QUINE'S NATURALISM AND BEHAVIORISMS

TONY CHENG

Abstract: This paper investigates the complicated relations between various versions of naturalism, behaviorism, and mentalism within the framework of W. V. O. Quine's thinking. It begins with Roger Gibson's reconstruction of Quine's behaviorisms and argues that it lacks a crucial ontological element and misconstrues the relation between philosophy and science. After getting clear of Quine's naturalism, the paper distinguishes between evidential, methodological, and ontological behaviorisms. The evidential and methodological versions are often conflated, but they need to be clearly distinguished in order to see whether Quine's argument against mentalism is cogent. The paper argues that Quine's naturalism supports only the weakest version of behaviorism, that is, the evidential one, but this version is compatible with mentalistic semantics. Quine's opposition to mentalism is an overreaction against the behaviorist camp. By contrast, Jerry Fodor's objection to José Luis Bermúdez is an overreaction from the opposite direction.

Keywords: behaviorism, language of thought, meaning, mentalism, naturalism, Quine.

1. Naturalism, and Quine's Version of It

"Naturalism" is a familiar but ambiguous term in contemporary philosophy. To this one might say that almost every doctrine in philosophy is championed by multiple philosophers; as a result, all the doctrines have plenty of versions. It is not surprising, therefore, to learn that "naturalism" is another ambiguous philosophical doctrine. In my view, the distinguishing characteristic of this doctrine is that when someone regards herself as a naturalist, others cannot confidently attribute any specific thesis to her without making heavy assumptions. One might treat natural science as the only avenue to truth, or one might emphasize the importance of embedding oneself into nature, or perhaps one is simply against anything supernatural, to mention a few approaches.[1]

---

[1] One can consistently hold that there is nothing supernatural and that natural science is *not* the only way to truth, for "there is nothing supernatural" is ontological and "natural science is not the only way to truth" is epistemological. This is one of the views I argue for in this paper. The crux of my argument rests on the relation between *a priori* philosophy and natural science.

The first view is about the way we pursue knowledge, the second about the way we live, and the third about what there is. We need to assume one of these divergent interests in order to conjecture which view one actually holds. Compare this with physicalism. Physicalism is manifestly an ontological thesis in the literature, no matter how one construes it. The differences between versions of physicalism are about detailed formulations. When someone claims he is a physicalist, what he advocates is relatively clear.

Quine is one of the famous naturalists of the twentieth century. To avoid the ambiguity of "naturalism" sketched above, I shall formulate his version carefully below. Roger F. Gibson Jr.'s reconstruction (1995) of Quine's naturalism is often cited, so I begin with it:

(1) Quine "rejects the traditional quest for a first philosophy, i.e. the quest for a ground somehow outside of science upon which science can be justified," and
(2) Quine "accepts science as the final arbiter concerning questions of what there is." (Gibson 1995, 426)

"First philosophy" in the present context means "*a priori* or experiential ground outside of science upon which science can either be justified or rationally reconstructed" (Gibson 2004, 181). The formulations above are *epistemological* in nature: they remark on the roles of *a priori* philosophy and natural science, two putative ways of pursuing *knowledge*. But, I submit, Quine's naturalism also contains an *ontological* tenet, namely, that (3) everything that exists is natural in character. Quine writes: "With Dewey I hold that knowledge, mind, and meaning are part of the same world that they have to do with, and that they are to be studied in the same empirical spirit that animates natural science" (1969b, 26). The first half of the sentence is the ontological claim I just mentioned, and the second half seems to be Gibson's point (2).[2] My primary aims in this article are twofold. First, I argue that Gibson's formulation of (2) is misleading.[3] Second, I criticize Quine's major argument against mentalism concerning meaning. Generally speaking, mentalism is the thesis that mental items, whatever their natures, are the primary locus of meaning (more on this in section 3). In this argument Quine bases his behaviorism on naturalism and thinks he can rebut mentalistic semantics with this behaviorism. I shall distinguish

---

[2] Here Quine only mentions knowledge, mind, and meaning, but his naturalism is more thoroughgoing than this.
[3] I do not claim that Gibson misunderstands Quine at this point. My contention is rather that although Gibson might himself perfectly understand this thesis, the way he formulates it seems to overemphasize the role of natural science. Before we enter section 3 below, some modifications of the relevant formulations are needed.

three kinds of behaviorism and argue that the only version supported by naturalism is too weak to reject mentalism. The other two versions are indeed inconsistent with mentalism, but they gain no support from naturalism, at least Quine's version of it.

2. Philosophy and Natural Science

My target in this section is Gibson's point (2) above: namely, "Quine accepts science as the final arbiter concerning questions of what there is." This statement amounts to saying that we should tackle the questions of ontology only through science, fundamentally speaking. More recently, Gibson (2004) reformulates this tenet of Quine's naturalism. He writes: "It is up to science to inform us about what exists" (2004, 181).[4] The reason for this, he continues, is that "all is not lost with the passing of first philosophy, for science remains" (182). Here he seems to assume, though unwittingly, a dichotomy between science and *a priori* philosophy. Because, according to Quine, *a priori* philosophy has gone by the board, and there is a dichotomy between natural science and *a priori* philosophy, all we can have in studying what there is, the thinking goes, is natural science. I think this formulation does *not* capture Quine's thought accurately. For one thing, if this really were what Quine had in mind in arguing for naturalism, this doctrine would face obvious counterarguments; for another, this formulation does not fit well with some of Quine's remarks (see below).

