Quine on Indeterminacy

Forthcoming in Ernest Lepore and Barry C. Smith (eds) *The Oxford Handbook of Philosophy of Language*, (Oxford: Oxford University Press) with whom copyright resides.

W.V. Quine's thesis of the indeterminacy of translation is the theory which launched a thousand doctorates. During the 1970s it sometimes seemed to be as firmly entrenched a dogma among North American philosophers as the existence of God was among medieval theologians. Although now subject to much more by way of critical appraisal, Quine's work is still, rightly, at the forefront of contemporary philosophy of language. Moreover though propounded and defended by the doyen of analytical philosophy, as hard-nosed a logician as one can find, Quine's questioning of the determinacy of meaning is of interest to a much wider audience than logicians. Indeed the idea of indeterminacy of meaning has more than a whiff of smoke-filled cafés on the banks of the Seine about it, though Quine's arguments for indeterminacy belong firmly to the tradition of logical empiricism.

It is very tempting, of course, to apply a little reflexivity and deny that there is any determinate thesis of indeterminacy of translation; to charge Quine with championing a doctrine which has no clear meaning, or which is hopelessly ambiguous. Such a charge is, I will argue in §I false. His meaning is fairly clear and there is widespread agreement on what the thesis amounts to. In the second section I will look at Quine's 'argument from below', for indeterminacy, in §III at the 'argument from above', with concluding remarks in §IV.

§I

The *locus classicus* for the exposition of Quine's thesis of indeterminacy of translation is Chapter Two of *Word and Object* (Quine, 1960). Quine starts with an 'uncritical' presentation of the doctrine:

> two men could be just alike in all their dispositions to verbal behavior under all possible sensory stimulations, and yet the meanings or ideas expressed in their identically triggered and identically sounded utterances could diverge radically, for the two men, in a wide range of cases. (Quine, 1960, p. 26.)

(Women, of course, were language-less, in the early 1960s.) However he rejects this version as meaningless:

> a distinction of meaning unreflected in the totality of dispositions to verbal behavior is a distinction without a difference. (ibid.)

This makes it look as if some form of behaviourism is a background presupposition of Quine's argument and he does say:

> We are concerned here with language as the complex of present dispositions to verbal behavior (op. cit, p. 27; see also 1987, p. 5).

Quine certainly eschews 'mentalism', if we define this as a rejection of the supervenience of semantics on behaviour. For 'no distinction of meaning without a difference in behaviour', just is, in slogan format, the supervenience of semantics on behaviour. Quine moves, therefore, to a less mentalistic formulation of the indeterminacy thesis:

> the infinite totality of sentences of any given speaker's language can be so permuted, or mapped onto itself, that (*a*) the totality of the speaker's dispositions to verbal behavior remains invariant, and yet (*b*) the mapping is no mere correlation of sentences with *equivalent* sentences, in any plausible sense of equivalence however loose. (ibid.)

What could the plausible sense of equivalence be? A third formulation is introduced to help clarify:

> The same point can be put less abstractly and more realistically by switching to translation. ... [M]anuals for translating one language into another can be set up in divergent ways, all compatible with the totality of dispositions, yet incompatible with one another. (ibid.)

The discussion then focuses on the thought experiment of radical *translation*, of the predicament of a linguist faced with a community speaking a language which has no discernible affinities with a known language (a bit like Aberdonians, but even more so). Quine introduces his famous example of 'Gavagai':

> A rabbit scurries by, the native says 'Gavagai', and the linguist notes down the sentence 'Rabbit' ... as tentative translation (op. cit. p. 29).

This might make it look as if the indeterminacy thesis is concerned solely with *translation* between languages, perhaps only with translation of alien, putatively incommensurable, cultures; or that it is an epistemological thesis, expressing sceptical doubts as to whether we can ever know what others mean, at least if they speak a radically different language.

This would be a grave mistake. Quine's second formulation, in terms of permutations of one's own language, is the most fundamental and perspicuous one. His thesis is not an epistemological one but a metaphysical one and it concerns an indeterminacy in the meaning of the expressions of one's own language–

> On deeper reflection, radical translation begins at home. (Quine 1969*a*, p. 46; see also 1960, p. 78.)

Indeterminacy of translation is merely a corollary of the main thesis, albeit one which is pedagogically useful. The radical translation thought experiment helps one to bracket mentalistic assumptions and focus on the purely behavioural data which are available to linguist and language learner alike.

