

Quine on Explication and Elimination

MARTIN GUSTAFSSON
Stockholm University
Stockholm
Sweden

I Ordered Pairs and Mental Entities

In section 53 of *Word and Object*, Quine argues that the set-theoretical explications of the concept of the ordered pair offered by mathematicians such as Wiener and Kuratowski give us a model for the clarification of philosophically troublesome ideas. According to Quine, ordered pairs might seem indispensable in science. But at the same time they have appeared unclear to many philosophers, who have argued that an extensional treatment of the logic of relations can be satisfactory only to the extent that we can give a transparent and substantial explanation of what an ordered pair really is.¹ Quine cites Peirce as someone who tries to meet this sort of demand. The explanation Peirce offers is mentalistic, but it is not the mentalism that Quine regards as the most fundamental problem with Peirce's account. Rather, Quine finds the very idea that a substantial philosophical explanation is needed mistaken. According to Quine, what Wiener and Kuratowski offer is something very different: they free us from trouble by showing how to replace the notion of ordered pair with a clearer substitute that fulfills all the relevant scientific functions of the old concept. Thus, instead of directly confronting those aspects of the idea of ordered pair that confuse us, Quine argues that we do not have to care about those confusing aspects. Allegedly, explic-

1 See, for example, Russell (1903), 98. Interesting historical surveys of the troubles logicians and philosophers have had in trying to handle the notion of ordered pair are given in Dipert (1982) and Kanamori (2003).

tions like those of Wiener and Kuratowski dissolve rather than solve our difficulties: 'when explication banishes a problem it does so by showing it to be in an important sense unreal; viz., in the sense of proceeding only from needless usages' (Quine, 1960, 260).

Spontaneously, one might want to object that it is essential to ordered pairs that they can contain the same members and yet be different: $\langle x, y \rangle \neq \langle y, x \rangle$. Hence, it may be argued, no set-theoretical substitute can fully capture the sense in which ordered pairs are *ordered*. Quine, however, rejects all such talk of essences and senses. As I will show, this anti-essentialist attitude is intimately related to his view of the ontological import of explication procedures. According to Quine, an explication should help us reform our established theory of the world so that the resulting scheme assumes an ontology that is as sparse as possible. Allegedly, what Wiener and Kuratowski show is simply that whatever good is accomplished by talking of ordered pairs can be accomplished by talking of sets. And, Quine adds, '[a] similar view can be taken of every case of explication: *explication is elimination*' (Quine, 1960, 260).

What, exactly, does Quine mean by 'elimination' in this context? One thing he seems to have in mind is that explication shows a certain vocabulary to be scientifically superfluous. There is, in this sense, a terminological elimination involved. More importantly, Quine thinks explication makes possible the elimination of ontological commitments. The explication procedure allows us to abandon an old theory where certain existential assumptions are made, in favor of a new, streamlined version in which those assumptions are not made.

It may therefore seem reasonable to classify Quine as an *eliminativist* with regard to the objects purportedly referred to by the explicated terms. And this label is all right, as long as one keeps in mind that Quine is an eliminativist only in a rather unusual sense of the word. In fact, given a more common use of the term 'eliminativism,' Quine is not an eliminativist at all. For being an eliminativist in the usual sense of the word involves distinguishing one's own standpoint from views that *identify* the objects purportedly referred to by the explicated terms with the objects referred to by the terms doing the explication. For example, in the case of ordered pairs, an eliminativist in the usual sense of the word takes successful explication to show that ordered pairs do not exist, *as opposed to* showing that they are identical with sets. What makes Quine's eliminativism unusual is that he does not acknowledge this opposition. He thinks the alleged distinction between explication as elimination and explication as identification is without substance.