I want to avoid an uncharitable objection to the dichotomy in advance: one might find this dichotomy obviously false, for it seems to neglect mathematics, logic, social sciences, and the like. This objection is not difficult to answer. Quine's naturalism does not rigidly define which subjects are sciences, which are not. A subject can be tolerated by Quine's naturalism as long as it respects empirical data, for the thrust of this naturalism aims at so-called *a priori* pure conceptual analysis. Indeed, there is a substantial debate about the nature of social science, but Quine's naturalism does not take any specific position in this debate. Furthermore, mathematics and logic can also be encompassed by his naturalism, for Quine accepts pragmatism too, and in practice science cannot do without them. One can challenge Quine's pragmatism, of course, but at least within Quine's system of thought mathematics and the like can be explained by it. The dichotomy has prima facie plausibility because there is a traditional metaphilosophical

---

[4] The complete sentence is this: "It is up to science to inform us about what exists and how we come to know what exists." I omit the second half of the sentence because epistemology is not my focus in this article.

thought that philosophy is distinct from the rest of our inquiries of the world due to the fact that it consists in *a priori* conceptual analysis. Gibson's dichotomy seems to instantiate this thought implicitly: philosophy is *a priori*, and science is *a posteriori*; the former has gone by the board, so it is up to science to inform us about what exists.

My first counterargument directed to (2) is as follows. I grant, at least in this paper, that there is no *a priori* philosophy, no Archimedean point outside the whole inquiry. We have to, in Quine's own words, *work from within* our ongoing theories. It does not follow, however, that all we can have is natural science. First of all, some statements are remote from the interface between the whole theory and the tribunal of experience. Even if confirmation holism is true, it follows only that these sorts of statements, notably logic and mathematics, are experience loaded; it does *not* mean that we can devise experiments for them. Confirmation holism is a theoretical point; it tells us that although logic and mathematics do not wear empirical contents on their sleeves, they *do* have such contents. Nevertheless, to have empirical bearings is one thing, to be encompassed by natural science is another. Logic and mathematics simply have too little in the way of empirical contents to be studied by natural science. The same holds for many philosophical/ontological theses. For example, we cannot conduct any experiment for the existence of the external world, since every experiment involves our observations, and skeptics can always say that we cannot vindicate that those experiments do not consist in our own sense impressions, ideas, and the like. In such contexts, any intricate experiment is as useless as G. E. Moore's arms. We need logic, mathematics, and philosophy, though *a posteriori* ones.

My second argument against (2) is that though in some philosophical debates we do have ideas to design experiments for them, there still are reasons they cannot be studied by natural science. Consider Derek Parfit's teletransporter (1986). We do not have such technology today, but even if we invent it in the future, we still cannot carry out the experiment and thereby reach significant conclusions concerning the problem of personal identity, since the experiment involves people's lives. Or consider Frank Jackson's Mary (1982). Again, it is arguable that the experiment is technically impossible. But even if we overcome all the technical difficulties in the future, we still cannot carry out that experiment, for the simple reason of humanity. Moreover, experiments involving clairvoyance and other supernatural powers, as in the context of epistemological internalism/externalism, are probably empirically impossible. That's why we need *thought* experiments. I conclude that (2) is false because there are clear examples in which science cannot solve the problems on its own: we cannot envisage experiments for those philosophical statements that are remote from the tribunal of experience, and we cannot carry out experiments for those

philosophical theses that involve empirical, technological, or moral impossibilities.[5] *A posteriori* philosophy is required.

We have seen that (2) faces some obvious objections. That means at least one of its premises ought to be rejected. The premise at fault, I submit, is the dichotomy between natural science and *a priori* philosophy. Compare the situation with mathematics. If confirmation holism is true, it follows that mathematics is in fact *a posteriori*, but it does not follow that we should renounce mathematics altogether. Similarly, it follows from confirmation holism that there is no *a priori* philosophy, but this does not mean that philosophy has gone by the board altogether. Why shouldn't we engage in *a posteriori* philosophy? When it comes to ontology, there is no reason to confine ourselves to natural science. As in the quote from Quine above, when it comes to the ontology of meaning he holds that meaning is "to be studied in the same empirical spirit that animates natural science." This is *not* to say that natural science is the only avenue to study meaning; rather, it urges that we should bear "empirical spirit" in studying meaning, which amounts to saying that the study of meaning is *a posteriori* in nature.

There is other evidence that Quine does not accept the dichotomy in question; for example, he says: "I see philosophy not as an *a priori* propaedeutic or groundwork for science, but as continuous with science. I see philosophy and science in the same boat" (1969a, 126). "Philosophy is continuous with science" means that philosophy is not a purely conceptual work; it is *a posteriori* in character. Nevertheless, it does not follow that science is the final arbiter concerning questions of what there is. Quine's naturalism does contain an element of natural science, but this element does not identify ontology with science. Rather, it urges the view that philosophy and natural science are in the same boat. Thus, we can restate three central tenets of Quine's naturalism as follows:

(A) There is no *a priori* first philosophy (as the foundation of natural science).
(B) Philosophy is continuous with natural science; they both contribute to ontology.
(C) Everything that exists is natural in character.

If we define naturalism partly in terms of the claim that everything is natural, it raises the question of what the criterion for naturalness is. Quine's criterion would seem to be that something is natural if its existence is posited as an element of our best theory of the world.