In what sense, then, do the permutations map sentences to non-equivalent sentences whilst leaving behavioural dispositions untouched? If Quine is indeed assuming some form of behaviourism, this is puzzling. Say that two sentences of a speaker's language are *behaviourally equivalent* iff the totality of the speaker's verbal dispositions towards the one is the same as that towards the other (spelling this out in detail has complications which we will pass over). Then Quine in the above quotation seems to be saying that there are behaviourally equivalent sentences which are nonetheless non-equivalent in some plausible sense. What sense can this be?

Could it be that they differ objectively in meaning, though they are behaviourally equivalent? A mentalist could say this, could take the thesis to be a rejection of the supervenience of meaning on verbal behaviour. But Quine cannot say this since for him any distinction in meaning must be reflected in a distinction in dispositions to verbal behaviour. On the other hand, to say only that they are syntactically distinct sentences is merely to affirm the existence of synonyms which Quine, by the time of *Word and Object,* (Quine, 1960) sees is fairly platitudinous.[1] Similarly dismissed by Quine as platitudinous are the theses that translation is often rough, there being no precise synonym, and that many sentences are vague (see again 1960, pp. 41, 73-4).

Quine's indeterminacy thesis is far more radical than this. One more radical claim which might be taken to interpret the thesis is the assertion that two behaviourally equivalent sentences can be non-equivalent in the sense of *intuitively non-synonymous*. That is, they can be objectively alike in meaning yet we think they differ; our beliefs about synonymy are fallible (1960, pp. 36, 63). This seems plausible for a behaviourist, though it goes against what one might call an extreme Cartesian view of the mind. On the latter view, the mind, including the meanings our mind gives to words, is transparent to us so that we have infallible, privileged access to all our mental states and can thus tell whether two words express, on our lips, the same idea or not.[2]

---

[1] In 'Two Dogmas of Empiricism' (Quine, 1951) Quine notoriously expressed a strong scepticism about the notion of synonymy. But later, in *Word and Object* (1960), he introduces a notion of 'stimulus synonymy' which, he says, is quite close to our intuitive notion of synonymy in the case of highly observational sentences. Two syntactically distinct sentences can, for Quine, be stimulus synonymous for a given individual (and for a linguistic community, if thus synonymous for each speaker in the community).

[2] Plausible though the anti-Cartesian view is, it is rejected by Wittgenstein, (1922, 4.243) and, following Frege, Dummett (1981, p. 95).

However the more of our intuitions Quine holds to be erroneous, the more radical (and less plausible) his position becomes. In fact the following version of indeterminacy:

> countless native sentences admitting no independent check …may be expected to receive radically unlike and incompatible renderings under the two systems. (1960, p. 72).

illustrates just how radical Quine's doctrine is. It embodies the thesis Quine most often has in mind when arguing for indeterminacy of meaning:– two sentences can be behaviourally equivalent yet necessarily distinct in truth value– the one is true if and only if the other is false. A variant of this thesis applied to names is the thesis of *ontological relativity* or *inscrutability of reference*:[3] two names can be behaviourally equivalent and yet stand for different objects, two predicates can be behaviourally equivalent yet true of different things.

But how can this possibly be? If sentence p is behaviourally equivalent to sentence q then surely, for the behaviourist, p means the same as q. Yet if they are incompatible in the above sense, we have P iff not Q, where P and Q are the sentences named by p and q respectively. But the following rule R is surely constitutive of the notions of meaning and truth:

$$\frac{p \text{ means the same as } q}{p \text{ is true iff } q \text{ is true.}}$$

From the premiss that two sentences mean the same we can conclude that the one is true iff the other is (relative to a background context which removes any ambiguity and fixes reference for any context-relative terms). As Tarski noted, 'p is true iff P' and 'q is true iff Q' are constitutive of the concept of truth. Putting all these things together (using the symmetry of 'iff') we derive, from the premises that p means the same as q and that P iff not Q:

Q iff q is true [Tarski]; iff p is true (R); iff P [Tarski]; iff not Q

And the transitivity of 'iff' (A iff B and B iff C entails A iff C) gives us Q iff not Q which leads, in standard logic, straight to contradiction. More directly, from rule R we conclude that the one sentence is true iff the other is whereas Quine maintains that the one is true iff the other is false; and these two claims are surely contradictory.

