Philos Stud
DOI 10.1007/s11098-006-9029-0

ORIGINAL PAPER

What we know now that we didn't know then: reply to critics of *The Age of Meaning*

Scott Soames

Abstract Author's response to critical essays by Brian Weatherson, Alex Byrne, and Stephen Yablo on *Philosophical Analysis in the Twentieth Century, Volume 2 The Age of Meaning*

Keywords Philosophical analysis · Ordinary language · Necessary aposteriori · Contingent apriori · Truth · Meaning · Perception · nowledge · Skepticism · Behaviorism · Indeterminacy · Wittgenstein · Ryle · Strawson · Austin · Quine · Davidson · Kripke

The ordinary language school: reply to Weatherson

Brian Weatherson distinguishes the book I wrote from the one he wishes I had written. The former is a series of intensive case studies of the most important and revealing work of the period, evaluated with an eye to separating strengths from weaknesses, and extracting enduring lessons. The latter would have focused more on the broader philosophical community—who influenced whom, why some were acclaimed and others neglected, and how work of thinkers I didn't discuss—e.g. Goodman, Sellars, and Chisholm—related to work of those I did. Although such points have merit, their inclusion would, as Weatherson notes, have required dropping material essential to my purpose.

With this bow to reality, he turns to the book I did write. However, it soon becomes clear that he is still dreaming of his phantom volume. In discussing Grice, he raises two main questions.

(i) Is the ordinary language school of philosophy really dead?
(ii) If so, did it die as I suggest—with Grice delivering the final blow to an approach that was already faltering under two decades of accumulated difficulties?

S. Soames (✉)
Mudd Hall of Philosophy, USC School of Philosophy, 3709 Trousdale Parkway, Los Angeles, CA 90089-0451, USA
e-mail: soames@email.usc.edu

Although the questions are coy, there is a straightforward sense in which the answer to both is ‘‘yes’’. When a school of philosophy dies the work done in it doesn't lose all importance or influence. Despite the death of logical positivism, echoes of it can be found in subsequent work—e.g., Quine's holistic verificationism, Van Fraassen's constructive empiricism, Dummett's intuitionist semantics, Hare's non-cognitivism, and Blackburn's expressivism. Although these writers aren't positivists, aspects of their thought carry on certain themes of positivism.

The same point can be made about the ordinary language school. Once a reasonably cohesive and self-conscious approach to philosophy characterized by rough and ready adherence to T1–T3, the school flourished for a time, and eventually died.

T1. All philosophical problems are linguistic, arising from the misunderstanding and misuse of language.
T2. Meaning, in so far as it is central to philosophy, is not to be studied from a theoretical or scientific perspective. Instead, philosophers must attend to subtle aspects of language use, and show how misuse of words leads to particular philosophical confusions.
T3. Illuminating philosophical analyses almost never state necessary and sufficient conditions for the application of a term; instead they trace the intricate and philosophically significant web connecting the use of that term to the uses of other, related terms.[1]

To my knowledge, no major philosopher today adheres to T1–T3. Kripke, whom Weatherson dubs ‘‘the last ordinary language philosopher,’’ repudiates all three. His affinity with ordinary language philosophy consists primarily in respecting what we pre-theoretically think, and distrusting unmoored, revisionary speculation. In this he is closer to Moore than to ordinary language philosophers. What about adherents of the ‘‘Canberra Plan’’? Although they do give dismaying signs of accepting T1, they don't accept T2 or T3—and no one would mistake their philosophical outlook for that of Ryle or Austin. There are, of course, some similarities between those Weatherson mentions and some ordinary language philosophers, but that just shows that the death of an approach to philosophy doesn't mean that all the concerns of its practitioners vanish along with it, never to appear again in other forms. As for his observation that ‘‘the analysis of knowledge industry ... seemed to putter along much the same before and after the demise of ordinary language philosophy,’’ surely, it is not being suggested that anyone who offers an analysis of an everyday notion is, thereby, an ordinary language philosopher. In the case of knowledge, a passing acquaintance with some of the more elaborate and technical attempts to solve the Gettier problem should be enough to disabuse one of the idea that ‘‘the analysis of knowledge industry’’ of the seventies and eighties was the work of the ordinary language school. For one thing, the search for necessary and sufficient conditions for knowledge that characterized the industry cut strongly against the grain of T3.

In short, the ordinary language school is no more—which does not mean that its classics are consigned to oblivion. Although the leading proponents of the school were sometimes guided by mistaken methodological and meta-philosophical views, they were also good philosophers with real insights, whose work didn't always fully conform to those views. In the book, I stress this point about Wittgenstein, Ryle, and

[1] See Wittgenstein on games and family resemblances (discussed on pp. 16–17, 26–27), and Ryle on conceptual analyses (discussed on pp. 79–80).

Austin, but Weatherson's example of Strawson is also apt. *Introduction to Logical Theory* contains valuable ideas that are largely independent of the flawed methodological perspective that guided it. All too often, however, similarly flawed perspectives did adversely affect the work of even the best ordinary language philosophers. Strawson's defective performative analysis of truth—singled out by Grice and discussed at length in my chapter 5—is a classic example of what can go wrong when the systematicity of meaning, and pragmatic implicatures, are overlooked. A similar error infects one of Austin's main arguments (discussed on pp. 174–178) for a central thesis of *Sense and Sensibilia*—that knowledge of objects arising directly from perception often does not rest on evidence. A different kind of obliviousness to the systematicity of meaning—the so-called Frege-Geach point—is shown in chapter 6 to tell against Hare's performative analysis of *good*, while Ryle's *Dilemmas* is found to contain serious errors due to his adherence to T1 and T3, and to his the assimilation of necessity and apriority to analyticity. Recurring flaws like these doomed the school, despite the fact that it produced some enduring classics of the analytic tradition.