---

[5] One might think that we can overcome the concern about morality. After all, we can still carry out the experiments if we are immoral. Technical and empirical impossibilities, however, cannot be avoided this way.

This is not, to be sure, the only reasonable criterion one can be sympathetic with. John McDowell (1996), for example, has argued for a more relaxed conception of naturalness, based on the Aristotelian notion of second nature. Since the present paper investigates internal questions concerning Quine's system only, I shall not go into the details of McDowell's relevant thinking.[6]

With these theses in hand, we are ready to see the relations between naturalism, behaviorism, and mentalism. Before moving on, a clarification concerning Gibson's interpretation and mine is called for. Gibson characterizes Quine's naturalism as a set of epistemological commitments, while I have argued that for Quine naturalism encompasses ontological commitments as well. Some readers might think that there is no genuine conflict between the two, since proper ontology, according to Quine, falls outside our best total theory of the world, and since epistemology, in turn, is best in virtue of adhering to the best epistemological criteria, it is prior to ontology. I agree that there is *no* strict contradiction between Gibson's remarks and mine. It is just that according to Gibson's actual formulations the role of ontology is lost. What I would like to stress is that although it is true that for Quine epistemology is prior to ontology in the relevant sense, still, the significant place for ontology should be properly acknowledged. It should also be noted that Gibson never says explicitly that there is absolutely no role at all for philosophy in our investigations of what there is. This vicious implication should, however, be clearly avoided, which is what I have done above.[7]

3. Naturalism, Behaviorism, and Mentalism

In this section my focus is on Quine's naturalism about meaning. First we need to get clearer about Quine's major target, mentalism. Broadly speaking, it includes any theory of meaning that makes reference to mental phenomena as irreducible, fundamental elements of the theory. It considers mental items the primary locus of meaning. Sometimes a stronger further claim is made: namely, that private language is possible. This further commitment should be clearly separated from the core of mentalism.[8]

Now, after declaring what his naturalistic position is in "Ontological Relativity," as quoted in section 1 above, Quine proceeds as follows: "Meanings, therefore, those very models of *mental* entities, end up as

---

[6] I thank two reviewers for urging me to make these ideas explicit. The same goes for other thinkers' versions of naturalism, such as those of John Dewey (1988 [1929]), John Herman Randall (1958), Justus Buchler (1955), and many others.

[7] I thank a reviewer for pressing me on this point.

[8] I thank a reviewer for reminding me of this.

grist for the *behaviorist's* mill. Dewey was explicit on the point: 'Meaning ... is not a psychic existence; it is primarily a property of behavior.... Uncritical semantics is the myth of a museum in which the exhibits are meanings and the words are labels'" (1969b, 26–27, my italics). The inference is this: given naturalism, we should adopt behaviorism and thereby reject mentalism in the study of meaning. Raffaella de Rosa and Ernest Lepore spell out this line of thought as follows: "Quine is well known for his naturalism.... For meaning to be investigated empirically, it must be made public. Meaning becomes a property of behavior.... On Quine's view, the requirement that meaning is essentially public and social in nature relates to the identification of meaning with evidence, that is, to some sort of verificationism, which in Quine's case explicitly takes the form of behaviorism" (de Rosa and Lepore 2004, 68). From naturalism one has the publicity of meaning, and this, the thought goes, leads to Quine's behaviorism-cum-verificationism. Against this I urge that naturalism entails only a minimal version of behaviorism, and, *pace* Quine, this behaviorism is by no means verificationism of any sort. As a result, the bridge connecting naturalism and anti-mentalism collapses.

This discussion has some significance because most of Quine's commentators, as far as I know, do not challenge this line of reasoning. For example, Gibson writes in agreement that "once language is understood in this naturalistic way, as a social art to be studied empirically, it is *immediately obvious* that there cannot be any useful sense to the claims that language is private or that meaning is private" (1982, 64, my italics). In his introductory book on Quine, Alex Orenstein remarks: "[Quine's] argument is in part that of a behaviourist, that private ideas are 'pointless and pernicious' in the scientific study of language, and that we should dispense with them in favour of publicly observable linguistic behaviour" (2002, 122). De Rosa and Lepore's article, from which I quoted above, is another instance. Some might think that by "claims that language is private" Gibson means mentalistic semantics, for after these remarks he immediately refers to the museum myth criticized by Quine. Some problems may arise with this identification by Gibson, because a mentalist does not necessarily allow a private language, as indicated above. Fortunately, this complication bears no significance in the present context. Similarly, by "private ideas" Orenstein refers to mentalism; he chooses this narrow locution presumably because he is targeting old empiricists. It should be clear that a mentalist need not commit private ideas or images in his picture; John Searle (1983) and Jerry Fodor (1987) are good examples.

Let me at this point introduce some varieties of behaviorism (which has plenty of versions). In what follows I sketch only three relevant versions and try to locate Quine's view(s) appropriately. All the versions of behaviorism I formulate below are about semantics. They are:

(a) *Evidential* behaviorism: in constructing and testing theories of meaning, we can *use* only observable behaviors *as evidence*.
(b) *Methodological* behaviorism: in constructing and testing theories of meaning, we can *make reference* only to observable behaviors.[9]
(c) *Ontological* behaviorism: in constructing and testing theories of meaning, we can *make reference* only to observable behaviors, because *there is no* unobservable (mental) entity to be referred.

