

Modality & Other Matters: An Interview with Timothy Williamson

Timothy Williamson (**TW**) has been the Wykeham Professor of Logic at Oxford since 2000. His books include *Vagueness* (Routledge, 1994), *Knowledge and its Limits* (Oxford, 2000), and *The Philosophy of Philosophy* (Blackwell, 2007). He has also published numerous articles on epistemology, philosophy of language, philosophy of logic and metaphysics. He taught at Trinity College Dublin, from 1980–1988.

The interviewer, Paal Antonsen (**PA**), is a philosophy student at Trinity College Dublin.

PA: In this interview I want to get an overview of some of your philosophical perspective. I will ask three main questions: the first about the difference between philosophical and non-philosophical knowledge, the second concerns the epistemology of modality, and the third is on the emerging metaphysical picture.

I

PA: In your recent book, *The Philosophy of Philosophy* (3), it is said to be an important part of your view that a common assumption, which you call ‘philosophical exceptionalism’, is rejected. Could you explain what you mean by this term, and what the alternative position ‘philosophical anti-exceptionalism’ involves?

TW: By ‘philosophical exceptionalism’ I mean the view that philosophy is quite unlike other forms of intellectual inquiry, that there is something very special about philosophical inquiry. This was a

common view in a lot of 20th century of philosophy, held for example by the logical positivists and both the early and the later Wittgenstein. The idea was that the aim of philosophy was something like conceptual clarification rather than gaining knowledge of how things are. Sometimes it is said that philosophy is supposed to lead to understanding, but not to knowledge. There is a whole family of versions of this view. By 'philosophical anti-exceptionalism' I mean the negation of that view, that philosophical inquiry is a form of intellectual inquiry not so different from what goes on in other disciplines. This is not to say that it should be assimilated to any particular non-philosophical form of inquiry, such as physics, because there are wide variations among non-philosophical forms of inquiry. The differences between philosophy and other disciplines are not unlike the differences between physics, mathematics, history, economics, biology and linguistics, which are themselves very marked. In particular, I am suggesting that the aim of philosophy is, in the long run, to acquire some sort of knowledge of how things are. I also think that the alleged dichotomy between knowledge and understanding is a false one: you can't understand why something is so without knowing why it's so.

PA: A corollary of anti-exceptionalism is that there is no principal difference between philosophical and non-philosophical knowledge (and understanding). From this one may infer, as indeed Quine seems to have done, that the notion of 'metaphysical necessity' is simply spurious. If there is nothing special about philosophical knowledge, then there is no special knowledge of metaphysical modality involved either. How do you think that metaphysical necessity may be salvaged and how does it relate to non-metaphysical truths?

TW: I think the inference from the assumption that there is no very radical difference between philosophical and non-philosophical knowledge to the spuriousness of metaphysical necessity is fallacious. The notion of metaphysical necessity is not itself an epistemic idea.

It is simply an idea about how things could have been. Something is metaphysically necessary if it couldn't have been otherwise, in the most unrestricted sense. In other words, if something really is metaphysically necessary, it would have obtained no matter what had obtained. What I am suggesting is that the kinds of ways we have of acquiring non-philosophical knowledge are capable of being applied to give us philosophical knowledge, including knowledge of metaphysical necessity. The particular link I am making is between the sorts of methods we use to evaluate counterfactual conditionals outside philosophy and the notion of metaphysical necessity. I characterize metaphysical necessity as what would have been the case, whatever had been the case. That is, in effect, itself a generalized counterfactual conditional. As far as I can see, the methods we have for evaluating counterfactual conditionals in general are appropriate even in special cases like that of metaphysical necessity. Of course, to some extent they need to be refined for the special cases we are interested in when doing philosophy, but refining ways of acquiring knowledge does not mean switching to radically different ways of acquiring knowledge.

II

PA: Continuing this thought on the link between metaphysical modality and counterfactual conditionals takes us over to the issue concerning epistemology of modality. A crucial step in your formulation of philosophical methodology involves the logical equivalence between modal claims and counterfactual conditional claims. There are different ways of formulating it, but one way to explicate the equivalence is of the form **(1)**:

$$(1) \quad \Box A \leftrightarrow (\neg A \Box \rightarrow (A \& \neg A))$$

Informally, it reads as saying ‘**A** is necessary if and only if, if **A** weren’t the case a contradiction would’. Under standard interpretations of modality and counterfactual conditionals (1) is valid. However, you think that (1) also plays an epistemological role, namely that given such a logical equivalence one can reduce the modal epistemology to counterfactual epistemology. Could you explain why you think that knowledge of modality is a special case of knowledge of counterfactuals?