I now turn to Weatherson's complaints about my treatment of particular philosophers, starting with Ryle. His first complaint focuses on Ryle's argument that seeing a tree is not an inner physiological state, because my knowing that I see a tree doesn't require knowing physiological facts. I criticize this argument for conflating a metaphysical question about the essential nature (necessary identity) of my seeing the tree with an epistemological question about what knowing that I see a tree involves. Although Weatherson agrees with the criticism, he takes the mistake to be "basically verbal," since he thinks that Ryle's position can trivially be revised to avoid it. The revision makes the functionalist point that since seeing a tree is a multiply realizable state, it is not identical with any of its concrete realizations. However, this is *not* a revision Ryle can accept. The crucial point for Ryle is not just that seeing a tree cannot in general be identified with a single neurological realization, but also that particular instances of the state, like my seeing a tree a minute ago, can't either. This belies functionalism. For the functionalist, talk about our perceptions is analogous to talk about an automobile's carburetor. Just as talk about what a carburetor does is functionalist talk about causally efficacious internal structure—each particular instance of which is identical with a physical particular under the hood—so talk about our perceptions is supposed to be functionalist talk about causally efficacious internal structure, each token of which is identical with a particular neurological event. Since Ryle would reject any such identification as a category mistake, his view is not salvageable in the way Weatherson imagines.

Weatherson's next point is better. He observes that I wrongly labeled, as *a form of verificationism*, Ryle's argument that if the beliefs and desires of an agent A were internal states, unobservable to other agents, then such agents could never know, or even plausibly conjecture, that A believed or desired anything. The argument assumes that hypotheses about the mental that can't be conclusively verified or falsified by observation can't be *known* or *justifiably believed*. However, it does *not* assume that they are meaningless, and so should not have been labeled, *a form of verificationism*. Weatherson is right—the label was sloppy—even though the critique of Ryle's argument was sound.

Weatherson's final point about Ryle concerns his behaviorism. He begins with the humorous acknowledgement that, *of course*, Ryle must be counted as a behaviorist because what we mean in philosophy by 'behaviorist' is determined by the reference-fixing description 'one who shares Ryle's view of the mental'—as if this were an arbitrary linguistic stipulation, with little behind it. He ends by acknowledging that in fact there is nothing arbitrary about such a stipulation; *of course*, Ryle was a behaviorist, since he was not an eliminativist about the mental and he rejects the view that mental states are brain states—leaving nothing else for them to be but behavioral dispositions. Between these two acknowledgements, Weatherson criticizes me for calling Ryle a behaviorist. But if Ryle really was a behaviorist, what is wrong with calling him one?

Weatherson seems to be worried about three points: (i) Ryle didn't think of himself as a behaviorist and didn't like being called one; (ii) Ryle's dispositional analyses of mental terms freely invoke other mental terms, without (as I emphasize) any systematic effort to show that the chain of analyses bottoms out in entirely behavioral and non-mentalistic language; and (iii) Ryle is in general no friend of *reduction*—in the sense of the reduction of one theory to another (as in the logicist reduction discussed in Volume 1). All of these points, are, I believe, of a piece—not only with one another but with Ryle's insistence that it is wrong to expect philosophical analyses to yield necessary and sufficient conditions for the application of a term. In keeping with this view, he would reject the claim that there are any truths $\ulcorner \forall x\ (Mx \Leftrightarrow Bx) \urcorner$ that count as definitions or analyses, where M is a formula containing a mentalistic notion like 'believe' and B is an entirely behavioral and physicalistic formula. Moreover, Ryle would see in this rejection the rejection of behaviorism, and the explanation of (ii) and (iii).

It is, I think, because he thought of behaviorism in this narrow sense that Ryle didn't regard himself as a behaviorist. Nevertheless, he *was* a behaviorist in the properly more expansive sense of believing that talk of the mental is, in the end, nothing more than talk of behavior.[2] The problem is that the dispositional analyses of mental terms on which he spilled so much ink do *nothing* to establish this. Perhaps he thought that his critique of the Cartesian picture—and of conceptions of mental states as causally efficacious internal states of any kind—already established that talk of the mental is nothing more than talk of behavior. Recall his famous analogy (discussed on pp. 94–96) between (i) the view that talk of the mental is talk about something internal that stands behind observable behavior and (ii) the view that talk of the university is talk about something beyond the visible buildings, people, activities, and their coordination. Just as (ii) is, in Ryle's view, self-evidently absurd, so is (i). In both cases, he seemed to think, talk of one kind (of the mental, or of the university) is, at bottom, nothing more than talk of things of another kind (of behavior, or of the buildings, people, and activities that make up the university)—even though there are no *definitions* or *analyses* $\ulcorner \forall x\ (K1x \Leftrightarrow K2x) \urcorner$ reducing the one kind of talk to the other. Thus, he mistakenly seems to have concluded, his ''version of behaviorism'' was uncontentious, and unworthy of the name.

I next turn to Weatherson's complaint that my attribution to Austin of an anti-skeptical goal in *Sense and Sensibilia* is unfounded, and his suggestion that Austin's aim was merely to advance direct realism about perception. It is true that Austin was a vigorous opponent of sense-data, and a harsh critic of attempts to analyze the perception relation we bear to ordinary objects in terms of any supposedly more basic relation involving appearances. However, to stop there is to miss his attack on

[2] This is all I meant in the chapter about Ryle's need to show that talk of the physical reduced to talk of the behavioral.

the privileged epistemic role of appearance statements, expressed by (1), in justifying knowledge attributions.

1. Knowledge of ordinary objects in our environment requires empirical evidence provided by statements about how things perceptually appear. These appearance statements constitute a maximally secure epistemic foundation encompassing crucial evidence for claims about the ordinary objects around us.

Although Austin recognizes special difficulties with versions of (1) that take appearance statements to be about sense data, he makes it clear that he regards (1) as a dangerously incorrect invitation to skepticism, no matter how appearance statements are construed. For this reason, he mounts two main arguments against it. One (discussed on pp. 174–178) attempts to show that certain statements—like the statement *It's a pig*, made looking at a pig directly in front of one—are known without one's having evidence for them at all. The other (discussed on pp. 187–192) tries to show that it is wrong to think of appearance statements as having greater epistemic security than that of the ordinary material-object statements they are supposed to justify.