The significance and originality of Quine's conception of explication is particularly striking when it comes to his account of what it means to give a physicalist explication of mental concepts. According to Quine, the point of positing mental states and events can only be that such

positing has a systematizing function in theories of human behavior, analogous to the way in which the positing of elementary particles in physics allows us to give systematic predictions and explanations of the behavior of physical objects. But, he asks, would not the positing of bodily states and events have a similar systematizing function? '[S]urely,' he continues,

as much organization could be achieved by positing merely certain correlative physiological states and events instead. Nor need we spot special centers in the body for these seizures; physical states of the undivided organism will serve, whatever their finer physiology. Lack of a detailed physiological explanation of the states is scarcely an objection to acknowledging them as states of human bodies, when we reflect that those who posit the mental states and events have no details of appropriate mechanisms to offer nor, what with their mind-body problem, prospects of any. The bodily states exist anyway; why add the others? (Quine, 1960, 264)

The spirit of this passage may seem straightforwardly eliminativist, in the usual sense of the word. Apparently, Quine is arguing that we should not assume the existence of mental states and events, since bodily states and events suffice to handle the relevant data and produce the predictions we want. Thus, he might seem to be contradicting a supposedly alternative line of thought, according to which what happens when we abandon the old mentalistic theory in favor of a new, physicalist successor is *not* that we drop the assumption that mental states and events exist, but that we reject the mistaken assumption that mental states and events are distinct from physical states and events. According to this seemingly alternative way of viewing the matter, the mistake of the old theory is that it involves certain false claims about mental states, claims according to which mental states are not physical. But, the argument goes, abandoning those false claims does not amount to banishing mental states from the universe. Instead, it means coming to understand better the true nature of such states.

This latter conception characterizes standard forms of the so-called identity theory of mind. If Quine were an eliminativist in the usual sense, one would expect him to reject the identity theory as false. This, however, he does not do. On the contrary, he sees no substantial conflict at all here. Indeed, he explicitly says that 'the distinction between an eliminative and an explicative physicalism is unreal' (Quine, 1960, 265). In Quine's view, the supposedly fundamental difference boils down to a matter of rhetoric. To insist on one of these alternatives, he says, is just to insist on 'a way of phrasing matters' (Quine, 1960, 261).

At least before the 1980s, Quine tended to prefer the eliminativist idiom. His most explicit reason for doing so is that the eliminativist way of talking provides a more secure protection against the allure of men-

talism. '[P]roduct though the identity theory is of hard-headed materialism,' Quine says,

we must beware of its sedative use to relieve intellectual discomfort. We can imagine someone appealing to the identity theory to excuse his own free and uncritical recourse to mentalistic semantics. We can imagine him pleading that it is after all just a matter of physiology, even if no one knows quite how. This would be a sad irony indeed, and the repudiation theory has the virtue, over the identity theory, of precluding it. (Quine, 1975, 95; cf. Stemmer, 2001, Gibson, 2004a)

The mere fact that Quine prefers the eliminativist way of talking, however, does not mean that he can be classified as an eliminativist in the usual sense of the word. In fact, Quine's justification of eliminativist talk in terms of what is rhetorically advisable makes it quite clear that his eliminativism is not the usual one. Again, eliminativism typically involves the view that there is a difference that is *not* just a matter of rhetoric between claiming that mental states do not exist and claiming that mental states are identical with physical states. Insofar as eliminativism and the identity theory are defined as standing in such substantial opposition to one another, Quine is neither an eliminativist nor an identity theorist. According to Quine, using explication as a philosophical method means refusing to take sides in conflicts of this sort.

This aspect of Quine's conception of explication has not received much attention in the literature. Quine is often and more or less routinely classified as an 'eliminative materialist,' while the details of his actual viewpoint are ignored. In what follows, my aim is to provide a more faithful account of his conception. In section II, I look more closely at Quine's paradigm example of ordered pairs. In section III, I clarify the philosophically more central case of mental entities. In the course of my investigation, I exhibit the connection between Quine's conception of explication and other fundamental aspects of his philosophical outlook. I highlight the importance of his doctrine of ontological relativity, and consider his general view of the relation between language, experience and reality upon which that doctrine rests. In this connection, I also stress the crucial significance of Quine's commitment to extensionalism.

II A Closer Look at the Case of Ordered Pairs

The case of ordered pairs is unusually clear-cut, says Quine, since the notion of ordered pair was more or less deliberately introduced as subject to one very simple postulate, namely,

- (1) If $\langle x, y \rangle = \langle z, w \rangle$, then $x = z$ and $y = w$.