TW: I think some people have slightly misinterpreted what I was saying in the book as the idea that if two things are logically equivalent then it just follows that the way we know one is the same as the way we know the other. That isn’t strictly true. If we’re not aware of the logical equivalence we may use different methods. We may be aware of the logical equivalence in principle without using it in practice. It is certainly not as straightforward as that. It also doesn’t seem to be psychologically plausible that the way we evaluate modal claims about possibility and necessity is by actually articulating an equivalence of the kind that you mentioned, so that after evaluating the counterfactual claim we apply to the equivalence to make the corresponding modal judgement. The connection I see is a little bit subtler than that. It is something like this: suppose we manage to give a decent account of how it is we acquire knowledge of counterfactuals. This is going to involve postulating various sorts of cognitive mechanisms that will involve reasoning processes, the use of the imagination and so on. If we have such an account, the question arises, do we need to postulate any further cognitive capacities in order to explain our knowledge of modality? What I am in effect arguing in the book is that once we see what range of cognitive capacities is needed to explain our knowledge of counterfactuals, we see that those capacities are already sufficient to explain our knowledge of modality. That is not going by some inference that involves an equivalence like (1) as a premise. It is simply that the same capacities can be applied directly to the modal judgments, and applied in a way that we would expect given that the equivalence

(1) holds. In other words, the applications of these cognitive capacities will be the same for the left hand side as for the right hand side. That doesn't mean we are explicitly analysing modalities in terms of counterfactuals, but once we see that no further cognitive capacities need be postulated in order to explain our knowledge of modality, it becomes utterly implausible to suppose that nevertheless we are using some extra unnecessary cognitive capacities when we are evaluating claims of modality. In particular, whereas counterfactual conditionals have all sorts of practical applications, and it is quite easy to see why it would have been evolutionary valuable to develop capacities for evaluating counterfactuals, claims of metaphysical necessity and possibility are generally of little interest to anyone except philosophers. It would be very, very strange indeed if somehow evolution had given us these special cognitive capacities whose only application was to claims of interest to metaphysicians.

PA: Not only can you say that it seems sufficient to stipulate the ability to do counterfactuals to modality, it also seems that people who are arguing about particular modal claims fall back to counterfactual terms. In particular, I have in mind Kripke's preface to the second edition of *Naming and Necessity*. When challenged on the point of modal epistemology, he reverts back to talk about counterfactual scenarios. It seems that whenever people have to talk about specific cases, they have to revert back to talk about particular counterfactual situations. Also, as you were mentioning the role of evolution, both Dennett and Dawkins have pressed the importance of developing the capacity to evaluate counterfactual conditionals in order to predict future outcomes. Dennett also seems to think that evolution can explain the capacity for counterfactual conditionals, and thus provide a naturalized way of explaining the capacity for evaluating modal claims. Would you say that since people always revert back to counterfactual talk when explaining modality, that sort of goes to show your point as well?

TW: It's grist to my mill. One has to be a little bit careful, because the cognitive processes that are in fact employed, perhaps sub-personally, when people arrive at a judgment are not necessarily the same as the kinds of justification that people produce when they are pressed to do so. One thing that cognitive science has produced evidence for is that there is a considerable element of rationalization in the sort of justification we produce, partly because we often don't have access of an introspective kind to the processes we are actually using. One can't simply rely on what people say when they are pressed, because it may not correspond to the actual cognitive processes involved in the acquisition of the relevant beliefs. In this particular case however, there is as far as I can see no need or plausibility in postulating some alternative cognitive mechanism. The sort of reliance on counterfactuals that people do, as you say, typically make in such situation may indeed reflect the process by which they actually reached these judgments. But we need corroboration from several different directions to make the case convincing.

PA: It was mentioned that you want to sever a link between the epistemology and the metaphysics of modality. There is, however, a view that modality must in any case be more attached to epistemology. Some will be critical of the principle of bivalence as applied to counterfactual claims—the principle of bivalence being the view that every proposition is determinately either true or false. If one holds bivalence for modal claims then, by **(1)**, one must also hold a similar condition for counterfactual conditionals. But the denial of a counterfactual may be represented in two different ways. Hence, we have two versions of conditional bivalence.