> Austin notes the connection with skepticism near the end of *Sense and Sensibilia*, in discussing Warnock's linguistic version of Berkeley.
>
> Although Warnock insists that neither he nor Berkeley has any intention of casting doubt on the judgments we ordinarily make, of arguing for any brand of philosophical scepticism, this procedure of representing forms of words [ordinary material-object statements judged true on the basis of perception] as *in general* vulnerable is, of course, one of the major devices by which the sceptical theses have commonly been insinuated. ... What Warnock is really trying to do ... is to produce ... a *minimally adventurous* form of words, by use of which we can always stick our necks out as little as possible. And in the end he arrives at the formula, 'It seems to me now as if ...' ... But Warnock doesn't leave it at that; he goes on to say that statements about 'material things' are not *the same* as sets of statements about how things seem—the two kinds of statements are related as *verdicts* to *evidence*. ... But this comparison is really quite disastrous. It clearly involves falling in with a number of mistakes we mentioned earlier on [in the discussion of Ayer]—with the idea, for instance, that statements about 'material things' *as such* are always, have to be, based on evidence, and that there is a particular other kind of sentence [appearance statements] the business of which is to be evidence-providing. ... Warnock's comparison also leads directly to just the kind of 'scepticism' which he is officially anxious to disavow. ... To give a verdict on evidence is precisely to pronounce on some matter on which one is not a first-hand authority. So to say that statements about 'material objects' are in general like verdicts is to imply that we are never, that we can't be, in the best position to make them ... But to put the case in this way is to make it seem quite reasonable to suggest that we can never *know*, we can never be *certain*, of the truth of anything we say about 'material things'... But how absurd it is ... It is just this kind of comparison which does the real damage. (138–141)

As this passage makes clear, Austin's concern with perception was intimately linked to a concern with how it provides knowledge, and to his view that non-realist, appearance-based, treatments of perception open the door to disastrous and absurd forms of skepticism. It must be remembered that *Sense and Sensibilia* is an attack on two main works—Ayer's *The Foundations of Empirical Knowledge*, in which a sense-data analysis of perception is placed in service of a phenomenalist capitulation to skepticism, and Warnock's *Berkeley*, in which a sympathetic ordinary-language face is put on Berkeley's skeptical phenomenalism (Ayer, 1940; Warnock, 1953). To read *Sense and Sensibilia* oblivious to Austin's evident intent of demolishing the epistemically disastrous effects of misguided philosophies of perception is to fail to come to grips with one of its central themes.

Regarding Wittgenstein, Weatherson objects to the connection I draw between the deflationary conception of philosophy in the *Investigations* and the identification of necessity, apriority, and analyticity. I argue that Wittgenstein's deflationism is a consequence of (i) his view that the problems of philosophy are linguistic, and so are to be solved by linguistic analysis, (ii) his rejection of the Tractarian conception of analysis as involving hidden logical forms, and (iii) his social conception of meaning and deflationary conception of rule-following, understood in terms of socially conditioned agreement about the use of words. Point (i) is a holdover from the *Tractatus*, where the necessary, the apriori, and the linguistically, and logically, true are explicitly identified. The *Investigations* is similarly explicit that truths that are not empirical or scientific are apriori, and therefore linguistic in nature. (PI 109) Since, like other major philosophers of the period, he took the necessary and the apriori to be one, Wittgenstein did, as I maintain, identify necessity, apriority, and analyticity.[3]

It is true that in the *Investigations* Wittgenstein doesn't pay much attention to necessity, *per se*. Nor, so far as I know, does he argue directly from the necessity of philosophical claims to their apriority, and then to their analyticity. Thus, Weatherson seems to think, recognition of the Kripkean necessary aposteriori would not have affected his meta-philosophical position. I am not so sure. Although such a recognition may not have disrupted any direct argumentative route to (i), showing that necessity cannot be reduced to analyticity, would, I suspect, have encouraged the suspicion that apriority can't either—as it has for us—and in that way threatened (i). Even more important, there were already good grounds at the time Wittgenstein wrote to reject the linguistic conception of the apriori. As shown in chapter 12 of Volume 1, Quine's argument in "Truth by Convention" was a powerful (though underappreciated) attack on the idea that analyticity explains apriority (Quine, 1936). Weatherson emphasizes that some analytic philosophers *argued*, rather than simply assumed, that the way to explain the necessary apriori was to identify it with the analytic. That's true. But if the philosophers of the period were so open to argument on this point, some defender of the linguistic conception of the apriori should have cogently answered Quine. That none did suggests that the identification of necessity, apriority, and analyticity had, by the forties, attained the status of a dogma.

The relevance of Quine to my evaluation of Wittgenstein's linguistic conception of philosophy underscores another way in which Weatherson is off target. He notes the far-reaching importance I attach to the progress we have made in distinguishing

[3] In assessing the quotation of a Wittgensteinian remark about the synthetic apriori, I would emphasize his mention of *the analysis* of the concept of prime number. Analysis, understood in the old logical manner, was something that Wittgenstein gave up after the *Tractatus*. According to the later Wittgenstein, something might hold by virtue of meaning, even if its truth is not guaranteed by any old-style analysis (and so would be judged synthetic by the old standard).

logical, analytic, apriori, and necessary truth—and also my recognition of Kripke's contribution to this progress. However, he fails to note my recognition of the contributions of others. In Quine's case, these include his 1936-argument against the linguistic conception of the apriori, and his 1951-argument against the linguistic conception of the necessary. Wittgenstein, among others, ignored them at his peril.

Weatherson's underestimation of the significance of Wittgenstein's and Ryle's identification of the necessary and the apriori with the analytic, and his neglect of my discussion of Quine's contribution to the ultimate disentanglement of these three types of truth, leads him to complain that I exaggerate the importance of Kripke. However, he gives no examples of any such exaggeration apart from the problematic remarks he makes about Ryle and Wittgenstein that I have already dealt with. Instead, he tries to make his point with a string of pejoratives—*Whig history*, and Kripke as *hero* and *deus ex machina*—as well as with sarcasm—"Soames gives us no inkling of where theories of direct reference came from save from the brilliant mind of Kripke." What is one to say about the evident irritation animating these remarks?

To me, it appears to have led him to misrepresent the role Kripke plays in my history. A few factual observations may help. First, though Kripke certainly blazed the trail, he has never been a direct reference theorist in the sense (intended by Weatherson) that David Kaplan, Nathan Salmon, and I have been. Second, Weatherson's repeated implication that I uncritically give Kripke more than his due ignores my extensive criticism in chapters 15–17 of important aspects of his account of the necessary aposteriori and the contingent apriori. Third, as Weatherson well knows, but apparently can't accept, the aim of my already lengthy volumes made it impossible to trace the antecedents not just of Kripke's views on reference, but of many of the important ideas I discuss. Fourth, although Weatherson says that "there's no discussion of the possible connections between Wittgenstein's later theories and direct reference," precisely this point *is* discussed on pp. 18–22.