Wiener proposed that we identify the ordered pair $\langle x, y \rangle$ with the set

$\{\{x\},\{y,\emptyset\}\}$.² Identical sets have the same members; hence, if $\{\{x\},\{y,\emptyset\}\} = \{\{z\},\{w,\emptyset\}\}$, then $x=z$ and $y=w$. So, Wiener's construal takes care of postulate (1). Admittedly, Wiener's substitute does not have 'the same meaning' as the original concept. But this is not a pertinent objection to the explication procedure, for explication in Quine's sense involves no synonymy claims. On the contrary, that there are certain differences between the original concept and the substitute might be said to be the very point of the procedure. Again, one of the aims of explication is precisely to get rid of those confusing characteristics that make the original concept different from the new one, and that might tempt us to ask questions which are better treated as extraneous don't-cares.

An important consequence of Quine's view of explication is that one and the same concept might be explicated in different ways, all of which are equally acceptable. As long as the proposed substitutes do the relevant theoretical work, everything is all right. Thus, Kuratowski's explication of the notion of ordered pair, according to which $\langle x,y \rangle$ is identical with $\{\{x\},\{x,y\}\}$, is just as good as Wiener's. As Quine puts it, Wiener's and Kuratowski's construals 'conflict with one another only out among the don't-cares' (Quine, 1960, 260).

This relaxed attitude toward different explications of one and the same concept — 'Which is right? All are' (Quine, 1960, 260) — might seem exceedingly tolerant. For two explications of one and the same concept are not just different in meaning. Often they are incompatible even at the extensional level. Thus, Wiener identifies the ordered pair $\langle 2,7 \rangle$ with the set $\{\{2\},\{7,\emptyset\}\}$, whereas Kuratowski identifies the same pair with $\{\{2\},\{2,7\}\}$. Now, if we hold that these explications are equally correct, then don't we, in effect, subscribe to the claim that $\langle 2,7 \rangle$ is identical with *both* $\{\{2\},\{7,\emptyset\}\}$ and $\{\{2\},\{2,7\}\}$? Which, by the symmetry and transitivity of identity, entails that $\{\{2\},\{7,\emptyset\}\} = \{\{2\},\{2,7\}\}$; which is absurd.

This objection involves a misunderstanding of Quine's point (cf. Benacerraf, 1965). When Quine claims that Wiener's and Kuratowski's explications are equally right, he is not saying that the sentences ' $\langle 2,7 \rangle = \{\{2\},\{7,\emptyset\}\}$ ' and ' $\langle 2,7 \rangle = \{\{2\},\{2,7\}\}$ ' are both true. Rather, what he means is just that the streamlinings of theory offered by these explications do not differ significantly in terms of simplicity and systematicity, and that no one is better than the other at helping us foresee new observations in the light of previous sensory experience. That is, Wie-

2 This is Quine's rendering of Wiener's definition. In fact, Wiener worked within Russellian type theory, and a more faithful rendering of his definition is: $\langle x,y \rangle = \{\{x\},\{\{y\}\}\}$, where x and y are of the same type and \emptyset is the null class of the next type. Cf Wiener 1914 and Kanamori 2003, 290.

ner's and Kuratowski's explications are equally right in the following sense: no matter which alternative we choose, we can get rid of the assumption that ordered pairs exist besides sets, and yet continue to exploit all parts of theory where the old notion of ordered pair used to do its work. With similar gains in theoretical expediency, the suggested explications show how to abandon the assumption that ordered pairs exist *in addition to* sets. The ontological commitments of ordinary set-theory suffice to account for the scientific function of ordered pairs.

Even if they are in principle superfluous, it will be convenient to retain expressions of the form ' $\langle x,y \rangle$ ' as parts of our vocabulary. Accepting Wiener's explication means treating, say, ' $\langle 2,7 \rangle$ ' as interchangeable with ' $\{\{2\},\{7,\emptyset\}\}$,' whereas accepting Kuratowski's explication means treating ' $\langle 2,7 \rangle$ ' as interchangeable with ' $\{\{2\},\{2,7\}\}$.' These different construals are both fine, but mutually exclusive. What we are not allowed to do is treat ' $\langle 2,7 \rangle$ ' as interchangeable with *both* ' $\{\{2\},\{7,\emptyset\}\}$ ' and ' $\{\{2\},\{2,7\}\}$.' Given Wiener's explication, $\langle 2,7 \rangle$ is identical with $\{\{2\},\{7,\emptyset\}\}$ and not with $\{\{2\},\{2,7\}\}$. Given Kuratowski's equally justified explication, $\langle 2,7 \rangle$ is identical with $\{\{2\},\{2,7\}\}$ and not with $\{\{2\},\{7,\emptyset\}\}$.