- (2)** $(\mathbf{A}\Box \rightarrow \mathbf{B}) \vee \neg(\mathbf{A}\Box \rightarrow \mathbf{B})$
- (3)** $(\mathbf{A}\Box \rightarrow \mathbf{B}) \vee (\mathbf{A}\Box \rightarrow \neg\mathbf{B})$

The strong denial of a counterfactual conditional in **(2)** is a negation of the counterfactual conditional as a whole, whereas the weaker denial of a counterfactual conditional in **(3)** is a negation of the conse-

quent. In some logics, (2) and (3) collapse, but you probably don't want that. Could you first explain how you see a difference in epistemology between affirming and denying a counterfactual conditional?

TW: This was something I wrote about in an article on bivalence and subjunctive conditionals, while here at Trinity. Let me make a few qualifications. One is that I would consider the principle of bivalence just to be the principle that every proposition is either true or false. I wouldn't add the 'determinately' qualification, which I think you have taken from Dummett. In my view, that just confuses the issue, but I don't think it's so crucial here. My views about the logic and semantics of counterfactuals are not primarily based on epistemological considerations, because, unlike Dummett, I am a realist. I don't think the meaning of counterfactuals is to be articulated in terms of the conditions under which we can know or assert them. I would rather evaluate these issues in directly logical, semantic and metaphysical terms. One point that I'd like to make is that it seems quite clear what the falsity of a counterfactual conditional is equivalent to; it is equivalent to the truth of its negation. The negation of a counterfactual conditional has to be represented as a negation of the whole thing. There is a further issue as to whether negating a counterfactual conditional is equivalent simply to negating its consequent. In other words, whether 'It is not the case that, if it had rained, the match would have been cancelled' is equivalent to 'If it had rained, the match would not have been cancelled'. There is actually a very straightforward consideration which shows that it is not always the case that negating a counterfactual conditional is equivalent to negating its consequent. This comes from considering counterfactual conditionals with impossible antecedents. First, let's take a principle which virtually all would accept, which is that a conjunction counterfactually implies its conjuncts. If **A** & **B** had been the case, then **A** would have been the case, and also if **A** & **B** had been the case, then **B** would have been the case. If we now consider the case where the antecedent is a contradiction, it gives us that if **A** & \neg **A** had been the case, **A** would have been the case, and

also that if $\mathbf{A} \ \& \ \neg\mathbf{A}$ had been the case, then $\neg\mathbf{A}$ would have been the case. If negating the consequent is equivalent to negating the whole thing, the second of those would have been equivalent to the negation of the first and we would have a contradiction. But we have to assert both those counterfactual conditionals, because they are both instances of the truism that a conjunction counterfactually implies any one of its conjuncts. This establishes, on pretty uncontroversial grounds, that negating a counterfactual conditional is not ipso facto equivalent to negating its consequent. It shows, from a syntactic point of view, that we can't think of the negation as in effect operating on the consequent. The only question, I think, is whether there are good reasons for independently accepting conditional excluded middle, which is the principle you called **(3)**. One could accept **(3)** even if one thinks that the negation of a counterfactual is different from the result of negating its consequent. **(3)** is a principle that has been defended by Stalnaker. It is valid on his logic of counterfactuals, and he has defended it without any confusion about what the contradictory of a counterfactual conditional is. In my view, his defence is unconvincing and unnecessary. We have perfectly good logics for counterfactuals, like the one David Lewis gave, where **(3)** is invalid. The cases to consider are cases where the antecedent seems to be in some way completely neutral between the consequent and the negation of the consequent, maybe cases where, say, indeterminism holds. If we have 'If the coin had been tossed, it would have come up heads' and 'If the coin had been tossed, it would have could come up tails' and therefore 'not heads', and there is nothing to choose between them. It's not that we don't know which is true, but as it were, reality itself doesn't decide in favour of either of them. One way we might think about this is that in order for a counterfactual conditional to be true, there has to be some sort of connection between the antecedent and the consequent. Let's not now try to specify what sort of connection that is required. It seems that there would be cases where the antecedent lacks that connection to the consequent, but also lacks that connection to the negation of the consequent. In those cases conditional excluded middle would

fail, but that is not a failure of bivalence, because if what's required for a counterfactual conditional to be true is that there is a connection of the right kind between the antecedent and the consequent, then all that is required for it to be false is the absence of such a connection. It is not required that there be some alternative connection going the opposite way. We could compare them to existential claims. What's required for 'There is a talking donkey' to be true is just that there be such a donkey. That means that what is required for that sentence to be false is simply that there be no such donkeys. It is not required that there be some other kind of donkey, that prevents all other donkeys from talking. It is only required that there be no talking donkeys at all. Similarly, if a counterfactual informs us that there is a connection of a certain kind, what it amounts to for it to be false is simply that there be no such connection.

