

COUNTERFACTUALS AND MODAL EPISTEMOLOGY

TUOMAS E. TAHKO
University of Helsinki

Summary

What is our epistemic access to metaphysical modality? Timothy Williamson suggests that the epistemology of counterfactuals will provide the answer. This paper challenges Williamson's account and argues that certain elements of the epistemology of counterfactuals that he discusses, namely so called background knowledge and constitutive facts, are already saturated with modal content which his account fails to explain. Williamson's account will first be outlined and the role of background knowledge and constitutive facts analysed. Their key role is to restrict our imagination to rule out irrelevant counterfactual suppositions. However, background knowledge turns out to be problematic in cases where we are dealing with metaphysically possible counterfactual suppositions that violate the actual laws of physics. As we will see, unless Williamson assumes that background knowledge corresponds with the actual, true laws of physics and that these laws are metaphysically necessary, it will be difficult to address this problem. Furthermore, Williamson's account fails to accommodate the distinction between conceivable yet metaphysically impossible scenarios, and conceivable and metaphysically possible scenarios. This is because background knowledge and constitutive facts are based strictly on our knowledge of the actual world. Williamson does attempt to address this concern with regard to metaphysical necessities—as they hold across all possible worlds—but we will see that even in this case the explanation is questionable. These problems, it will be suggested, cannot be addressed in a counterfactual account of the epistemology of modality. The paper finishes with an analysis of Williamson's possible rejoinders and some discussion about the prospects of an alternative account of modal epistemology.

1. *Introduction*

Timothy Williamson (2005, 2007a, 2007b) has recently suggested that the epistemology of metaphysical modality is a special case of the epistemology of counterfactuals. This account has already received a substantial amount

of attention.¹ However, one aspect of Williamson's account deserves further examination, namely the status of what he calls 'background knowledge' and 'constitutive facts'. These will be the focus of our discussion, with special attention to the role of physical possibility and conceivability. First, Williamson's account will be sketched, and an analysis of the role of background knowledge in his account will be provided. Second, it will be argued that unaccounted modal knowledge is present in background knowledge, and that this conflicts with Williamson's assimilation of modal knowledge to counterfactual knowledge. Hence, Williamson's account of modal epistemology faces a serious problem. This will be the subject of the second section. Further, in the third section it will be argued that counterfactuals cannot serve the role in modal epistemology that Williamson postulates for them, if previous modal knowledge is present in background knowledge. This is due to problems in distinguishing between conceivable and metaphysically possible scenarios, and conceivable but metaphysically *impossible* scenarios. In the fourth section we will analyse the role of physical possibility with regard to constitutive facts, and in the fifth section Williamson's reply to an objection which is related to the one raised in this paper will be considered. Two possible rejoinders that are available for Williamson will be discussed in the sixth section, before finally considering the prospects for an alternative account of modal epistemology in the seventh section.

The formal aspects of Williamson's account, i.e., the manner in which he derives metaphysical modality from the logic of counterfactuals, will not be discussed here. If our analysis of the nature of background knowledge and constitutive facts is correct, then it makes little difference whether Williamson's formal presentation of counterfactuals and metaphysical modality is feasible, as the modal content of the scenarios that Williamson discusses as well as our epistemic access to it precedes the counterfactual analysis. That is, the arguments about to be presented suggest that we must already have epistemic access to metaphysical modality before we can even formulate a counterfactual account of it. The final section of the paper focuses on the source of this previous modal knowledge in particular.

1. See for instance Jenkins (2008) and Roca-Royes (2011a, 2011b). The critique of Williamson's account developed here is on similar lines as Roca-Royes', but the focus will be somewhat different.

2. Summary of Williamson's account

Williamson's starting point is that the epistemology of metaphysical modality must somehow derive from our ordinary cognitive capacities. Specifically, he thinks that our ordinary cognitive capacity to handle counterfactuals also includes a capacity to handle metaphysical modality (2007b, 136). According to Williamson, this capacity is our only epistemic route to metaphysical modality. It is perhaps best to illustrate Williamson's account with one of his examples:

You are in the mountains. As the sun melts the ice, rocks embedded in it are loosened and crash down the slope. You notice one rock slide into a bush. You wonder where it would have ended if the bush had not been there. A natural way to answer the question is by visualizing the rock sliding without the bush there, then bouncing down the slope. You thereby come to know this counterfactual:

[CF] If the bush had not been there, the rock would have ended in the lake. (Williamson 2007b, 142.)