What comes to the surface here is the intimate connection between the possibility of mutually incompatible yet equally justified explications and Quine's doctrine of ontological relativity, according to which '[t]o say what objects someone is talking about is to say no more than how we propose to translate his terms into ours' (Quine, 1981, 20). Before Wiener and Kuratowski, mathematicians employed the unexplicated notion of ordered pair. If, after Wiener and Kuratowski, we purport to say what those old mathematicians were talking about, we propose a particular translation into set-theory — one among various. The point is that there is no such thing as *first* deciding what the mathematicians were talking about and *then* settling for a particular translation. We cannot, as it were, rise above the activity of translation and ask, from such a detached viewpoint, What were those old mathematicians *really* referring to? Saying that they were talking about sets of the form $\{\{x\},\{y,\emptyset\}\}$ is just to translate their pronouncements in terms of the Wiener explication; saying that they were talking about sets of the form $\{\{x\},\{x,y\}\}$ is just to translate their pronouncements in terms of the Kuratowski explication; and that is all.

Hence it might be misleading to say that identity statements of the form ' $\langle x,y \rangle = \{\{x\},\{y,\emptyset\}\}$,' or ' $\langle x,y \rangle = \{\{x\},\{x,y\}\}$,' *justify* the transition from the old theory to a streamlined successor theory. Again, the transition from the old theory in which the notion of ordered pair had not yet been explicated to a streamlined successor theory of either the Wiener or the Kuratowski variety, is justified merely by the fact that the successor theory enables us to organize and predict sensory input in a more efficient manner. There is no additional justification consisting in the fact that $\langle x,y \rangle$ is somehow *really* identical with, say, $\{\{x\},\{y,\emptyset\}\}$.

In fact, if we prefer, we may account for the desired streamlining in terms of *replacement* rather than identification. This means opting for a purely eliminativist rhetoric: instead of saying that what the old mathematicians called ' $\langle x, y \rangle$ ' is identical with what we call ' $\{\{x\}, \{y, \emptyset\}\}$,' we say that what those old mathematicians purported to talk about did not exist. According to this way of phrasing matters, explication is not a matter of identifying ordered pairs with sets, but of showing that we can talk of $\{\{x\}, \{y, \emptyset\}\}$ *instead of* $\langle x, y \rangle$, the latter being treated as a non-existent entity on a par with Santa Claus. As long as we remember that such eliminativist jargon is neither more nor less justified than a construal in terms of identification, everything is all right. Our choice has no deeper foundation than rhetorical taste and terminological convenience.

All this hangs together with some of Quine's most central philosophical ideas. According to Quine, the basic connection between language and reality is not the referential connection between objects and words, but the causal connection between sensory stimuli and the observation sentences that we assent to or dissent from as prompted by those stimuli. This means, among other things, that what fundamentally matters to the empirical adequacy of a translation manual is not its particular specifications of reference, but its capacity to do justice to the manner in which the language that is being translated constitutes a response to sensory input. Postulating connections between words and objects is important only insofar as it helps the translator capture the ways in which sentences are related to the irritation of sensory surfaces. The translator's fundamental concern will be the patterns of sentence-to-sentence and sentence-to-stimuli links, as those patterns are exhibited in the behavior of the speakers of the language. As long as the translation manual provides a sufficiently good model of the relevant behavioral regularities, its ascriptions of reference are allowed to fall as they may. And if there are models that all do justice to the behavioral data, but that ascribe reference in mutually incompatible ways, there is no objective ground for picking one rather than another. Again, Quine thinks there is no absolute way of making ontological ascriptions: specifications of what someone is talking about are always relative to a manual of translation (Quine, 1986, 460 and 1992, 52; see also Hylton, 2004).