Similarly given 'the referent of $\langle t \rangle = t$' (another disquotational Tarskian truth about reference- here substitutions for parameter "$\langle t \rangle$" canonically name substitutions for parameter "t") plus the analogue R' of R (if two names mean the same, their referents are identical), the assumption that t means the same as u yields (with 'Ref($\langle t \rangle$)' standing for 'the referent of $\langle t \rangle$'):

---

[3] For Quine, these two terms are pretty much synonymous, at least for him (1992, pp. 51-2).

$$t = \text{Ref}(\langle t \rangle) \text{ [Tarski]}; = \text{Ref}(\langle u \rangle) \text{ } [R\text{'}]; = u \text{ [Tarski]}$$

so that $t=u$ even though, according to Quine we can have $t \neq u$. For example, if 'Bugs Gavagai' names your pet rabbit then Quine holds, as we shall see, that the singular term 'Bugs Gavagai's left ear' is identical, in point of objective meaning, with 'Bugs Gavagai' hence, by the above argument, Bugs Gavagai = Bug's Gavagai's left ear, even though we know they are distinct, one being a proper part of the other. A similar argument can be given for the indeterminacy of the extension of predicates: 'gavagai' can be interpreted as true of all and only rabbits or as true of all and only the undetached parts of rabbits, and so on.

Is Quine's position simply contradictory then? There is an explicit answer for reference, in the doctrine of ontological relativity, though one he seems to shy away from in the case of truth. For reference, Quine takes the above argument to be a *reductio ad absurdum* of some of the underlying assumptions, in particular, of the assumption that meaning and reference are absolute. Rather, insofar as sentences without empirical content can be said to have meaning at all, it is only relative to some one among many possible interpretations of the language into some background language (the *meta-language* in which we talk of the *object language* in question; it may be a completely different one or an extension of the object language):

> unless pretty firmly conditioned to sensory stimulation, a sentence S is meaningless except relative to its own theory; meaningless inter-theoretically. (1960, p. 24.)

> reference *is* nonsense, except relative to a coordinate system. ... What makes sense is to say not what the objects of theories are, absolutely speaking, but how one theory of objects is interpretable or reinterpretable in another. (1969*a*, pp. 48, 50.)

What Quine's view seems to comes down to, then, is this. If p and q have no empirical content, they do not have meanings but have one or other interpretation imposed or projected onto them, although never the same one in the same projection, if the equivalence P iff not Q holds. Similarly a term such as 'Gavagai' may have a determinate meaning, to do with rabbit features being present– 'It's rabbitish', as it were– but it does not segment occasions into rabbits, rather than undetached rabbit parts, and so forth. Such a segmentation is *our* projection onto a world which, in itself, does not come packaged into separate objects.

Relativity alone, however, will not save Quine from contradiction. The argument above will go through with the various notions relativised to an interpretation I: 'means$_I$ the same as', 'is true$_I$' and so on. The Tarskian schema then becomes p_I is true$_I$ iff P, where p_I names the

sentence which translates P according to interpretation I.[4] Further, rule R then becomes R^*:

$$\frac{p \text{ means}_I \text{ the same as } q}{p \text{ is true}_I \text{ iff } q \text{ is true}_I.}$$

In order to block the inference to:

Q iff q_I is true$_I$ [Tarski]; iff p_I is true$_I$ (R^*); iff P [Tarski]; iff not Q

granted that p and q are alike in all objective aspects of meaning, we need in addition to deny that if sentences have the same meaning then there is some interpretation I in which they mean$_I$ the same.

The claim that p means the same as q, then, Quine has to read as something like:

> There is a sentence r such that there is an interpretation I in which p is interpreted by r and also an interpretation I^* such that q is interpreted by r in I^*; but it need not be the case that $I = I^*$.

If, therefore, we step back from our object language at time t and talk about it at $t+1$ in what is in effect a metalanguage, we can say that there is an interpretation I of 'rabbit' in which it refers to, or is true of, all and only the rabbits. But there is also an equally good interpretation I^* in which 'rabbit' is true of all and only undetached proper parts of rabbits. There is, though, no interpretation I^{**} in which 'rabbit' is true of all and only the rabbits *and* all and only the undetached proper parts of rabbits, that is indeed absurd.

The upshot of Quine's indeterminacy thesis is the relativity of reference and, if he is consistent, of truth. Hence Quine must be placed firmly in the camp of the anti-realists. True, Quine is prepared to assent to current scientific theory and hence to affirm its truth, since he accepts p is true iff P; he has a deflationary, disquotational view of truth (Quine, 1960, pp. 24-5). But he is wrong in thinking this makes him a realist, in anything like the traditional sense. An instrumentalist who accepts current scientific theory (as instrumentalists generally did) and is prepared to accept the legitimacy of a disquotational conception of truth (as most were, post Tarski) would not disagree in the least with Quine on truth. The realist, however, believes that our theoretical conjectures are determinately and absolutely true or determinately and absolutely false, whether or not we have any means of finding out which. But for Quine, once we reach the theoretical realms where meaning, he claims, is indeterminate, truth and reference are relative, not absolute.