III

PA: A natural progression from bivalence is to turn over to more general metaphysics. Let's think in terms of a familiar Quinean perspective: science introduces entities of certain kinds, and philosophy regiments the language of science into something like a first order language, where the objects in the domain of quantification are the things said to exist—hence the slogan 'to be is to be the value of a variable. From your perspective, this provides an impoverished ontology. How do you see yourself as differing from a traditional view of this kind?

TW: First, let me say something that seems right about Quine's perspective. That is that the existential quantifier of first order logic is the right way to express the claim that there are things of a certain kind. I'm taking existence here just as a matter of being

something or other, and not requiring a further property of concreteness. I think that, properly interpreted, Quine's claim that to be is to be value of a variable is correct. What I don't think is that it is as informative as Quine suggests, because I don't think that the metalinguistic element in talking about values of variables is really adding anything useful. It's not that things are because they are the values of variables, it is rather that they are values of variables because they are. The metaphysics is preceding the semantics, and not the other way around. I differ from Quine on several other points. First, I think Quine is being much too narrow in insisting that in order to evaluate the ontological commitments of a theory we have to regiment the theory in first order logic. There are many plausible and intelligible semantic constructions, for example modal operators, higher order quantifiers and operators for propositional attitudes, such as belief and knowledge. These are not devices of first order logic, but nevertheless are entirely legitimate. In my view, we're distorting a theory if it is first formulated in those terms, and then we insist on trying, in some procrustean way, to find a first order analogue of it. Suppose we take claims of possibility. If we try to put them in first order terms, we may have to go to a first order language in which we quantify over possible worlds. Then we'll understand claims of possibility as claiming, amongst other things, that there is a possible world of a certain kind. But that is not a commitment of the original claim, which is in good standing before it's regimented, and the regimentation may be unfaithful to what the philosopher is putting forward. I therefore think we need to be much more liberal than Quine was in what sorts of language we think it is legitimate for philosophers to formulate their theories in. A second point is that Quine's emphasis on natural science, which you hinted at when you talked about science introducing entities, seems to be too narrow. There is considerably more legitimate diversity within intellectual inquiry than Quine allows. Physics may be one of our highest achievements, but it is not the case that any form of knowledge needs to be derived from physics in order to count as part of our best theory of the world. So I would repudiate Quine's emphasis

on fundamental natural science to the exclusion of all other forms of inquiry, as the proper test of ontology. I would also insist that the emphasis on ontology to the exclusion of other forms of metaphysical commitment seems to me lopsided. Quine himself is aware that ontological commitment is not the only form of commitment; he does talk of ideology as well as ontology. I think that once one is more liberal about the sorts of languages that it is legitimate to theorize in one has to realize that there are all sorts of claims that are formulated, using semantic operations that are not first order. They nevertheless involve metaphysical commitments, because they involve commitments to how things really are. It is only that in the strict sense the commitments they involve are not ontological commitment, because they are not commitments to the effect that there are some things of a certain kind. I guess the fourth point on which I differ from Quine as you represented him is a more technical one. The way you formulated it was that we talk about objects in the domain of quantification. When we're quantifying unrestrictedly, which is what we do in metaphysics, we're quantifying so widely that, for reasons connected with Russell's paradox, we can't allow that all the things we are generalizing over (which are all the things there are) could fit into any single domain. No set is big enough to hold them all. We are therefore quantifying over everything, and not the set that contains everything, because there is not such set. But actually, Quine himself doesn't insist on this talk of domains, so that's not really a difference with Quine.