Where there is agreement about relevant criteria of identity, questions about reference, difference and identity are routinely resolved. Consider the claim that Wiener's and Kuratowski's explications are incompatible. This claim is accepted by all who understand the language of set-theory. Sharing the language of set-theory involves agreeing upon identity criteria for sets — criteria which entail that, in all cases where $x \neq \emptyset$, $\{\{x\}, \{y, \emptyset\}\} \neq \{\{x\}, \{x, y\}\}$. Thus, if they were asked, Wiener, Kuratowski, the old mathematicians who used the non-explicated concept of ordered pair, as well as modern readers of the present paper, can all be presumed

to agree that Wiener and Kuratowski assign different objects to expressions of the form ' $\langle x, y \rangle$.'

By contrast, there are no shared criteria of identity by means of which it can be decided whether or not, say, $\langle 2, 7 \rangle$ is identical with $\{\{2\}, \{7, \emptyset\}\}$. Importantly, what Quine is willing to count as behavioral evidence is of no help here, since the interpretation of such evidence presupposes that the relevant criteria of identity have already been determined. For example, let us suppose that mathematicians before Wiener would not have assented to sentences such as ' $\langle 2, 7 \rangle = \{\{2\}, \{7, \emptyset\}\}$,' or ' $\{2\} \in \langle 2, 7 \rangle$.' Indeed, for the sake of argument, let us suppose that those mathematicians explicitly dissented from such sentences. If Quine is right, this does not falsify the claim that their term ' $\langle 2, 7 \rangle$ ' referred to the set $\{\{2\}, \{7, \emptyset\}\}$. According to Quine, nothing stands in the way of construing their dissent as the manifestation of false beliefs about $\{\{2\}, \{7, \emptyset\}\}$, rather than as evidence that they did not use ' $\langle 2, 7 \rangle$ ' to talk about $\{\{2\}, \{7, \emptyset\}\}$. On the other hand, nothing makes such a construal *more* legitimate than a construal according to which $\langle 2, 7 \rangle$ is not identical with $\{\{2\}, \{7, \emptyset\}\}$ but with $\{\{2\}, \{2, 7\}\}$, or a construal according to which $\langle 2, 7 \rangle$ does not exist at all.

The usual, non-Quinean sort of eliminativist is not satisfied by being told that according to Wiener's explication, $\langle 2, 7 \rangle$ is identical with $\{\{2\}, \{7, \emptyset\}\}$, whereas according to Kuratowski's explication, $\langle 2, 7 \rangle$ is identical with $\{\{2\}, \{2, 7\}\}$. Nor is he satisfied by being told that his own claim, that $\langle 2, 7 \rangle$ does not exist, is just as fine as such identity claims. For he wants to say that what is *really* the case is that the $\langle 2, 7 \rangle$ which the old mathematicians purported to talk about does not exist. Indeed, it is precisely at this supposedly deeper level that the conflict between this usual sort of eliminativist and his opponent is meant to take place: they try to quarrel about the nature and existence of ordered pairs in a non-relative, absolute sort of way. According to Quine, this quarrel is idle.

It must be kept in mind that Quine's belittling attitude toward the seemingly fundamental worries about the identity and existence of ordered pairs is perfectly compatible with his claim that explication involves the elimination of ontological commitments. That such elimination occurs is not under debate. Quine agrees with both usual eliminativists and their opponents that set-theoretical explications of the notion of ordered pair involve the transition from a theory that is committed to the view that ordered pairs exist in addition to sets, to a new, streamlined theory that involves no such assumption. What usual eliminativists and their opponents are debating is what it means to give up this assumption. The eliminativists think it means denying that ordered pairs exist. Their opponents think it means discovering that ordered pairs are identical with sets. Quine's attitude is that insofar as this disagreement is seen as something else than a matter of rhetorical preference, it can and should be discarded.

This attitude is a manifestation of Quine's extensionalism. If he is right, the debate between the standard eliminativist and his opponent cannot achieve its apparent significance without recourse to intensional, essentialist discourse. In effect, the usual sort of eliminativist is arguing that being distinct from sets is such a central feature of ordered pairs (or of the concept of ordered pair, or of the meaning of the term 'ordered pair') that to give up the assumption that ordered pairs exist in addition to sets is, *a fortiori*, to give up the assumption that ordered pairs exist. By contrast, his opponent is arguing that being distinct from sets is not such a central feature of ordered pairs, and, hence, that someone who gives up the assumption that ordered pairs exist in addition to sets does not have to deny that ordered pairs exist. According to Quine, and *pace* both the standard eliminativist and his opponent, the most reasonable and fruitful thing we can do is stop worrying about the essence of ordered pairs, or about the concept of ordered pair, or about the meaning of the term 'ordered pair.' Quine regards it as a virtue of successful explication that it shows how laying such worries aside can be conducive to scientific progress and clarity.