Just as a diagram can be read as a rabbit looked at one way, a duck another, though the objective figure is the same, so a theory can be true interpreted one way, false interpreted another though the objective facts

---

[4] The trivial homophonic interpretation H in which each sentence translates itself will always be admissible hence so too will the disquotational schema p is true$_H$ iff P; in this transparent case, we can drop the subscript on 'true'.

(that is, for an empiricist such as Quine, the *empirical* facts) and the meaning of the theory remain the same. Hence there is no sense to the claim that it is one or the other independently of us: as regards the theoretical component of a theory whose empirical consequences are true, it is *we*, by our way of reading that theoretical component, who make it true$_I$ or false$_{I*}$. The objective world is, for Quine, an ensemble of occasions possessing, as wholes, objective, observable features. But the segmentation of such occasions into distinct objects possessing non-observable, underlying micro-structures is a human construction answering to no corresponding objective reality; it is a colourisation of an intrinsically monochrome scene, as it were, and one which could equally well be effected in a number of different ways (though perhaps not by us). Such a view embodies a relativistic anti-realism which Quine himself shrinks from, at least as regards the notion of truth. (Quine, 1960, pp. 24-5. 1975, p. 327-8.)

§II

Indeterminacy, then, is a bold thesis with far-reaching metaphysical consequences. Does Quine give us good reason for thinking it true? He has two main arguments for indeterminacy of meaning, which he terms the argument from below and the argument from above, respectively (1970, p. 183). The argument from below is his argument for the relativity of reference. It hinges on the assumption that the only meaning a sentence can have, on its own, is empirical meaning which he characterises in terms of his concept of *stimulus meaning*. A sentence's stimulus meaning is a pair consisting of the affirmative stimulus meaning together with the negative stimulus meaning. The former is the set of stimulations which would prompt assent to the sentence on being queried on it, the latter the set of stimulations which would prompt dissent. In *Word and Object*, Quine treated stimulations as physical events or 'patterns' just outside the sensory organs (1960, p. 31). Translation then should match sentences with approximately identical stimulus meanings: the natives would assent to 'Gavagai' on being prompted with pretty much the same stimulations as we would assent to 'Rabbit', likewise for dissent.

Quine's empirical meanings, then, are not distal objects:

> It is important to think of what prompts the native's assent to "Gavagai?" as stimulations and not rabbits. Stimulation can remain the same though the rabbit be supplanted by a counterfeit. (1960 p, 31.)

and similarly a rabbit may fail to stimulate assent because of poor lighting etc. Later on, he despairs of intersubjective stimulations; placing them just outside the sensory organs will not work because of differences in orientation and anatomy among different subjects (Quine, 1974, p. 24). He therefore re-defines stimulations as patterns of firings of sensory receptors (1992, pp. 2, 40) and accepts, because of the lack of homology of sensory nerve networks (Quine, 1974, p. 24, fn. 2), that there can be no

intersubjective stimulations. This is a major change in his position: translation of observation sentences is no longer based on objective sameness and difference of stimulus meanings but has become a much more hermeneutic matter of empathetic placing of oneself in the subject's shoes and figuring out what translation makes best sense from that perspective (Quine, 1992, p. 42).

Could Quine not have held firm to a naturalistic account of observational meaning? He acknowledges, in response to Davidson's suggestion of a more distal meaning in external physical objects:

> I could place the stimulus out where Davidson does without finessing any reification on the subject's part. But I am put off by the vagueness of shared situations. (1992, p. 42.)

If, though, we are prepared to accept a little vagueness, we could define occasions as chunks of spacetime centred on a linguistic subject, whilst refraining, for Quinean reasons, from segmenting them further into distinct objects or bodies of a more traditional sort. We can think of the occasions, then, as what prompt assent or dissent for that particular speaker.[5] Granted this, define

> a is stimulus-equivalent with b just in case for any occasion O which prompts a speaker to a verdict, a is a part of O iff b is.

Here we need to a charitable reading of the vagueness of 'occasion'. I take it that when an occasion prompts assent to 'rabbit', then though one half, say, of the rabbit is not visible, it is still part of the prompting occasion. At least that will be so if the speaker would have been surprised to discover the rabbit was a cardboard cut-out or a robot whose hidden side reveals all the electronic innards. Similarly the sighting merely of a rabbit's ear with the rest of the body obscured may prompt assent to 'rabbit' but the whole rabbit is again part of the prompting occasion.