Finally, a general point. I wouldn't be doing philosophy if I didn't think that it progressed, and that as a result we know more now than we did a century ago. For that reason, I don't view the history of philosophy as a story of the irresolvable clash of defensible but irreconcilable views, in which all we can do is present the passing parade of who said what, and who influenced whom. Instead, I see it as an identification of what philosophical progress has been made, and an account of how it was made. For those who think this is "Whig history" I pose two questions. First, if you don't think that progress is made in philosophy, or that history should chronicle it, why should we be interested in the subject, or its history, at all? Second, if you agree that history is centrally concerned with real philosophical progress, but you think I have misidentified it, what have been the most important developments in the past century, and what shortcomings in the philosophy that preceded them does our more advanced knowledge allow us to spot?

Quine and Davidson: reply to Byrne

Byrne begins with my account of Quine's two main argumentative routes to the indeterminacy of translation. The first assumes that if translation is determinate, it is determined by purely behavioral truths; the second assumes that if translation is determinate, it is determined by all physical truths. Although Byrne agrees that neither route is successful, he doubts my contention that the second can be extracted from what Quine says. There seem to be two reasons for his doubt. First, he accepts my contention that when *determination* is made explicit and disambiguated, the physicalistic route fails for reasons that are arguably the same as those that would defeat a similar argument for the indeterminacy of just about any theory couched in ordinary language, as opposed to the language of physics. But surely, Byrne thinks, Quine's argument for the indeterminacy of translation should turn on something special about language. I agree, it should. But this tells against my interpretation of Quine only if what is obvious to Byrne and me was equally obvious to Quine. It wasn't, since he thought that there is an important sense in which claims about translation and meaning are not determined by physical truths, while non-intentional truths of other disciplines are. Second, Byrne questions whether there is textual support for my contention that Quine presented his physicalistic route to indeterminacy as a supposedly unanswerable challenge to non-behaviorists to specify physical facts, beyond stimulus meanings, that would determine translation. There is such support.

In *Words and Objections* Chomsky complains that Quine's indeterminacy thesis is simply the behaviorist application to the science of language of an unjustifiable restriction of the *facts* constituting its subject matter to observable events providing *evidence* for its claims (Chomsky, 1969). Since other sciences are not so restricted, Chomsky argues that there is no justification for imposing the restriction on linguistics. Quine's response is that there is no such imposition, since the indeterminacy of translation remains, even if the linguist is granted access to all physical facts.

> In respect of being underdetermined by all possible data, translational synonymy and theoretical physics are indeed alike. The totality of possible observations of nature, made and unmade, is compatible with physical theories that are incompatible with one another. Correspondingly the totality of possible observations of verbal behavior, made and unmade, is compatible with systems of analytical hypotheses of translation that are incompatible with one another. Thus far the parallel holds. ... Where then does the parallel fail?
>
> Essentially in this: theory in physics is an ultimate parameter....Though linguistics is of course a part of the theory of nature, the indeterminacy of translation is not just inherited as a special case of the underdetermination of our theory of nature. It is parallel but additional. Thus, adopt for now my fully realistic attitude toward electrons and muons and curved space-time, thus falling in with the current theory of the world despite knowing that it is in principle methodologically under-determined. Consider, from this realistic point of view, the totality of truths of nature, known and unknown, observable and unobservable, past and future. The point about indeterminacy of translation is that it withstands even all this truth, the whole truth about nature. This is what I mean by saying that, where indeterminacy of translation applies, there is no real question of right choice; there is no fact of the matter even to *within* the acknowledged under-determination of a theory of nature. (302–303)

This is Quine's second route to the indeterminacy of translation, according to which meaning and translation are not determined by physical truth. Since Quine's behaviorism told him that non-behavioral, physical truths were irrelevant anyway, allowing his opponents access to them didn't worry him. As an argumentative strategy, this move was appealing, since it didn't require defending explicitly behaviorist premises, or singling out linguistics (and psychology) for special treatment among the sciences. As time went on, fewer and fewer philosophers were satisfied with the behaviorist premise that since we learn language on the basis of behavior, the linguistic facts learned can't transcend behavior. The vulnerability of this premise was emphasized in an influential article by Michael Friedman, after which the argument from physicalism became the main focus of attention (Friedman, 1975 cited on p. 246). The problem with that argument is, of course, that Quine didn't make the notion of P determining Q explicit, and the most obvious analyses of it—Q being a necessary, apriori, or counterfactual consequence of P, Q being explained by P, and Q being a logical consequence of P together with true bridge principles connecting the vocabulary of Q to that of P—all either clearly fail, beg the question, or are too underspecified to yield a definite result.

In light of this failure, Byrne suggests that Quine's idea might find better expression in a third route to the indeterminacy of translation. The idea is that (i) translation can be determinate only if it is determined by how words are used, but (ii) how words are used doesn't determine translation. The burden of this route to indeterminacy is in explaining *how words are used*. Byrne notes both an overly narrow behaviorist interpretation and an overly broad semantic interpretation, on which the new route fails. He then suggests that a more plausible case might be made if the use of an expression is taken to include its role in silent contemplation. I doubt it. If the psychological states involved in specifying the use of an expression in contemplation include the broadly intentional states of believing, intending, and imagining, then Quinean indeterminacy will presumably be resolved by the attribution of content to those states, and (ii) will fail. On the other hand, if such contentful states are excluded from the determination base—and all that it is added to publicly observable verbal behavior are inaudible whispers, and silent syntactic apprehensions—then there is no reason to think that (i) and (ii) can jointly be established, short of equivocation on the determination relation of the sort that undermines the second route to indeterminacy.