PA: Let me follow up on the last comment of quantification, together with the issue of modality. Two modal principles that stand out as crucial in your view of modal metaphysics is the Barcan formula (BF) and its converse (CBF). In their existential version, they may be stated as follows.

$$\begin{array}{ll} \mathbf{BF} & \diamond \exists x A \rightarrow \exists x \diamond A \\ \mathbf{CBF} & \exists x \diamond A \rightarrow \diamond \exists x A \end{array}$$

The claim **BF** reads informally as saying that ‘If there could have been something that met the condition **A**, then there is something that could have met condition **A**’, and the claim **CBF** reads with the opposite direction. If **BF** is true, the everything that could have existed must in some sense already exist. One motivation for rejecting **BF** is that, on many understandings of the sense of ‘exist’, it contains a metaphysical commitment that is highly implausible. How can the Barcan formulas be reconciled with or improve on our intuitions about existence?

TW: The first thing we need to do is to distinguish between material existence and existence in general. Numbers may be an example of things that exist without having material existence. By existence I mean ‘being something or other’. What is overwhelmingly plausible is that it is contingent what material things there are. There could have been more or fewer material things than there in fact are. That is consistent with **BF** and **CBF**, because it is consistent with the idea that if, for example, the table we’re sitting at had not been material it would still have been something. It would just not have been something material. Initially, that sounds like a strange claim, and the immediate question is: well, what would it have been? I’m not talking about the table as being a collection scattered atoms. That is still something material, and even those atoms might never have been. What I’m suggesting is that if there had been no such table and no such atoms, there would still have been a possible table, in other words, something that could have been a table and that could have been a material thing. But in those circumstances, it would not have been material, it would not have been a table, and so it would not have been located in space and time. It would have been a merely possible table, a merely possible material thing. Once we see that that is what **BF** and **CBF** is claiming, when they are understood with the quantifiers completely unrestricted and not limited to material things, it is much less obvious that there is any inconsistency between what those formulas are committing us to and what common sense is in any position to

assert or have any special authority over. It seems that scientifically informed common sense may have some authority over what materially and contingent things there are, but once the question is what merely possible material things there are, that sort of abstract questions is not to be decided by common sense. That question is to be decided by systematic theory. Of course, it is not at all obvious such theoretical considerations themselves will favour **BF** and **CBF**, but what I have tried to argue in a series of works is that in fact they do. When we try to develop a systematic theory of metaphysics of quantified modal logic, we get into a mess unless we accept **BF** and **CBF**. So, in the end, we are driven, not by some immediately intuitive consideration, but by considerations of the needs of systematic theory into accepting these formulas.

PA: When talking about the impoverished ontology of Quine you also mentioned the impoverished language. In your latest work you have been focusing on metaphysical issues with the use second order modal logic. Could you say something about why you think that future work in metaphysics will profit from the application of logics of this kind, and also perhaps what you take to be the most pressing question(s) that is opened up with this new method of analysis?

TW: Second order logic, depending on how we interpret it, allows us to generalize over, as well as over things, over the properties those things have. It also allows us to generalize in a plural way, a point that Peter Geach originally made. If you take a sentence like 'Some critics admire only one another' you can't formalize it in first order logic even though it sounds like a first order claim on first hearing. The plural quantifier 'some critics' needs something like second order logic in order to formalize it. I think second order logic is quite widely, though not universally, accepted for such purposes and for technical reasons it seems a much more natural background for a lot of mathematics for set theory and arithmetic than first order logic. At the same time, most philosophers are happy to use modal operators for possibility and necessity. So second order modal logic

is simply the result of combining these different resources that separately probably the majority of philosophers acknowledge the need for. Of course, the question is whether these two sets of logical notions interact in any interesting way. I think what people have not realized is how rich the new questions that arise concerning the interaction between them are. A simple example is that we have these two ways of interpreting second order logic; one in terms of plurals and one in terms of properties. In the non-modal case, it is not obvious it makes any difference which one you use, but as soon as you introduce modal operators, you see these two forms of interpretations of second order quantification interact quite differently with the modal operators. We can use the interaction with modality as a way of separating different interpretations of second order logic. They might both be equally legitimate interpretations, but if they are different we need to keep them separated and it is modality that separates them. But I am also interested in second order modal logic because it provides a wider setting for the evaluation of formulas such as **BF** and **CBF**. It is a more expressive language than first order modal logic and therefore allows us to ask more questions and bring in a wider variety of systematic considerations to bear on the evaluation of these formulas. In effect, it gives us a wider evidence base for reaching our conclusions, even about formulas like **BF** and **CBF**, which are themselves formulas of first order modal logic. Once we consider how they interact with second order modal logic, we can make a more informed judgement of what their overall theoretical effects are and in my view that judgement comes out in favour of the two forms.