Next define an admissible permutation of an interpretation I to be a permutation of the domain of individuals of the interpretation which maps individuals to stimulus-equivalent individuals. If p is any such permutation then from any interpretation I we can create a new interpretation I_p in which if α is the referent of singular term t in I then $p(\alpha)$ is its referent in I_p, if $\{x: \varphi x\}$ is the extension of predicate F in I then $\{x: \exists y \varphi y \ \& \ x = p(y)\}$ is its extension in I_p.[6] It is easy to show that I and I_p are materially equivalent, that is, S is true in I iff true in I_p, for all sentences S.[7] And it is evident (by induction on sentence complexity) that

---

[5] Or better, relative to a speaker with a given set of dispositions to respond to sensory input. Given different dispositions, the same occasion might elicit different responses, indeed the segmentation of the world into occasions might have to be done differently.

[6] It is routine to extend this to second-order languages and modal languages.

[7] Note that this will be true even if the variant interpretation does not change the interpretation of what Quine calls the 'apparatus of individuation', terms for identity, definite and indefinite articles, numerals and attributions of number– the number of Fs

any stimulation which would be expected to prompt assent (dissent) on interpretation I would be expected to prompt assent (dissent) on interpretation I_p:– call this stimulus equivalence. One criterion for two interpretations I and I^* to be equally good (call this equivalence of interpretations) is simply that they pass the two constraints of material and stimulus equivalence.

My reconstruction of the Quinean argument of *Word and Object* (1960, §12) is thus that it concludes that there is a multiplicity of equally good interpretations of the singular terms and predicates of our language, each pair related by an admissible permutation as above, but no pair of which can be combined into a single coherent interpretation. There is nothing in our verbal behaviour which could differentiate between an interpretation of 'gavagai' as segmenting occasions into [rabbits, against a background], on the one hand, versus [a grouping of undetached rabbit parts against a background] on the other.

A virtue of this reconstruction is that it answers a counter-argument of Gareth Evans (Evans, 1975). He contends that Quine considers only the simplest one-word contexts in which names and general terms might occur. In contexts such as 'white gavagai' we might have hard behavioural evidence that 'gavagai' segments occasions up in one particular fashion (assuming that their 'white' is stimulus synonymous with our 'white'). For instance, if natives dissent from 'white gavagai' in the presence of a largely brown rabbit with a white patch but assent in the presence of a mostly white rabbit, we can conclude 'gavagai' is not true of any and every undetached rabbit part.

Evans seems to me to be onto a genuinely explanatory account of the nature of at least some forms of predication. Nonetheless any admissible permutation function p which maps objects only onto stimulus-equivalent objects will get round Evans' point. For the extension of 'gavagai' in the variant interpretation I_p will be the set of all p-images of rabbits, so including perhaps some rabbit feet; the extension of 'white' will be all p-images of white things. Suppose p, for example, is the identity permutation save that it permutes our mostly brown rabbit, Bugs Gavagai, with its white front left foot, swapping the two round. So Bugs' foot satisfies 'gavagai' according to I_p but not 'white' hence I_p is entirely consonant with the speakers dissent from 'white gavagai' in the presence of that rabbit.

There are two directions of response to Quine's argument (as construed above). One could hold that there are more constraints on the equivalence of interpretations than the two constraints of material equivalence and stimulus equivalence. A mentalist, for example, may take the argument as showing not that reference is indeterminate but that

---

is n– and so on. Some of Quine's earlier arguments appealed to compensating re-interpretations of such terms, see 1960, p. 53, 1969*a,* pp. 32-3. See 1964, p. 215 for 'reduction' construed as material equivalence given by an effective mapping.

behavioural supervenience, or perhaps even supervenience on the physical, is false. Alternatively one may hold that there are even fewer constraints. Quine's argument from below, we shall see, abandons material equivalence, though only for theoretical terms. Both Davidson (1979, p. 229) and Putnam (1981, pp. 32-5, 217-8) in effect drop stimulus equivalence and retain only material equivalence, this generated by arbitrary permutations; and in this they are followed, to a great extent, by Quine himself in some of his later work.[8]

If material equivalence is the only constraint on equivalence of interpretations then indeed reference is wholly indeterminate, since any permutation will produce a materially equivalent one and so 'Hilary Putnam' is as correctly interpreted as referring to the Andromeda Galaxy as to Hilary Putnam. That this conclusion has been accepted is a tribute to the engaging tendency of philosophers to transform a reductio ad absurdum of a cherished assumption into a proof of that very absurdity. For the assumption that material equivalence is the sole criterion for equivalence of interpretations is vastly less plausible than the thesis that the Andromeda galaxy is nothing like as good a candidate for the referent of 'Hilary Putnam' as Hilary Putnam himself.