Byrne is on firmer ground when he turns to the question of whether Quine is a radical eliminativist about ordinary meaning and reference, as well as attitudes like *asserting* and *believing*. In chapter 11, I use Quine's doctrines about what is, and what is not, determined by the physical truths, together with a simple formulation of his physicalism—*All genuine truths are determined by the physical truths*—to argue that he is such an eliminativist. However, this statement of physicalism is arguably too simple. The complications, mentioned in a footnote, are discussed at length in the final section of my ''Indeterminacy of Translation and the Inscrutability of Reference'' (Soames, 1999a) . There, I offer a maximally charitable way of interpreting Quine according to which he is not quite a complete eliminativist about our ordinary semantic and psychological notions, and since he isn't, his view is not defeated by the mere fact that he has asserted, believed, or argued for *something*.[4] In

[4] The interpretation (pp. 362–364) is based on a formulation of physicalism according to which '''Gavagai' refers to x'' comes out (as Byrne recommends) undefined for the value of 'x' when 'x' is assigned something (like a rabbit, rabbit-stage, etc.) in the indeterminacy range of the term. By contrast, the formula is *not true of* anything, like a telephone, outside that range—so 'gavagai' *doesn't* refer to telephones. However, '''Gavagai' refers either to rabbits or to .... or to undetached rabbit parts'' is *true* when ''rabbits or to ... or to undetached rabbit parts'' exhausts the range of indeterminacy, as is ''Gavagai refers to something,' even though all of its instances are either not true or undefined.

the end, however, this is no real victory—since, as I argue there, and as Byrne agrees in his comment, Quine turns out to be close enough to being a complete eliminativist that his own assertions, beliefs and arguments end up being self-undermining in a slightly different way.

In turning to Davidson, Byrne questions my interpretation of his argument against the possibility of alternative conceptual schemes. As I see it, the argument relies on a flawed conception of truth in which instances of *'S' is true iff S* provide us with our understanding of 'true', which is then extended to sentences of other languages by translating them into English. It follows on this approach that we can't make sense of the ascription of truth to untranslatable sentences. Since the intelligibility of the idea that there are meaningful sentences to translate depends on the intelligibility of ascribing truth conditions to them, Davidson concludes that the idea of an alternative conceptual scheme expressed in a language wholly untranslatable into ours makes no sense. Nor is it coherent to imagine that such a conceptual scheme might be expressed in a language partially translatable into ours, since, he argues, translation requires fundamental agreement—from which it follows that the extent to which a language is translatable into ours, the conceptual scheme of its speakers must agree with ours.

The chief errors in the argument are its reliance on a flawed conception of truth, and its exaggeration of the extent of agreement required by translation. Byrne admits this criticism is effective, if the notion of truth employed by Davidson is the kind of restricted, Tarskian notion I take it to be. However, Byrne doubts that it is, or needs to be, and suggests an alternative interpretation. The idea is to admit that our notion of truth allows us to make *sense* of the idea of meaningful sentences we can't translate, but to insist that we couldn't have *evidence* of their meaningfulness without being able to find out whether they were true, which requires being able to translate them. The conclusion is then, not that there couldn't *be* meaningful but untranslatable languages, but that we couldn't have *evidence* of such.

Unfortunately for Davidson, the argument remains unsuccessful, even on Byrne's interpretation. One can easily imagine the practical success of obviously intelligent symbol-using creatures providing evidence that many of their sentences must be true, even if we can't translate them, and so can't determine whether their conceptual scheme agrees with ours. In addition, as Byrne intimates, nothing in the argument successfully rules out a chain of partially differing conceptual schemes, expressed in languages partially translatable into one another, connecting substantially different, and untranslatable, schemes at either end.

As for the interpretive question, Byrne is right that more needs to be said about Davidson's transition from the argument against alternative conceptual schemes expressed in wholly untranslatable languages to the argument against alternative conceptual schemes expressed in partially untranslatable languages. This can be done by adding a new element to my interpretation. The argument Davidson actually gives against wholly untranslatable languages does confuse Tarski's notion of truth with the notion needed for theories of meaning. As I argue in chapter 13, he was still confusing Tarski's notion with the one he needed in ''Radical Interpretation,'' (Davidson, 1973), published the year before ''On the Very Nature of a Conceptual Scheme (Davidson, 1974)''[5]. Thus, the confusion I ascribe to him in my

[5] See Soames (1999b) pp 102–107 for a discussion of Davidson's confusion about Tarski's notion of truth in ''Truth and Meaning,'' originally published in (1967), reprinted in *Inquiries into Truth and Interpretation*.

interpretation of the latter was one from which he continued to suffer—as indicated by his suggestion that we have no notion of truth that is applicable to sentences we can't translate.

> And the criterion of a conceptual scheme different from our own now becomes: largely true but not translatable. The question whether this is a useful criterion is just the question *how well we understand the notion of truth, as applied to language, independent of the notion of translation*. The answer, I think, is that *we do not understand it independently at all*. (194 OVNCS, my emphasis)

This passage is immediately followed in the text by an invocation of Tarski that incorporates the flawed conception of truth identified in my interpretation.

The new element to be noted is that Davidson's infamous interpretivism—according to which agreement between the interpreters and the interpreted is *constitutive* of the correctness of ascriptions of meaning and belief—provides him with an independent reason for refusing to credit those who can't be interpreted with meaning or believing anything. For Davidson, ⌜x means or believes that S⌝ is true of an agent A only if it follows from the best holistic, agreement maximizing interpretation of A. In light of this, one would expect 'x means or believes something' to be true of A only if some claim ⌜x means or believes that S⌝ follows from such an interpretation. When A is completely uninterpretable, one draws a blank, making it natural for an interpretivist to hold not just that we don't have *evidence* that A means or believes anything, but that it makes no sense to suppose that A does.

Although Davidson doesn't give this argument in reaching his desired conclusion about the impossibility of wholly uninterpretable agents speaking untranslatable languages, he could have. However, he does give a version of it when arguing, in the last few pages, against alternative conceptual schemes expressed in languages that are only partially translatable. There, he maintains that fundamental agreement between interpreters and interpreted is constitutive of even partial translation, thereby ruling out the possibility of alternative conceptual schemes. At this point, the argument has, as Byrne rightly notes, shifted—for we now have a single interpretivist line that can be applied to both the wholly and partially untranslatable cases, without the worrisome consequence that we can't make sense of the idea of true but untranslatable sentences. My guess is that Davidson—who apparently wasn't worried about this consequence—was unclear about the connection between the two applications of this interpretivist line, and felt the need to bolster his case against wholly untranslatable agents by bringing in his flawed notion of truth. There was, in any event, good reason to worry about what I am here calling "the single interpretivist line." When that line is taken, the argument *assumes* that what it is for an agent to mean or believe anything at all is for us to be able to interpret what he means or believes by using a theory the correctness of which is *defined* in terms of agreeing with us about most things. Not only is this assumption questionable, but the use of it to draw Davidson's conclusion about the impossibility of alternative conceptual schemes might well be deemed to beg the question. Be that as it may, the important interpretive point is that the faulty conclusion he reaches about alternative conceptual schemes is not just that we couldn't have evidence for them, but that the very idea that they are possible is conceptually confused.