Version (a) is weaker than (b), and (b) is weaker than (c). Each in the sequence implies those previous to it, and adds something to them. Quine sometimes embraces (c), but at other times he limits his claims to one or another of the weaker versions, to the exclusion of the stronger ones. The distinction between (b) and (c) is relatively clear: they are identical except that (c) adds an ontological reason why one should not make reference to those entities. By contrast, the distinction between (a) and (b) is less obvious, so it is worthy of considering the details for a while. The following example, borrowed from Gilbert Harman (1977), will sharpen the distinction. A physicist observes a vapor trail in a cloud chamber, and she conjectures that there goes a proton. Vapor trails are observable, and they are the physicist's evidence for the existence of unobservable protons.[10] This story is compatible with the evidential version, for the physicist does *not use* any unobservable thing *as evidence*: a proton is not her evidence, it is her explainer of a vapor trail. But the story is *not* compatible with the methodological version, since the physicist *does refer* to something unobservable. Notice that this is not a refutation of (b), because it is a doctrine about meaning, not about particles. The example is merely an analogy. The point of this example is to illustrate the distinction between taking unobservable things as evidence and referring to unobservable things in theories.

With clear distinctions between the three versions in hand, we can locate Quine's position or positions and see how they relate to naturalism and verificationism. Quine is often equivocal when talking about his behaviorism. As a result, he seems to accept all three versions I formulated above. He adopts (a) when he asserts: "Empiricism as a theory of truth thereupon goes by the board, and good riddance. As a theory of evidence, however, empiricism remains with us" (1981, 39). His behaviorism here, as a tenet of his empiricism, is also a theory of

---

[9] That is, we cannot make reference to mental items. This can mean at least two things. First, we cannot make reference to mental items at all and, second, we cannot make reference to mental items as irreducible. Here I mean to keep (b) vague enough to encompass a broader territory. I thank a reviewer for pressing me on this point.

[10] This may not be a proper example today, for science has advanced very much since the 1970s. This should not blur the point of the example. Presumably science will discover smaller particles in the future, but the distinction I draw remains.

evidence. As Dagfinn Føllesdal writes: "Evidential behaviorism, and which I regard as that of Quine, is a position concerning *evidence*: the only evidence we can build on in our study of man … is empirical evidence, in particular the observation of behavior" (1990, 98). The objection Quine has to intention-based semantics shows, however, that he is a methodological behaviorist too, for he urges that it is not legitimate to make reference to intention or any other unobservable items in theory of meaning. This is where verificationism comes in, as de Rosa and Lepore rightly point out. The spirit of verificationism is that the target's ontology is exhausted by *our* epistemology of it. Intention is not observable *for us*, so it cannot *appear in* our theory of meaning.

Compare evidential behaviorism's attitude toward intention: intention is not observable *for us*, so it cannot *be evidence* for or against our theory of meaning. Notice that "evidence" is an epistemological concept: evidence available to A is bound with A's epistemological power. That's why evidential behaviorism is an uncontroversial claim. But the same does not hold in methodological behaviorism. To say this is not to say that methodological behaviorism is false. The only thing I want to stress is that the methodological version is not as uncontroversial as the evidential one. All the worse, in his later writings Quine sometimes talks as if he does not admit the existence of mental phenomena. For example: "Physically construed, [a mental state] is a state of nerves" (Quine 1985, 6). To this extent he is an ontological behaviorist. Notice that Quine is not inconsistent in adopting all these versions, for (c) entails (b), and (b) entails (a). Now, in making fundamental reference to mental states in a theory of meaning, H. P. Grice's theory would be inadmissible on (b) or (c), but admissible on (a), which requires only that intentions and other mental phenomena cannot be evidence for any particular theory of meaning. Such things could nonetheless be posited, even at a fundamental level, on the basis of their usefulness in helping us explain the evidence, much as unobservable, fundamental subatomic particles are posited by physicists to explain the phenomena they study.[11]

I want to say something more about evidential behaviorism and its relation to the methodological one. The word "evidence" is ambiguous. Consider a conversation between John and Tom. Tom says to John: "You say that someone has been stalking you for a week; show me the *evidence*." John replies: "I never see or hear him, but I know he is somewhere around me. Last week in a car accident I hurt a teenager. His friends told me that they will retaliate." Tom shrugs his shoulders: "Well, this might be a good *reason* you are so afraid, but a week has passed and nothing has happened. If you never see them, you should

---

[11] I come back to the Gricean view at the beginning of section 4. I thank a reviewer for urging me to expand on this point.

forget about the threat." Tom might be wrong about his advice, but the point is that he is drawing a distinction between reason and evidence. In this sense, for something to be evidence, it must be seen or perceived in some other ways. This, however, is not the only sense of "evidence." Sometimes the concept "evidence" contains the concepts "reason," "explanation," and the like. So we have at least two senses of "evidence."

If Quine's evidential behaviorism involves the wide sense of "evidence," it will face a certain kind of counterexample. Wide evidential behaviorism has it that in constructing and testing theories of meaning, we can only use observable behaviors as *reasons*. The counterexample runs as follows. Suppose we have a semantic hypothesis. Suppose further that we find this hypothesis inconsistent with a certain well-established theory in another field. The trouble is that one may reasonably reject the semantic hypothesis on the ground that it is inconsistent with some well-founded theory. One may be wrong to do so for certain reasons; for example, there is no genuine inconsistency between them or we have better reasons to renounce the well-established theory. It is undeniable, however, that at least in some cases one's rejection of the semantic hypothesis on the ground of its inconsistency with other theories is a right move. In this kind of case, the reason we give up our theory of meaning is *not* anything observable, given that in the present context "observable" means "perceivable by sense organs."[12] I would charitably suggest we not attribute wide evidential behaviorism to Quine.