It will not do to respond that the terms featuring in any additional constraints could themselves be re-interpreted in different ways: 'stimulus-equivalence' could be re-interpreted so that the Andromeda Galaxy and Hilary Putnam are stimulus-equivalent.[9] Certainly if one assumes from the outset that no term has determinate reference one will be hard put to show that many terms have determinate reference. But we are engaged here in a sub-species of naturalised epistemology in which we are trying to explain a special type of knowledge: of meaning. Our task is to assume determinate reference and extension relations, e.g. between 'Hilary Putnam' and Hilary Putnam (and nothing else), 'stimulation' and stimulations and so forth and then go on to show how speakers could grasp a language with such reference relations. This is not, prima facie, an impossible feat so long as one does not deny the theorist the right to assume determinate reference relations at the outset. But it is certainly no trivial feat as Quine's original arguments, for a more moderate form of indeterminacy of reference, show. For these arguments suggest strongly that one is forced to give up either supervenience of semantics on the behavioural, perhaps indeed on the physical or else give up determinacy of reference. Quine, of course, chooses the latter course.

---

[8] See for example, Quine, 1995, pp. 71-3, Quine, 2000, pp. 419, 420. General permutations– 'proxy functions'– emerged in Quine, 1964.

[9] Davidson (1979, p. 237) and Putnam, (1978, p. 126 and 1981, p. 36) respond along those lines against which see Kirk, (1986, pp. 118-127).

§III

Quine's argument from above is an argument for the indeterminacy of meaning of theoretical sentences, an argument which he originally based on the *underdetermination of theory by evidence* (see, in particular, Quine, 1970 and 1975). This occurs when we have two theories T and T* which are empirically equivalent– that is if T entails an empirical sentence E so does T* and vice versa– yet incompatible (so that there is a theoretical, non-empirical sentence A such that T entails A but T* entails ~A).[10]

Trivial examples of underdetermination arise when we take a given theory and swap round two theoretical terms– e.g. swap 'electron' with 'molecule' in the axioms of the theory so we now end up saying that molecules are smaller than electrons, have negative charge etc. Quine focuses on more complex examples in which we know of no simple permutation of predicates which would turn T into a 'merely terminologically' different T*. Suggested examples of such theories include two versions of Newton's gravitational theory, one interpreting gravitation in terms of fields of force which exist at every point in space, the other in terms of action at a distance. These are clearly incompatible, yet observationally there would be no difference between the two. Another example is a theory T which says time is linear but cyclical with infinitely many exact repetitions of each epoch, whilst T* says there is only one epoch but the topology of time is circular. Or T posits a 'multiverse' comprising a vast plurality of mini-universes each expanding from a Big Bang but according to different parameters, only one set of which determines the mini-universe we are in and can observe. And so on.[11]

How does underdetermination lead to indeterminacy? Quine's argument is that since T and T* have the same empirical content, in addition to the trivial identity mapping translating P by itself for all members of T and T*, there will also be translations of T into T* and T* into T which preserve empirical content (do not map a sentence into one with distinct empirical content) and so are equally good, yet map

---

[10] More complexly one might add the requirement that T is empirically equivalent to T* only if, for all observation sentences E, the probability (or degree of confirmation or some such) of T given E equals that of T* given E; but I will stick with the simpler, purely deductive, definition which is to be found in Quine, 1970, p. 179 (expressed, equivalently, in terms of compatibility rather than entailment).

[11] In the last two cases, although *we* cannot make any crucial experiments distinguishing the two hypothesis, it may be that different sets of observations sentences, construed as abstract objects which need be grasped by no observer, are true in each case. This problem can be avoided if we image a multiverse in which all mini-universes bar our own last only a few nano-seconds or have tiny spatial dimensions. There might be theoretical reasons for positing such a multiverse, for example resolving the 'fine-tuning' problem of why the 'brute' parameters of the universe seem so exactly fitted for the development of stable complex molecules and so life.

sentences (such as the conjunction of the axioms of T) onto incompatible sentences (such as the conjunction of the axioms of T*).

Some objected to Quine that there was no more here than the usual scientific uncertainty:– in linguistics as in physics the theory outruns the evidence. Quine replied (e.g. at 1970, p 180, 1987, pp. 9-10) that the thesis is a stronger metaphysical one. Fix all the physical facts– choose either T or T* as the correct theory; there is still no unique correct interpretation and hence no determinate fact of the matter. A counter-reply to this is that perhaps the physical facts are not all the facts or, if this is so by definition of 'physics', perhaps future physics will differ from current physics in such a way as to render indeterminacy implausible. Quine would accept this possibility, seeing his philosophical views as part of science, not prior to and more fundamental than it, and so fallible like the rest of science.