Kripke: reply to Yablo

Like Weatherson, Yablo finds me too critical of my target, and takes my criticism to stem from elements foreign to the philosophy I criticize. Unlike Weatherson, who faults me for imposing Kripkean standards on ordinary language philosophers, Yablo faults me for imposing "high semantic theory" on the untheoretical Kripke. Which is it, I wonder—that I uncritically treat Kripke as the god to which all analytic philosophy ascended, or that I unjustly criticize him by imposing a semantic framework that minimizes his most important achievements? No doubt the more ingenious of my critics will find a way of answering, "Both!" However, the truth is simpler. My history aims to identify enduring achievements of the 20th century, and to disentangle them from whatever flaws, errors, and obscurities may have accompanied their discovery. Important advances are often made by philosophers who have caught a glimpse of the truth about something big, but who, understandably, don't yet have it under complete control. As a result, they are not always fully faithful to their own deepest insights. In my opinion, this was true of Frege in some of his remarks about sense and reference, of Russell in some of his remarks about descriptions, of Tarski in some of his comments about truth, of Carnap and Quine in some of their views of the modalities, and of Kripke in some of his discussions of the necessary aposteriori and the contingent apriori. Because of this, insights of an essentially Kripkean nature can sometimes be used to correct not only the missteps of earlier philosophers, but also those of Kripke himself.

A good place to begin is with Kripke's observation about what one means when one uses a name the referent of which is fixed by a description.

> So suppose we say, 'Aristotle is the greatest man who studied with Plato'. ...If ... we merely use the description to *fix the referent* then that man will be the referent of 'Aristotle' in all possible worlds. The only use of the description will have been to pick out to which man we mean to refer. But then when we say counterfactually 'suppose Aristotle had never gone into philosophy at all', we need not mean 'suppose a man who studied with Plato, and taught Alexander the Great, and wrote this and that, and so on, had never gone into philosophy at all', which might seem like a contradiction. **We need only mean, 'suppose that *that man* had never gone into philosophy at all**'. (57, *Naming and Necessity*, my emphasis)

Here, Kripke suggests that when the referent of n is fixed by a description denoting a man m, what one means when one utters (2a) is what would be expressed by (2b).

2a. (Suppose that) n is F.
b. (Suppose that) *that man* is F(said demonstrating m).

Assuming this, one needs neither Millianism nor "high semantic theory" to conclude that (3a) should be true only if (3b) is true, which, in turn, requires the truth of the *de re* ascription (3c).

3a. Jones supposes/means/says/knows that n is F.
b. Jones supposes/means/says/knows that *that man* is F. (said demonstrating m)
c. Jones supposes/means/says/knows that x is F
(relative to an assignment of m to $'x'$)

Since the same holds for names the referents of which are not descriptively fixed, all Kripke's putative examples of the necessary aposteriori and the contingent apriori involving names are instances of *de re* knowledge—as are those involving natural kind terms, where the knowledge is of kinds (which is more complicated). Since Yablo doesn't discuss such terms, I too will put them aside.[6]

Yablo ignores this explanation of the Kripkean *de re*, given on pp. 400–402, where it is couched in a modest framework of propositions—which he regards as foreign to Kripke. However, the framework is only a convenience, since, as we have just seen, the point can be made without mentioning propositions. Moreover, the framework is one that Kripke should find congenial. Its assumptions are: (i) that some things are asserted, believed, and known, (ii) that these things may be contingently or necessarily true, (iii) that they are designated by (uses of) clauses like ⌈(the statement) that S⌉, and (iv) that ascriptions ⌈x asserts/believes/knows that S⌉ report that one asserts/believes /knows one of these things. Since it is central to Kripke's position that certain things that are knowable only aposteriori are necessary truths, he is committed to (i) and (ii). Since he concludes that "the statement that Hesperus is Phosphorus" is one of these things from the claim that "one cannot know apriori that Hesperus is Phosphorus" despite it's being a necessary truth, he is also committed to (iii) and (iv). As for the identity of "propositions," it is plausible to assume, as I did, that they are not identical with sentences that express them, but rather are what these sentences have in common. However, nothing in my critique depends on this. One can, if one likes, take propositions to be sentences, or utterances, provided one recognizes that ⌈x knows/believes that S⌉ as used in C may be true of an agent who has never heard of S, if the agent (justifiably) accepts some (true) sentence which in the agent's context has the same content that S has in C.

With this in mind, consider whether Pierre, who believes that Londres is pretty, thereby believes that London is pretty? (I here use a made-up dialect of English in which 'Londres' has been imported, with the same meaning it has in French.) As Kripke has taught us, this is a vexed question. A case can be made that Pierre does believe this, since 'Londres' and 'London' seem to have the same content. However, this result is counterintuitive, so one had best not jump to conclusions. The case for caution carries over to whether Pierre, who we may assume knows apriori that London is London, thereby knows apriori that London is Londres. But, surely, if there is reason for being cautious about this question, there is reason to be cautious about whether or not our knowledge that Hesperus is Phosphorus, or that Cicero was Tully, is apriori.

This is the light in which to view Yablo's remark, "It is unclear why Soames thinks Kripke had to *show* that Hesperus's identity with Phosphorus was aposteriori; who would ever have doubted it?" Who would have doubted it? Kripke, for one—as indicated by the uncertainty expressed footnotes 43 and 44 of "A Puzzle About Belief." (Kripke, 1979.) Even in *Naming and Necessity*, he thought it unobvious enough to spend the last two pages of lecture 2 explicitly arguing that it is not knowable apriori that Hesperus is Phosphorus. The problem is that his argument (on pp. 103–104) uncritically jumps from (i) to (ii).

[6] For discussion, Soames (2005b).

(i) Prior to having evidence supporting the astronomical discovery, we understood, but didn't accept, and wouldn't have been justified in accepting, *the sentence* 'Hesperus is Phosphorus'.
(ii) Therefore, we didn't then believe, and wouldn't have been justified in believing, that Hesperus was Phosphorus. Since knowledge that Hesperus is Phosphorus requires empirical justification, it isn't apriori.