Then how about narrow evidential behaviorism? The narrow version has it that in constructing and testing theories of meaning, we can use only observable behaviors as narrowly construed evidence. This version is true, but unfortunately it is *plainly* true: "Observable by sense organs" is built into the narrow sense of "evidence," so to say that we can use only observable behaviors as evidence is just stating a tautology, or something near enough. Evidential behaviorism is an unfortunate thesis: widely construed, it is false; narrowly construed, it is true but trivially true. Quineans often give much weight to this thesis. For instance, Føllesdal distinguishes evidential and ontological behaviorisms, and he argues that in arguing for the indeterminacy of translation, the evidential version is enough. After introducing evidential behaviorism, he notes that there is an argument for indeterminacy: namely, the conjunction of the Duhem thesis and verificationism. It is widely accepted that in the context of translation, behaviorism exemplifies verificationism, and by combining it with the Duhem thesis we can reach indeterminacy. The point I am trying to make is that evidential

---

[12] Here I am indebted to Timothy Lane.

behaviorism is not as substantial as some Quineans think. What they need is *methodological* behaviorism, which genuinely exemplifies verificationism. As I said, the distinction between the evidential and the methodological version is less obvious. One is correct to note that in deducing indeterminacy one does not need ontological behaviorism, the strongest version of this doctrine. Many people, however, tend to conflate evidential and methodological behaviorisms. The evidential version has the virtue of triviality; that is, it is not controversial. The methodological version has the virtue of instantiating verificationism; therefore it fits well with Quine's overall philosophical position. I have tried to show that one cannot have it both ways: evidential behaviorism and methodological behaviorism are different theses; the former is not controversial, but it is too weak to serve Quine's purposes; the latter can serve Quine's purposes, for it is a version of verificationism, but it is controversial and, as I shall argue, is not supported by naturalism.

Quine's naturalism has it that philosophy is continuous with science, which amounts to saying that both of them are *a posteriori* in nature. Evidential behaviorism can be derived, though trivially, from this tenet of naturalism, since evidence for *a posteriori* statements must be sensory evidence. This should not surprise us, for evidential behaviorism, narrowly construed, is almost a tautology. Moreover, this version of behaviorism is compatible with mentalism about meaning. Generally speaking, mentalistic semantics urges that we should treat unobservable mental states (for example, intention) as the primary bearer of meaning. Mentalists do *not*, however, take those unobservable items *as evidence* for or against semantic theories. Evidential behaviorism is so weak that almost everyone is willing to accept it. The controversial theses are methodological and ontological versions. Both are incompatible with mentalism because they do not refer to mental states, but unfortunately they are unsupported by naturalism, since there is no reason for *a posteriori* philosophy and natural sciences to confine themselves to observable items. If an unobservable item carries strong explanatory power, surely we can make reference to it in our theory.[13] There is nothing in naturalism that prevents us from doing so.

Then how about verificationism? Verificationism conflates evidence and meaning; it is exemplified by methodological behaviorism in the context of translation. Methodological behaviorism prevents us from referring to unobservable items in theories of meaning, for it is a special version of verificationism, which confines us to observable things in constructing semantic theories. Naturalism also lends no support to verificationism, simply because naturalism does not preclude intention or other mental states referred to in semantics, unless one regards them as certain

---

[13] Both Searle (1983) and Dennett (1987) insist on this point, though they give drastically different status to what is referred in intentional explanations.

non-natural items. Nonetheless, semantic mentalism is not forced to treat mental states that way. To summarize, there are (at least) three versions of behaviorism. The weakest is supported by naturalism, but it is too weak to refute mentalism. Two other versions of behaviorism, methodological and ontological, are indeed incompatible with mentalism. They, however, are not, *pace* Quine, entailed by naturalism. Besides, verificationism and methodological behaviorism are at root identical; neither gains any support from naturalism.

The reason the derivation of evidential behaviorism from naturalism is trivial is that evidential behaviorism is an innocuous thesis; it can stand on its own without support from naturalism. For this reason, one might feel that we need not assert that naturalism entails this version of behaviorism. To this I shall reply that I cannot find other plausible readings of "ontological relativity." Quine is explicit on the inference from naturalism to behaviorism. Methodological and ontological versions cannot, however, be derived from naturalism. The only possibility, as far as I can see, is to regard the evidential version as the behaviorism in question. This is an unwanted consequence for Quineans, since this version of behaviorism is too weak to repel mentalism, but I do not know what interpretation would be both plausible and substantial.

Now, three provisos for my objection. First, in my quotation from Quine on naturalism, he only talks about knowledge, mind, and meaning, but in fact his naturalism is a more thoroughgoing (therefore more controversial) thesis. Fortunately, other portions of his naturalism are irrelevant to my purposes in this paper, for the issue I want to tackle is whether naturalism precludes mentalistic semantics. It is enough to identify my opponent as holding that the study of meaning is subject to naturalism. Second, the line of thought I object to is from naturalism to behaviorism and finally to the denial of mentalism. I have argued that although the methodological and ontological versions are indeed inconsistent with mentalism, these two theses are both unsupported by naturalism. Therefore, the inference from naturalism to anti-mentalism collapses. Even if I am right about this, however, it is possible to find other resources to support these two versions of behaviorism and thereby refute mentalism. The third proviso is related to this. In my discussion above I construct no argument for the truth of mentalism, nor do I refute all arguments against mentalism provided by Quine. The only thing I have done is point out that naturalism supports a minimal version of behaviorism—namely, the evidential one—but this version does not preclude mentalistic semantics.