There is a puzzle about Quine arguing to indeterminacy from underdetermination, however, since underdetermination is surely a highly realist thesis:– it says there can be two theories which are empirically indistinguishable. But if they can also be equally explanatory, how could we ever know which, if any, is correct? From the sceptical standpoint of the realist (of a certain type), the answer is we cannot know; but from an empiricist viewpoint, surely such theories would be indistinguishable and hence not two theories but one. Quine's later work reveals an increasing acceptance of this argument and the related idea that there is no significant under-determination of theory by evidence. He sways between *ecumenical* views on truth– any two empirically equivalent theories can be rendered compatible by terminological readjustments and incorporated into a wider whole– and a *sectarian* view according to which we should plump for one such and reject the other as meaningless.[12] The latter seems his more favoured view but, reflecting on the situation, he acknowledges we can flip from sectarian adherence to one theory to similarly ardent adherence to the other (1992, pp. 99-100; real sectarians do not behave like this!); so his reflective position seems hard to distinguish from the ecumenical one.[13]

If realism if wrong and there is no genuine underdetermination, what of the argument from above? It still goes through because the background assumptions used to derive indeterminacy from underdetermination– *holism* and *verificationism*– are sufficient to yield indeterminacy on their own (see Quine, 1969*b*, pp. 80-81). Verificationism

---

[12] See Quine, 1992, §42. One reason against accepting a combined theory in which one renders the two formally incompatible but empirically equivalent theories consistent by terminological readjustment is that the resultant theory will be 'bloated'. Quine favours theories which are more elegant and have less 'fat'– this is what rules out adding 'The Absolute is Lazy' to an empirically acceptable theory– so long as they entail the right observational consequences in a simple, acceptable fashion.

[13] A very useful charting of Quine's oscillations on this is matter is to be found in Gibson, 1988, Chapter Five.

is the empiricist view that the only literal meaning is empirical meaning, empirical content. Holism is the view that no theoretical sentence has empirical meaning or, more moderately, that most theoretical sentences do not, though some (the conjunction of the axioms and boundary conditions of an empirical theory, for instance) do. This more moderate holism is Quine's later view and is extremely plausible. Together they yield the characteristic Quinean views that synonymy is an empty relation amongst most theoretical sentences, since they have no meanings of their own to relate, and that empirically equivalent theories T and T* can be correctly inter-translated by mappings which send a sentence P to an incompatible, intuitively non-synonymous, sentence Q.

For example, let N be the conjunction of Newton's three laws of motion plus his inverse square gravity law. N has no empirical content, it entails no empirical hypotheses independently of further auxiliary hypotheses and boundary conditions. Neither does its negation ~N: without further hypotheses we have no predictions as to which objects are violating the laws. So both have the same (null) meaning according to Quine! We did not need to appeal to underdetermination here.

Indeterminacy from above, and the resultant rejection of realism, follow from the highly plausible holist thesis, if verificationism is true. But verificationism is obviously incompatible with realism and fairly easily dismissed by the realist. Empirical content is not all there is to meaning even on fairly behaviouristic premises. Where O is an observational sentence then (O & O) and (O ∨ O) are both logically equivalent hence identical in empirical content. But they have different syntactic structures, one being constructed using &, the other ∨. Moreover Quine himself, in his verdict matrix theory of the connectives, provides a behaviouristic account of how such connectives have meaning and how they differ in meaning (1960, §13). If, then, we require for the synonymy of two sentences not only sameness of empirical content but also that operators with the same meaning occur at the same nodes in the structure of the sentences then we can fairly easily show how empirically equivalent sentences have different meanings. Quine's argument from above, then, is not incontrovertible, though it is much more difficult to show how empirically equivalent sentences could differ in truth-conditions[14] as well as in meaning.[15]

§IV

Quine's moderate holism explains his tolerant attitude, despite his behaviouristic outlook, to the failure of behaviourist reductions of key

---

[14] As (O&O) and (O ∨ O) do not.

[15] The idea here is close to Carnap's notion of 'intentional isomorphism'. Quine, 1960, §42, criticises the use of such an idea, see especially the paragraph pp. 205-6, but the criticism is arguably question-begging in that it assumes that the indeterminacy thesis is true.

notions such as assent, dissent and, in his philosophy of perception, perceptual similarity. As in any other science, theoretical concepts cannot be defined in more empirical (for instance, behavioural) terms– such is moderate holism. Nor can we always expect illuminating definitions in terms of other theoretical notions.