To justify this move, Kripke needs a premise allowing us to conclude ⌜We didn't believe, and wouldn't have been justified in believing, that S⌝ from the observation that we understood but didn't accept, and wouldn't have been justified in accepting, S. In short, his argument requires a principle of strong disquotation, or some functional equivalent.

Although such principles are false, they are not, as Yablo flippantly suggests, "insane." On the contrary, they are based on insights which, when properly formulated, tell us something important about the origins of *de re* attitudes. However, when so formulated, they fail to support Kripke's argument that it is not knowable apriori that Hesperus is Phosphorus. This is important, since in lecture 3 he generalizes that argument to all cases of the necessary aposteriori. Fortunately, this is not his only route to that destination. In addition, there is a sound, alternative route—based on apriori knowledge that certain properties (which can be known to be possessed by objects only aposteriori) are essential properties of anything that has them—which is sufficient to generate all of Kripke's putative examples of the necessary aposteriori, except for a small subclass of cases consisting of identity statements involving only *simple* names or natural kind terms (Soames, forthcoming, a).

Does this mean that it *is* knowable apriori that Hesperus is Phosphorus after all? In my opinion, the question is not univocal. Sometimes the sentence 'Hesperus is Phosphorus' may be used simply to assert, or report an agent's attitude toward, the proposition p it semantically expresses—which is the necessary, apriori truth also expressed by 'Hesperus is Hesperus'. However, on many other occasions, it is used to assert, or report an agent's attitude toward, a pragmatic enrichment q of p—where q is often (though not always) contingent, while being knowable only aposteriori (Soames, 2002, 2005a, forthcoming, b). These conclusions do indeed rest on "high semantic (and pragmatic) theory" that should neither be read into Kripke, nor presupposed in criticizing him. However, my chapters on Kripke are not guilty of these errors. Despite Yablo's repeated claim that I use the contention that it *is* knowable apriori that Hesperus is Phosphorus to criticize Kripke, in fact I do not. Instead, I raise doubts about the premises tacitly invoked in his argument.

> We need not here try to confirm or resolve those doubts, or to decide precisely what to say about potentially problematic cases ... It is enough to have shown that the premises needed, and tacitly used, in Kripke's argument are insecure, and cannot, without further investigation, be taken for granted in establishing his conclusion that it is not knowable apriori that Hesperus is Phosphorus. (393)

Suppose, however, one takes another tack. Can't one simply treat it as a datum that one may know that Hesperus is Hesperus, without knowing that Hesperus is Phosphorus—thereby removing the main obstacle to viewing the latter as both necessary and aposteriori? Yes one can, but only if one is prepared to explain how substitution of one name for the other changes the proposition semantically expressed by a sentence, or, equivalently, how substituting ‘Phosphorus’ for ‘Hesperus’ in

4. A knows that...Hesperus...

changes the sentences the understanding and justified acceptance of which is sufficient for satisfying the knowledge ascription. The dilemma is between embracing this theoretical option and adopting the alternative view that substitution of coreferential names preserves the proposition semantically expressed, while explaining how speakers use such substitutionally related sentences to assert and convey different information, and to express and report different beliefs. As I stress in my final assessment of the issue (393–395), both horns of the dilemma face daunting difficulties. The criticism of Kripke is not that he hasn’t resolved them, but that he has implicitly staked one of his two routes to the necessary aposteriori on a questionable theoretical option the correctness of which has not been assured. Fortunately, his other, essentialist, route remains unscathed, leaving only a small subset of cases on which one may wish to suspend judgment.

As for the contingent apriori, we have seen that whenever one can truly say

5a. A knows that if there is such a thing as D, then n is D

in a case in which the reference of n is fixed by D, the knowledge reported is or involves, on Kripke’s own account, de re knowledge of the object o that D denotes (expressed by (5b,c)).

5b. A knows that if there is such a thing as D then *that object* is D.
(said demonstrating o)
c. A knows that if there is such a thing as D, then x is D}
(relative to an assignment of o to ′x′)

Contrary to Yablo, there simply is no ordinary, Moorean, sense in which such *de re* knowledge is apriori. Whenever we know of a certain length l that if there is such a thing as stick S (and hence such a thing as its length), then l is the length of stick S, or of a certain man m that if there is some one person who invented the zipper, then m invented the zipper, or of anything else that it has a relevantly similar property, it is because we have empirical evidence justifying what we know. This evidence, which is not required for the truth of corresponding report

5d. A knows that if there is such a thing as D, then D is D,

cannot be created by simply engaging in a linguistic ceremony in which we stipulate the meanings or referents of words. Hence, all Kripke’s putative examples of the contingent apriori fail.[7] However, genuine examples can be generated by substituting

[7] My rejection of these examples does not, contrary to Yablo, involve rejection of Weak Disquotation (and Justification). Instead, I reject both the claim that one can understand a name the reference of which is fixed by a description without knowing of its denotation that it is so denoted, and Weak Linguisticism about the Apriori. (414–416).

actually-rigidified descriptions for his uses of names, thereby vindicating his guiding insight that there are contingent apriori truths.

Yablo disagrees with this story. In its place, he offers RLA as a truth that allows one to save Kripke's examples of the contingent apriori, and block an objection to the aposteriority of the claim that Hesperus is Phosphorus, while remaining true to everything in *Naming and Necessity*.[8]

RLA. If i′s understanding of S = i′s knowledge of S′s meaning = i′s knowledge of which proposition p it is that S expresses, provides i with a guarantee that p is true, then ⌈x knows apriori that S⌉ is true of i.

I object. For one thing, RLA wrongly counts ⌈x knows apriori that S⌉ to be true of an agent A in cases in which empirical evidence is required to justify A's knowledge, provided that having this evidence is part of what is involved in A's understanding S. Although sentences and propositions that have this property are epistemologically interesting, they are *not* genuine instances of the apriori.[9] For another thing, RLA, as understood by Yablo, leads to the result that (6a) is true, even though (6b) is false.

6a. We know apriori that one meter is the length of stick S (if S exists).
b. We know apriori that *that length* is the length of stick S (if S exists). (said demonstrating the length l of stick S)

This hard to square with Kripke's view—which is part of the data that Yablo's account is designed to preserve—that what we *mean* by (7a) and (8a) is what is meant by (7b) and (8b).

