowledge. If there is a priori knowledge of the proposi-

tion expressed by an implicit definer, its justification has to take some independent way.¹⁵ And second, implicit definers are *not epistemically analytic*. Knowing their meaning does not provide justification by virtue of meaning alone, for knowledge of meaning itself turns out to already require knowledge of facts about the world. It does not seem, therefore, as if the distinction between metaphysical and epistemic analyticity could be sustained—at least on the Boghossian model, the facts do return, and with a vengeance.

Conclusion

I have considered two classical Quinean ways of attacking the version of the analytic theory of the a priori that Boghossian is trying to salvage. On both counts, the attacks still seem very powerful. Epistemological or belief holism is not easily moderated by means of any principled distinction between empirical and non-empirical sentences, thus endangering the a priori in general. More particularly, there seems to be good reason to prefer accounts of meaning that do not require a distinction between meaning constitutive sentences or inferences and others. Conventional implicit definition accounts of meaning, even more particularly, have additional liabilities. Conventionalism, especially when it has to go implicit, needs to be saved from being an idle label. And in the crucial case of the logical constants, implicit definers do not even exist. Last, but not least, I have argued that even independently of the Quinean attacks implicit definers are, after all, not (epistemically) analytic. The prospects of implicit definition to provide an analytic explanation of a priori knowledge are, I am afraid, rather bleak.*

15. Bonjour makes a similar point; cf. Bonjour 1998, 50 f.

* For helpful comments on earlier versions of this paper I would like to thank Paul Boghossian, Sten Lindström, Peter Pagin, Åsa Wikforss and audiences in Berlin, Konstanz, Umeå and at a Birkbeck College (University of London) study weekend at Cumberland Lodge.

REFERENCES

- Block, Ned 1993: "Holism, Hyper-analyticity and Hyper-compositionality", *Mind and Language* 8, 125.
- Boghossian, Paul 1989: "The Rule-Following Considerations", *Mind* 98, 507–549.
- 1991: "Naturalizing Content", in: B. Loewer and G. Rey (Eds.), *Meaning in Mind. Fodor and his Critics*, Oxford and Cambridge, Mass., 65–86.
- 1993: "Does an Inferential Role Semantics Rest Upon a Mistake?", *Mind and Language* 8, 27–40.
- 1994: "Inferential Role Semantics and the Analytic/Synthetic Distinction", *Philosophical Studies* 73, 109–122.
- 1996: "Analyticity Reconsidered", *Noûs* 30, 360–391.
- 1997: "Analyticity", in: B. Hale and C. Wright (Eds.), *A Companion to the Philosophy of Language*, Oxford, 331–368.
- 2000: "Knowledge of Logic", in: P. Boghossian and C. Peacocke (Eds.), *New Essays on the A Priori*, Oxford, 229–254.
- Bonjour, Laurence 1998: *In Defense of Pure Reason*, Cambridge.
- Burge, Tyler 1992: "Philosophy of Language and Mind 1950–1990", *The Philosophical Review* 101, 3–51.
- Carnap, Rudolf 1930/1: "Die alte und die neue Logik", *Erkenntnis* 1, 12–26, quoted from the translation by I. Levi: "The Old and the New Logic", in: A. Ayer (Ed.), *Logical Positivism*, New York 1959, 133–146.
- Fodor, Jerry and Ernest Lepore 1992: *Holism. A Shopper's Guide*, Oxford.
- Fodor, Jerry and Ernest Lepore (Eds.) 1993: *Holism. A Consumer Update*, Amsterdam.
- Glüer, Kathrin 2002: "Alter Hut kleidet gut. Zur Verteidigung des semantischen Holismus", in: G. W. Bertram and J. Liptow (Eds.), *Holismus in der Philosophie*, Weilerswist, 114–128.
- 2002a: "Implizites und Explizites Regelfolgen", in: U. Baltzer and G. Schönrich (Eds.), *Institutionen und Regelfolgen*, Paderborn, 157–176.
- Harman, Gilbert 1996: "Analyticity Regained?", *Noûs* 30, 92–400.
- Pagin, Peter 1997: "Is Compositionality Compatible with Holism?", *Mind and Language* 12, 11–33.
- Quine, Willard Van Orman 1935: "Truth by Convention", in his *The Ways of Paradox and other Essays*, sec. ed., Cambridge, Mass. 1976, 77–106.

- 1951: “Two Dogmas of Empiricism”, in his *From a Logical Point of View*, sec. ed., Cambridge, Mass. 1980, 20–46.
 - 1954: “Carnap and Logical Truth”, in his *The Ways of Paradox and other Essays*, sec. ed., Cambridge, Mass. 1976, 107–132.
 - 1960: *Word and Object*, Cambridge, Mass.
- Stich, Stephen 1990: *The Fragmentation of Reason*, Cambridge, Mass. and London.