But if that is so, why not accept the notions of analyticity or synonymy, even though they are not definable or behaviourally reducible, and with them a whole host of other notions, such as concept, proposition, belief, all individuated in a fine-grained way? Quine might appeal to the supervenience of the psychological on the physical and claim to have shown that supervenience fails for all these notions; for example, synonymy relations are not fixed even when all physical facts are fixed. But he showed this only granted the very identification of meaning with empirical content which is being challenged as resting on reductionist premisses. Quine's ultimate answer here seems to be that he can see no point to the introduction of these notions (see Quine, 1986, p. 207).

Certainly, those traditional notions of analyticity and synonymy can play no role in the establishing of Quine's highly empiricist, anti-realist metaphysics, since they can be used to undermine it. This means that his radical empiricism, which he thinks of as breaking free from older metaphysics by being continuous with modern science and making no prior philosophical presuppositions, in fact itself presupposes, rather than establishes, the correctness of the old empiricist verificationist metaphysics. Nonetheless both the argument from above and from below yield important illumination even for those who do not accept that meaning is indeterminate. For they show that if synonymy is as fine-grained as we are pre-theoretically inclined to think it is, we must modify or abandon some central tenets of a naturalistic empiricism which many find highly attractive.

Alan Weir
School of Philosophical Studies,
Queen's University, Belfast,
BELFAST BT7 1NN
a.weir@qub.ac.uk

REFERENCES

Works by Quine cited:

Quine, W.V. 1951: 'Two Dogmas of Empiricism', in *From a Logical Point of View*, (Cambridge: Harvard University Press), 2nd revised edition (1961), pp. 20-46.

1960: *Word and Object*, (Cambridge Mass.: MIT Press).

1964: 'Ontological Reduction and the World of Numbers', in W.V. Quine, *The Ways of Paradox and Other Essays*, (2nd Edition), (Cambridge, Mass.: Harvard University Press, 1976.)

1969 *Ontological Relativity and Other Essays*, (New York: Columbia University Press).

1969*a*: 'Ontological Relativity' in W.V. Quine, 1969, pp. 26-68.

1969*b*: 'Epistemology Naturalized', in W.V. Quine, 1969, pp. 69-90.

1970: 'On the Reasons for Indeterminacy of Translation', *The Journal of Philosophy*, **LXVI**, pp. 178-83.

1974: *The Roots of Reference*, (La Salle, Illinois: Open Court).

1975: 'On Empirically Equivalent Systems of the World', *Erkenntnis*, **9**, pp. 313-28.

1986: 'Reply to Hellman', in L.E. Hahn and P.A. Schilpp (eds.) *The Philosophy of W.V. Quine*, (La Salle, Illinois: Open Court) pp. 206-8.

1987: 'Indeterminacy of Translation Again', *The Journal of Philosophy*, **LXXXIV**, pp. 5-10.

1992: *Pursuit of Truth*, (Cambridge, Mass.: Harvard University Press).

1995: *From Stimulus to Science*, (Cambridge, Mass.: Harvard University Press).

2000: 'Responses' in P. Kotatko and A. Orenstein (eds.), *Knowledge, Language and Logic*: *Questions for Quine*, (Dordrecht: Kluwer).

Useful Monographs on Quine:

Gibson, Roger, 1988: *Enlightened Empiricism*, (Tampa: University of South Florida Press).

Hookway, Christopher, 1988: *Quine*, (Cambridge: Polity Press).

Kirk, Robert, 1986: *Translation Determined*, (Oxford: Clarendon Press).

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

See also:

Miller, Alexander, 1998: *Philosophy of Language* (London: UCL Press), Chapter Four.

Other Works Cited:

Davidson, Donald, 1979: 'Inscrutability of Reference', in Donald Davidson, *Inquiries into Truth and Interpretation*, (Oxford: Clarendon Press, 1984, pp. 227-41).

Dummett, Michael, 1981: *Frege: Philosophy of Language*, (Second Edition), (London: Duckworth).

Evans, Gareth, 1975: 'Identity and Predication', *The Journal of Philosophy*, **LXXII**, pp. 343-363.

Putnam, Hilary, 1978: *Meaning and the Moral Sciences*, (London: Routledge, Kegan and Paul.)

1981: *Reason, Truth and History*, (Cambridge: Cambridge University Press).

Wittgenstein, Ludwig, 1922: *Tractatus Logico-Philosophicus*, (London: Routledge & Kegan Paul).