4. Mentalism and Evidential Behaviorism: Some Examples

In this section I turn to some examples that exemplify both mentalistic semantics and evidential behaviorism-cum-naturalism. Gloria Origgi and

Dan Sperber attempt to "fi[t] (post-) Gricean pragmatics into Millikan's conceptual framework" (2000, 158). Kim Sterelny (2003) pursues a similar Gricean line, and I shall describe it in somewhat simplified way.[14] Grice (1957) draws a distinction between natural and non-natural meaning. Human language belongs to non-natural sides, and Grice gives a threefold psychological picture of this sort of meaning. Let S be the speaker, H be the hearer, u be the utterance, and p be the content of u. The analysis is: (a) S utters u and intends that H form the belief that p; (b) S further intends that H recognize S's original intention; and (c) S still further intends that H form the belief that p at least partly on the basis of the recognition of original intention. Let's ignore all the details and problems inherited in this analysis, for I am not going to defend this kind of program here. In responding to Ruth Millikan's objection (1984) to this approach, Sterelny notes the distinction between understanding and acceptance. He argues: "It is in the interest of speakers to make the detection of syntactic structure and communicative intention as easy as possible, and it is in the audience's interest to recognize that structure and those intentions. Thus the proximate function of speech is to signal a communicative intention. It is in the interests of the audience to recognize that intention, whether or not it is also in the audience's interest to accept what is said. For identifying those intentions gathers important information in itself" (2003, 184). The point Sterelny is making is that even if the auditor wants to use or fool our speaker, he usually needs to understand the speaker's intention in advance. This feature gives some prima facie plausibility of the Gricean project. Origgi, Sperber, and Sterelny's projects are, undoubtedly, naturalistic; all of them accept evidential behaviorism.[15] Intention, as an unobservable item, does *not* figure in their semantics as evidence. The quotation above from Sterelny is an empirical generalization: it is generalized from his (and our) everyday experiences. To mention unobservable things in one's theory does not make the theory a mystical story. The naturalistic objection Quine has to mentalism is an overreaction; mentalism *may* be wrong, but the reason for this cannot be that it is inconsistent with his version of naturalism. They are perfectly compatible.

5. Taking Stock

Let me summarize my main claims and their rationales with the following statements.

---

[14] One orthodox Gricean in philosophy is John Searle. He also champions an intention-based semantics and holds a version of naturalism. For his modification of Gricean analysis, see Searle 1969 and 1983.

[15] It is quite possible that they do not accept specific claims in Quine's naturalism, but this would not affect my point, for my focus in this section is the relation between mentalism and evidential behaviorism.

Claim 1: Quine's naturalism consists in three tenets, as opposed to two.

Rationale: Gibson's reconstruction contains two tenets of Quine's naturalism. Both are epistemological in nature. There is, I submit, an ontological tenet underlining them: namely, that everything that exists is natural in character.

Claim 2: Gibson's point (2) faces two problems. They are derived from a false dichotomy.

Rationale: First, some statements are remote from the tribunal of experience. They have too little empirical content to be encompassed by natural science. Second, some experiments are empirically, technologically, or morally impossible. *A posteriori* philosophy can fill the gap.

Claim 3: Quine himself does not accept the false dichotomy.

Rationale: Quine thinks that philosophy is continuous with science. Both of them are *a posteriori*. The dichotomy fails to take *a posteriori* philosophy into account.

Claim 4: Quine makes the following inference: behaviorism follows from naturalism, and we should thereby reject mentalism.

Rationale: Quine explicitly says so in "Ontological Relativity" (1969b, 26).

Claim 5: There is a distinction between evidential behaviorism and methodological behaviorism.

Rationale: Quineans often hold that to dispel mentalism and to deduce the indeterminacy of translation, one need not adopt ontological behaviorism. The point is well taken, but they often neglect another genuine distinction, between evidential behaviorism and methodological behaviorism. The former is about observational terms, the latter about theoretical terms. Quineans need the latter, not the former.

Claim 6: There are (at least) two senses of "evidence." We should attribute narrow evidential behaviorism to Quine.

Rationale: Wide evidential behaviorism faces a kind of counterexample: when our semantic theory is inconsistent with other well-grounded theories, one is justified in abandoning the semantic theory. This reason is, literally speaking, unobservable. We should charitably attribute the narrow version to Quine.

Claim 7: Quineans cannot have it both ways when they invoke behaviorism.
Rationale: Evidential behaviorism is not controversial, but it is too weak to sustain any relevant substantial claim (such as indeterminacy, anti-mentalism, and so on). The methodological version can serve Quinean purposes, but it is debatable and not supported by naturalism (see claim 11 below).
Claim 8: Naturalism supports evidential behaviorism.
Rationale: *A posteriori* philosophy and natural sciences do not acknowledge any unobservable item *as evidence*. Narrow evidential behaviorism is just a corollary of naturalism.

Claim 9: Evidential behaviorism is compatible with mentalistic semantics.
Rationale: Mentalists do not take those unobservable items as evidence for or against semantic theories. Even dedicated mentalists, say Searle and Fodor, will accept this version. What they oppose are methodological and ontological behaviorisms.