7a. Suppose that one meter is the length of stick S (if S exists).
b. Suppose that *that length* is the length of stick S (if S exists). (demonstrating l)
8a. One meter is the length of stick S (if S exists).
b. *That length* is the length of stick S (if S exists). (demonstrating l)

Surely, if this is what we mean, then the truth of (9b) is a necessary condition for the truth of (9a).

9a. We mean/suppose that one meter is the length of stick S (if S exists).
b. We mean/suppose that *that length* is the length of stick S (if S exists). (demonstrating l)

Given this, one is hard pressed to deny that the same relationship holds in (10) and (11).

8 Although the truth of RLA wouldn't establish that our knowledge that *Hesperus is Phosphorus* is aposteriori, it would, Yablo thinks, render that that claim compatible with the apriority of our knowledge of *Hesperus is Hesperus*.

9 See *Reference and Description*, pp. 55–57, 66–67; also *The Age of Meaning*, 414–417.

10a. We say/assert/believe that one meter is the length of stick S (if S exists).
b. We say/assert/believe that *that length* is the length of stick S (if S exists). (demonstrating l)

11a. We know that one meter is the length of stick S.
b. We know that *that length* is the length of stick S (demonstrating l)

But now Yablo is in trouble. Since he admits that possession of empirical evidence supporting (8b) is required for the truth of (11b), it follows that possession of this evidence is also necessary for the truth of (11a). Since he further grants that (8a) and (8b) express the same proposition, he has to admit that in order for us to have knowledge that one meter is the length of stick S we must possess empirical evidence supporting the proposition that one meter is the length of stick S. Surely, this is incompatible with the claim—which he purports to derive from RLA—that we know apriori that one meter is the length of stick S. Thus, either RLA is false, or it is insufficient to get the Kripkean result he desires.

The only hope of blocking this argument is to deny that (11b) is a consequence of (11a). If one wishes to do this while remaining true to Kripke's *de re* account of what we mean by (7a) and (8a)—and with Yablo's admission that (8a) and (8b) express the same proposition—one must come up with a plausible answer to (i) or (ii) (or near variants).

(i) How can (11b) fail to be a consequence of (11a), when (9b) and (10b) are consequences of (9a) and (10a)?
(ii) How can (b) fail to be a consequence of (a) in each of (9–11), when (7b) and (8b) are what we mean by (7a) and (8a) (and (8a) expresses the same proposition as (8b))?

Needless to say, Yablo doesn't answer these questions. Since, in my opinion, they are unanswerable, I don't see how his attempt to preserve the full complement of Kripkean views can succeed.[10]

If, despite all this, one still hopes to save all Kripkean examples of the contingent apriori and the necessary aposteriori, one would, I think, do better not to fiddle with parochial accounts of apriori knowledge ascriptions, but to try to find an alternative to Kripke's *de re* treatment of what is meant by utterances containing names, including those the referents of which are fixed by description. Since I myself know of no such alternative that is not subject to crushing objections, this is not a strategy I favor. Be that as it may, if I am right about the difficulties discussed here, there is no

[10] Another way to put the same point involves the relationship between (11a) and (11c)—*We know that x is the length of stick S (if S exists)*, taken relative to an assignment of l to 'x'. Whereas one might be tempted to think that the truth conditions of ⌈A knows that α is F⌉ are sensitive not only to the proposition expressed by its complement, but also to the identity of the closed singular term α, one would not similarly be tempted to think that the truth conditions of ⌈A knows that x is F⌉ (relative to an assignment of o to 'x') are sensitive to the identity of the variable 'x' (over and above the proposition expressed by the complement). Since Yablo seems prepared to grant that the complement of (11c) expresses (relative to an assignment of l to 'x') the same proposition expressed by the complement of (11a), he has no basis for denying that that the truth of (11a) guarantees the truth of (11c). However, he also, quite rightly, seems to admit that the truth of (11c) requires us to have empirical evidence justifying the proposition that l is the length of stick S—from which we get the same *reductio* of his position as before.

plausible way of being completely true to Kripke, while getting all of his examples to come out precisely as he characterized them. Hence, if you don't like my relatively modest departures from Kripkean orthodoxy, you will have to find your own heresies to embrace.

References

Ayer, A. J. (1940). *The foundations of empirical knowledge*. New York: Macmillan.

Chomsky, N. (1969). Quine's empirical assumptions. In D. Davidson, & J. Hintikka (Eds.), *Words and objections*. Dordrecht: Reidel.

Davidson, D. (1973). *Dialectica, 27*, 313–328 [reprinted in Davidson (2001)]

Davidson, D. (1974). *Proceedings and addresses of the American Philosophical Association, 47* [reprinted in Davidson (2001)]

Davidson, D. (2001). *Inquiries into truth and interpretation.* Oxford: Clarendon Press

Friedman, M. (1975). Physicalism and the indeterminacy of translation. *Noûs, 9*, 353–373.

Kripke, S. (1979). A puzzle about belief. In A. Margalit (Ed.), *Meaning and use*. Dordrecht: Reidel.

Quine, W.V. (1936). Truth by convention. In O. H. Lee (Ed.) *Philosophical Essays for A.N. Whitehead*. Longmans, New York [reprinted in Quine, The ways of paradox. Random House, New York].

Soames, S. (1999a). *Canadian Journal of Philosophy, 29*, 321–370.

Soames, S. (1999b). *Understanding truth*. New York: Oxford.

Soames, S. (forthcoming, a). Kripke on epistemic and metaphysical possibility: Two routes to the necessary aposteriori. In A. Berger (Ed.), *Saul Kripke*. Cambridge: Cambridge University Press, available at usc.edu/~soames.

Soames, S. (forthcoming, b). The gap between meaning and assertion, available at usc.edu/~soames.

Soames, S. (2002). *Beyond rigidity*. New York: Oxford University Press.

Soames, S. (2005a). Naming and asserting. In Z. Szabo (Ed.), *Semantics vs. pragmatics* (pp. 356–382). Oxford: Oxford University Press.

Soames, S. (2005b). *Reference and description* (pp. 58–68). Princeton: Princeton University Press.

Warnock, G. J. (1953). *Berkeley*. New York: Pelican Books.

.