III Explicating the Mental

Quine's conception of physicalist explications of the mental is in all relevant aspects similar to his conception of set-theoretical explications of the notion of ordered pair. There are of course differences between the two cases, differences that Quine is aware of. One such difference is that present-day neuroscience does not provide the material needed for a detailed physicalist explication that identifies each mental state with a neurologically specified state of the body. There are, in this sense, no Wiener and Kuratowski of neurology. According to Quine, however, this should not stop us from explicating away the dualist view that there exists a separate sphere of mental entities. As I have already quoted him saying, '[I]ack of detailed physiological explanation of [mental] states is scarcely an objection to acknowledging them as states of human bodies.' Let us have a closer look at why Quine thinks a detailed physiological account is not required here.

Quine's strategy for taking us from a dualist theory to a theory which posits no non-physical mental entities has one limitation: it is applicable only to forms of dualism that do not assume disembodied minds. In other words, for Quine's strategy to be applicable, the dualist theory has to admit that for each particular mental state there is a corresponding bodily state. The relevant notion of correspondence is purely extensional: what has to be admitted is just that each mental state is coincident with some state of the body. And it is not required that we are able to give a detailed physiological specification of the bodily state. Indeed, it suffices if the

bodily state is characterized as the bodily state of *having* or *accompanying* a mind to which the relevant mentalistic predicates apply. For example, the bodily state that corresponds to feeling pain can be characterized simply as: the bodily state of having a mind that is in pain. If such correspondence between mental states and bodily states is granted, Quine thinks we can just ‘transfer the mentalistic predicates to the body, bypassing the purported mental substance’ (Quine, 1995, 85). Instead of accounting for John’s feeling pain in terms of his having a non-corporeal mind that is in a state of pain, we drop the reference to the non-corporeal mind and interpret the predicate ‘is feeling pain’ as being applicable directly to John’s body. Analogously, we treat ‘John’s mind is thinking about Fermat’s Last Theorem’ as replaceable by ‘John’s body is thinking about Fermat’s Last Theorem.’ In this sort of way, we eliminate all references to distinct mental entities. Mental predicates will of course remain, but they will now be taken to apply only to physical things. The supposed need to posit separate mental entities is thereby shown to be an illusion. (See also Quine, 1976, 243, and Quine, 1981, 18-19.)

Quine calls the resulting monism ‘effortless’ — ‘form without substance’ (Quine, 1995, 86). A more substantial explication would involve a detailed physiological account of the mechanisms of mental phenomena. In the case of sensations, such as pain, Quine thinks the future prospects for substantial explication are good: ‘neurologists no doubt understand its neural mechanism to a considerable degree, and pain can be identified with that’ (Quine, 1995, 86). Such identification would constitute an interpretation of all pain-talk as talk about the relevant neural mechanism. When it comes to the propositional attitudes, however, Quine admits that finding neural or other physiological or behavioral *ersatz* seems much less plausible: ‘The notion of an exhaustive class of states each of which qualifies as thinking about Fermat’s Last Theorem, and each of which is specifiable in purely physiological terms, seems discouragingly unrealistic’ (Quine, 1995, 87). Quine concedes that ‘we must perhaps acquiesce in the psychophysical dualism of predicates, though clinging to our effortless monism of substance’ (Quine, 1995, 87). It is sometimes suggested that Quine is here adopting a position first developed by Donald Davidson (see, for example, Orenstein, 2002, 168). What is today called anomalous monism, however, is a conception Quine has been advocating at least since 1954 (see, for example, ‘The Scope and Language of Science,’ reprinted in Quine, 1976, 228-45).³