Claim 10: Methodological and ontological behaviorisms do dispel mentalism.
Rationale: To be a mentalist, one has to make reference to mental items (entities or not) as the primary bearers of meaning. These two versions of behaviorism stop us from referring to unobservable things, so they are not compatible with mentalism.

Claim 11: However, methodological and ontological behaviorisms gain no support from naturalism.
Rationale: There is no reason for *a posteriori* philosophy and natural science to confine themselves to observable items. A naturalist can invoke unobservable items as theoretical posits, as long as these posits carry substantial explanatory powers.

Claim 12: Verificationism also acquires no support from naturalism.
Rationale: Again, naturalism does not preclude intentions or other mental states to be referred to in semantics, unless those unobservable things are construed as non-natural items by mentalists.

Claim 13: The theorists I have listed are all mentalists as well as evidential behaviorists.
Rationale: All of them treat intention (or something like that) as the primary locus of meaning, and all of them are evidential behaviorists-cum-naturalists. They do not invoke anything unobservable as their evidence in theorizing semantics.

These scaffold-like remarks do not exhaust my contentions and their rationales, but I believe that they will suffice as a summary of the main thoughts of this paper. Naturalism is an important thesis in contemporary philosophy, and it has espousers from miscellaneous philosophical camps; some are even outright incompatible with each other. So it is crucial to cash out the content and tenets of the doctrine of naturalism and the implicative relations between it and many other philosophical positions. This paper, which has so far focused mainly on Quine's version of the doctrine, serves to take up this conceptual task. The relations of implication I argued for or against are by no means exhaustive, and I hope to explore them more in the future. In the remainder of the paper I turn to a contemporary application of the distinction between versions of behaviorism.

6. Fodor's Misfire

In the previous sections I criticized a behaviorist overreaction to mentalism. Along the way I distinguished three kinds of behaviorism and connected them to naturalism and verificationism. I now want to apply these resources to a reverse overreaction. It is Jerry Fodor's review (2003) of José Luis Bermúdez's *Thinking Without Words* (2003).

The main objective of Bermúdez's book is to provide a framework for scientists when they try to explain the behaviors of nonhuman animal infants and human infants. The substantial claim underlying this objective is that, at least in some cases, psychological explanations are needed if we want to explain the behaviors of nonlinguistic creatures appropriately. In his review of the book, Fodor defends his language of thought hypothesis from Bermúdez's objections. The focus here will be Fodor's accusation that Bermúdez commits certain mistakes underlying behaviorism. I begin my discussion by presenting a set of distinctions made by Bermúdez.

Bermúdez distinguishes four dimensions of an *adequate* theory of nonlinguistic thought. They are:

(a) The *metaphysical* dimension of explaining how nonlinguistic thought is possible and what its vehicles might be.

(b) The *semantic* dimension of explaining the semantics of nonlinguistic thought.

(c) The *epistemological* dimension explaining how it is possible for us to identify the content of such thought.

(d) The *explanatory* dimension of elucidating the inferential and decision-making processes of nonlinguistic creatures in a way that underwrites the practice of giving psychological explanations of their behavior. (Bermúdez 2003, 31)

With these dimensions, Bermúdez argues that the language of thought hypothesis "fails to provide an epistemology for the thoughts of nonlinguistic creatures, [so] it cannot be the whole story" (2003, 31). And again: "The language of thought hypothesis cannot provide us with a wholly satisfactory account" (2003, 31). Since Fodor himself does not rebut the point that his hypothesis fails to fulfill the epistemological dimension, and since this point is irrelevant to my discussion, let's grant, for the sake of simplicity, that the language of thought hypothesis does fail to do this. To this point Fodor replies: "I think Bermúdez' insistence on this methodology is very surprising; and I think it is disastrous. . . . In fact, scientific theory don't usually provide 'operational criteria' for attributing the states, events, processes (whatever), that they purport to describe . . . string theory doesn't tell us how to tell whether string theory is true. . . . The *truth* of string theory depends on whether there are strings. But the *testing* of string theory depends on relations between strings and us . . . none of this epistemological stuff is relevant to whether there are strings; ontology is one thing, epistemology is quite another" (Fodor 2003, 5).

Fodor's accusation here is the conflation of epistemology and ontology in Bermúdez's thinking. According to my taxonomy of behaviorism, Fodor's target is methodological behaviorism. I have two rejoinders to Fodor's complaint. For one thing, even if Fodor's remarks on string theory are true, the situation *may* be different in theories of meaning. A methodological behaviorist or verificationist might reply that meaning has a very special mode of being, so the distinction does not hold in the realm of meaning. Therefore, people siding with Fodor need to provide arguments to rebut this claim. Fodor himself (1987) does offer an argument to this point. For another thing, and more important, Bermúdez is *not* a methodological behaviorist. Certainly, he is an evidential behaviorist; his works are filled with empirical data, and he treats those data as the only evidence for his theory. But he is not a methodological behaviorist, for he insists on giving psychological explanations, and therefore attributing mental states, to nonlinguistic creatures. In so doing he makes reference to mental states. Fodor's contention here is a misfire.

Furthermore, as I argued in section 3, evidential behaviorism is perfectly consistent with mentalistic semantics. In the present context, we

should say that Bermúdez's project based on evidential behaviorism is consistent with Fodor's semantic mentalism—namely, his language of thought hypothesis. Bermúdez's own remarks fit this point nicely: "The approach I develop is compatible with certain ways of developing [Fodor's] hypothesis. Supporters of the language of thought hypothesis can take the theory put forward in this book as an attempt to provide an epistemology and semantics for the application of the language of thought hypothesis to nonlinguistic creatures" (Bermúdez 2003, 31). It seems that Fodor regards Bermúdez's writing about his language of thought hypothesis as an *objection from methodological behaviorism*. That's why he invokes the example of string theory to make his reply. He is right that methodological behaviorism is incompatible with his approach, and that what he needs is a sharp distinction between epistemology and ontology, but he is wrong in thinking that Bermúdez is a methodological behaviorist. For further textual evidence, recall that what Bermúdez says is that the language of thought hypothesis "cannot be the *whole* story"; this hypothesis "cannot provide us with a *wholly satisfactory* account" (2003, 31, my italics). Bermúdez does not reject the language of thought hypothesis, he only exposes the inadequacy of it.

7. Diagnosis

In the course of this discussion, we have found that both behaviorists and mentalists tend to overreact to each other. The root of this situation, I submit, is the improper stigmas these two terms carry. In the modern history of philosophy, the doctrine of mentalism was often accompanied by Cartesianism, and there were philosophers—say, Gilbert Ryle (1949)—who correctly banished this doctrine but went too far. On the other hand, in the twentieth century behaviorism was often accompanied by verificationism or eliminativism. Therefore, mentalists like Fodor and Searle argue against behaviorism to their last breath. We can bring peace to all these controversies by removing those stigmas, as I have tried to do on this occasion. The main moral is that if we understand varieties of naturalism, behaviorism, and mentalism properly, a more sensible approach to semantics is open to view.

Department of Philosophy
University College London
Gower Street
Bloomsbury
London WC1E 6BT
United Kingdom
h.cheng.12@ucl.ac.uk

Acknowledgments

References

Bermúdez, José Luis. 2003. *Thinking Without Words*. New York: Oxford University Press.

Buchler, Justus. 1955. *Nature and Judgment*. New York: Columbia University Press.

Dennett, Daniel. 1987. *The Intentional Stance*. Cambridge, Mass.: MIT Press.

de Rosa, Raffaella, and Ernest Lepore. 2004. "Quine's Meaning Holisms." In *The Cambridge Companion to Quine*, edited by Roger F. Gibson Jr., 65–90. New York: Cambridge University Press.

Dewey, John. 1988 [1929]. *Experience and Nature*. Revised edition. In *John Dewey: The Later Works, 1925–1953, Volume 1*, edited by Jo-Ann Boydston. Carbondale: Southern Illinois University Press.

Fodor, Jerry. 1987. *Psychosemantics*. Cambridge, Mass.: MIT Press.

———. 2003. "More Peanuts." *London Review of Books* 25:1–9. (Page references are to the online version: https://www.lrb.co.uk/v25/n19/jerry-fodor/more-peanuts.)

Føllesdal, Dagfinn. 1990. "Indeterminacy and Mental States." In *Perspectives on Quine*, edited by Robert Barrett and Roger F. Gibson Jr., 98–109. Hoboken, N.J.: Wiley-Blackwell.

Gibson, Roger F., Jr. 1982. *The Philosophy of W. V. Quine*. Gainesville: University Press of Florida.

———. 1995. "Quine, Willard Van Orman." In *A Companion to Metaphysics*, edited by Jaegwon Kim and Ernest Sosa, 426–28. Oxford: Blackwell.

———. 2004. "Quine's Behaviorism cum Empiricism." In *The Cambridge Companion to Quine*, edited by Roger F. Gibson Jr., 181–99. New York: Cambridge University Press.

Grice, Paul. 1957. "Meaning." *Philosophical Review* 66:377–88.

Harman, Gilbert. 1977. *The Nature of Morality*. New York: Oxford University Press.

Jackson, Frank. 1982. "Epiphenomenal Qualia." *Philosophical Quarterly* 32:127–36.

McDowell, John. 1996. *Mind and World*. Cambridge, Mass.: Harvard University Press.

Millikan, Ruth. 1984. *Language, Thought, and Other Biological Categories*. Cambridge, Mass.: MIT Press.

Orenstein, Alex. 2002. *W. V. Quine*. Princeton University Press.

Origgi, Gloria, and Dan Sperber. 2000. "Evolution, Communication and the Proper Function of Language." In *Evolution and the Human Mind: Modularity, Language and Meta-Cognition*, edited by Peter Carruthers and Andrew Chamberlain, 140–69. Cambridge: Cambridge University Press.

Parfit, Derek. 1986. *Reasons and Persons*. Oxford: Oxford University Press.

Quine, Willard Van Orman. 1969a. "Natural Kinds." In *Ontological Relativity and Other Essays*, 114–38. New York: Columbia University Press.

———. 1969b. "Ontological Relativity." In *Ontological Relativity and Other Essays*, 26–68. New York: Columbia University Press.

———. 1981. "On the Very Idea of a Third Dogma." In *Theories and Things*. Cambridge, Mass.: Harvard University Press.

———. 1985. "States of Mind." *Journal of Philosophy* 82, no. 1:5–8.

Randall, John Herman. 1958. *Nature and Historical Experience*. New York: Columbia University Press.

Ryle, Gilbert. 1949. *The Concept of Mind*. London: Routledge.

Searle, John. 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

———. 1983. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.

Sterelny, Kim. 2003. *Thought in a Hostile World: The Evolution of Human Cognition*. Oxford: Blackwell.