3 Thanks to an anonymous referee for emphasizing Quine’s early advocacy of anomalous monism.

As in the case of ordered pairs, Quine holds that the choice between eliminativist and identity theorist construals of physicalist explications of the mental is only a matter of rhetorical taste and convenience. He agrees with both eliminativists and identity theorists that a physicalist explication takes us from an old, dualist theory according to which there are non-physical mental entities to a theory that does not postulate such a separate sphere of mental phenomena. What eliminativists and identity theorists disagree about, however, is whether this transition involves a repudiation of mental entities. Eliminativists presume that dualism is integral to the very idea of the mental, and, thus, that abandoning dualism commits us to the claim that mental entities do not exist. Identity theorists, by contrast, argue that being mental is compatible with being physical, and, indeed, that abandoning dualism in favor of physicalism means coming to understand better what mental entities really are. Quine thinks this controversy lacks substance, the only difference being that the identity theorist's way of talking sounds less harsh. 'The reduction of the mental to the physical ... can be characterized in either of two ways: as *explaining* or as explaining *away*. There is no difference, but the first phrasing has a gentler ring' (Quine, 1995, 86).

There might seem to be no room for this sort of relaxed attitude. One is tempted to object that there must be a fact of the matter as to whether the entity that dualists purport to refer to when they use an expression such as 'John's mind' exists. In Quine's view, however, such an objection mistakenly takes it for granted that when the dualist uses an expression such as 'John's mind,' what constitutes the identity of the entity in question has somehow already been determined. Quine's point is that such determination will have to be relative to a proposed translation of the dualist's verbal behavior. Such relativity is precisely what those who see a substantial difference between eliminativism and the identity theory are denying. Standard eliminativists and identity theorists both take it for granted that the relevant identity criteria are somehow inherent in the old, dualist way of talking, and are therefore settled antecedently to *our* decision to interpret the dualist idiom in a particular fashion. According to the usual sort of eliminativist, those antecedently determined identity criteria are such that, given a materialist world-view, the expression 'John's mind' lacks reference. According to the identity theorist, the criteria are such that, given a materialist world-view, 'John's mind' refers to a physical entity.

Quine is not denying the purely logical point that John's mind must either exist or not exist. His claim is just that it is only against the background of an agreement about the relevant criteria of identity that questions about the existence of John's mind can be given a determinate answer, and that, in the case at hand, there are no objective, extensionally characterizable facts by reference to which it is possible to resolve the

disagreement between the eliminativist and the identity theorist as to what the adequate criteria of identity are. The eliminativist insists that the relevant criteria of identity are such that they preclude an identification of John's mind with anything corporeal. The identity theorist, on the other hand, insists that John's mind is identical with John's body. He shares the eliminativist's materialism, but argues that the expression 'John's mind' does refer, and that the dualist is just mistaken, in an ordinary, factual sense, about the corporeal nature of John's mind. The problem is that, just as in the parallel case of ordered pairs, neither of these standpoints seem possible to substantiate without recourse to intensional discourse. In order to maintain the view that there is a substantial opposition between eliminativism and the identity theory, the eliminativist will, in effect, have to argue that mental entities are somehow essentially non-physical, or that the 'concept' of mental entities is such that talking of mental entities as physical somehow involves transgressing conceptual boundaries. Conversely, the identity theorist will have to presume that distinctness from physical entities is *not* essential to the mental, or that he can assent to 'John's mind is identical with John's body' without violating any conceptual restrictions. According to Quine, this need to invoke non-extensional discourse shows that the debate is without objective foundation. What may look like a genuine, scientific issue is in fact a rhetorical surface phenomenon. The choice between eliminativism and the identity theory depends only on which way of talking we like the best.

IV Conclusion

'Extensionality,' says Quine, 'is much of the glory of predicate logic, and it is much of the glory of any science that can be grammatically embedded in predicate logic. I find extensionality necessary, indeed, though not sufficient, for my full understanding of a theory' (Quine, 1995, 90-1). A quick way of summarizing what I have done in this paper is to say that I have clarified the interconnection between Quine's extensionalism and his conception of explication as philosophical method.

One aspect of this interconnection is that Quine regards the achievement of extensional theory as a central goal of explication procedures. He thinks science, properly regimented, is extensional, and explication is crucial to the achievement of such regimentation. Physicalist explications of mentalistic language may at first seem unsuccessful in this respect, since Quine admits that finding substitutes for the idioms of propositional attitude is hopeless. It should be noticed, however, that he is nevertheless convinced that we can give such idioms an innocuous extensional reading, at least in *de dicto* contexts, by treating believing, intending, and so forth, as relations between people and sentences (Quine, 1992, 67-73 and 1995, 91-8). If such extensional treatment were not possible, he

would hardly be prepared to count the social sciences, in which the idioms of propositional attitude are indispensable, as genuinely scientific.

But Quine's extensionalism manifests itself not only in his view that extensional theory is the goal of explication procedures. As I have emphasized throughout this paper, it also governs his conception of how, precisely, the ontological import of explication procedures is to be understood. According to Quine, Wiener's and Kuratowski's set-theoretical explications of the notion of ordered pair are both fine. To think that one of them does better justice to the notion of ordered pair than the other, or that none of them can do adequate justice to the notion of ordered pair, is to think, mistakenly, that there is something such explications fail to capture even if they manage to do full justice to the relevant uses of terms of the form ' $\langle x, y \rangle$.' A similar mistake is made by those who think there is a substantial difference between eliminativist and identity theoretical construals of how explication works. To claim that there is such a substantial difference is, in effect, to presume that an explication procedure is answerable to something else than extensionally characterizable features of the employment of the terms which get explicated. It is to presume that an explication procedure is answerable to something beyond such facts, something whose adequate characterization requires intensional language: essences, meanings, concepts, or whatever.

At least within the philosophy of mind, eliminativists and identity theorists usually proud themselves of being naturalists. If Quine is right, this self-understanding is in conflict with the very idea that eliminativism and the identity theory are rival conceptions. For there is no way to construe the alleged dispute as scientific, given the worldview of properly regimented contemporary science. Eliminativists and identity theorists may believe that, in criticizing each other, they are doing the science of the mind. Quine disagrees. To our best knowledge, the question they are debating has no scientific substance at all.⁴

Received July 2004

Revised December 2004

Revised June 2005

Revised November 2005

4 Earlier versions of this paper were presented at the University of Helsinki, the University of Oslo, Uppsala University and Stockholm University. I thank the audiences for thoughtful discussions. I also want to thank an anonymous referee of this journal for very useful and penetrating criticisms. Work on the paper was financed by the Swedish Research Council and the Faculty of Arts at Uppsala University.

References

- Benacerraf, P. 1965. 'What Numbers Could Not Be,' *Philosophical Review* 74 47-73.
- Dipert, R.R. 1982. 'Set-Theoretical Representations of Ordered Pairs and Their Adequacy for the Logic of Relations,' *Canadian Journal of Philosophy* 12 353-74.
- Gibson, R.F. 2004a. 'Quine's Behaviorism cum Empiricism,' in Gibson 2004b.
- _____, ed. 2004b. *The Cambridge Companion to Quine*. Cambridge: Cambridge University Press.
- Hahn, L.E. and P.A. Schilpp, eds. 1986. *The Philosophy of W.V. Quine*. La Salle, IL: Open Court.
- Hylton, P. 2004. 'Quine on Reference and Ontology,' in Gibson, 2004b.
- Kanamori, A. 2003. 'The Empty Set, the Singleton, and the Ordered Pair,' *The Bulletin of Symbolic Logic* 9 273-98.
- Orenstein, A. 2002. *W.V. Quine*. Princeton: Princeton University Press.
- Quine, W.V. 1960. *Word and Object*. Cambridge, MA: MIT Press.
- _____. 1975. 'Mind and Verbal Dispositions,' in *Mind and Language*, S. Guttenplan, ed. Oxford: Clarendon Press.
- _____. 1976. *The Ways of Paradox and Other Essays*. Revised and Enlarged Edition. Cambridge, MA: Harvard University Press.
- _____. 1981. *Theories and Things*. Cambridge, MA: Harvard University Press.
- _____. 1986. 'Reply to Paul A. Roth,' in Hahn and Schilpp, 1986.
- _____. 1992. *Pursuit of Truth*. Revised edition. Cambridge, MA: Harvard University Press.
- _____. 1995. *From Stimulus to Science*. Cambridge, MA: Harvard University Press.
- Russell, B. 1903. *Principles of Mathematics*. New York: Norton.
- Stemmer, N. 2001. 'The Mind-Body Problem and Quine's Repudiation Theory,' *Behavior and Philosophy* 29 187-202.
- Wiener, N. 1914. 'A Simplification of the Logic of Relations,' *Proceedings of the Cambridge Philosophical Society* 17 387-90.