According to Williamson, we come to know counterfactuals like (CF) with the help of our imagination. This is something that we do constantly in our everyday lives; we come up with counterfactual suppositions that accurately predict the behaviour of the physical world. However, we are immediately faced with a problem: how are we able to choose the correct scenario—that is, the scenario which most accurately reflects the behaviour of the rock in the absence of the bush—given that our imagination can come up with the wildest of scenarios, such as 'the rock rising vertically into the air, or looping the loop, or sticking like a limpet to the slope' (ibid., 143).

Indeed, our imagination² is unrestricted: everything from Disney fairy tales to philosophical zombies is supposedly imaginable, but whether there is any kind of helpful overlap between what is imaginable or conceivable and what is metaphysically possible is another question. There will of course be *some* overlap between them: on the face of it, everything that is metaphysically possible is also imaginable, at least by an ideal conceiver.³

2. Williamson (2007b, 163) mainly talks about imaginability as a tool for assessing counterfactuals, but he seems to think that conceivability and imaginability are both aspects of the same cognitive faculty. I will use these terms interchangeably since Williamson does so as well, although strictly speaking they could be distinguished.

3. Admittedly, this is open to debate, since the lack of any restrictions will generate so called

But since the range of imaginable scenarios is so vast, something more is needed. Accordingly, Williamson needs to explain how we restrict our imagination so that irrelevant scenarios can be ruled out. Here is what he says:

You do not imagine it those other [irrelevant] ways because your imaginative exercise is radically informed and disciplined by your perception of the rock and the slope and your sense of how nature works. The default for the imagination may be to proceed as ‘realistically’ as it can, subject to whatever deviations the thinker imposes by brute force: here, the absence of the bush. Thus the imagination can in principle exploit all our background knowledge in evaluating counterfactuals. (Williamson 2007b, 143.)

Williamson uses the term ‘background knowledge’ here for the first time, and immediately notes the difficulty in separating between background knowledge and what has to be removed with the help of imagination for one to be able to imagine the antecedent, i.e., the problem of cotenability. This is indeed an important point, and we will return to it shortly.

Let us consider the example at hand in more detail. Williamson claims that our imagination is somehow restricted by our perception of the situation and our ‘sense’ of how nature works. From this starting point, our imagination proceeds as ‘realistically’ as it can. A preliminary constraint that helps to restrict our imagination in the case of (CF) is our knowledge of how rocks generally behave in similar situations. Williamson talks about ‘folk physics’ (2007b, 146). Folk physics includes the general principles which govern our expectations of motion and other physical processes in real time. It enables us to make many useful predictions, but strictly

‘Kaplan’s paradox’ concerning possible worlds, which highlights the apparent inconsistency of the following three (plausible) claims (reconstructed in Chalmers 2011, section 9):

- (i) There are at least as many propositions as sets of worlds.
- (ii) There are at least as many worlds as propositions.
- (iii) There are more sets of worlds than worlds.

Chalmers (ibid.) discusses the paradox specifically with regard to epistemic modality and the idea that scenarios can have infinite extent. He notes that the paradox arises even if we replace worlds and propositions with scenarios and intensions. Any scenario could then be specified by an infinitary conjunction and sets of such scenarios by a disjunction of the conjunctions. An ideal conceiver could presumably entertain conjunctions and disjunctions of this type.

Chalmers suggests some options for overcoming this paradox, but this is not the place to go into the details. For our purposes, it is not crucial that epistemic modality is unrestricted; merely the fact that we can imagine numerous metaphysically impossible scenarios is sufficient for the argument at hand.

I’d like to thank an anonymous referee for bringing this issue to my attention.

speaking folk physics is false and Williamson emphasizes that we should not view it as a premise, but rather as ‘a locally but not globally reliable method of belief formation’ (ibid.). Accordingly, although folk physics is a part of our background knowledge, we should view it more like a pattern of inference rather than a premise in our counterfactual suppositions. Because of this, Williamson moves on to discuss simulation theories, as they may describe our capacity for counterfactual reasoning more accurately. He then applies the simulation idea to the case of (CF):

[W]e have various propensities to form expectations about what happens next: for example, to project the trajectories of nearby moving bodies into the immediate future (otherwise we could not catch balls). Perhaps we simulate the initial movement of the rock in the absence of the bush, form an expectation as to where it goes next, feed the expected movement back into the simulation as seen by the observer, form a further expectation as to its subsequent movement, feed that back into the simulation, and so on. [...] *The very natural laws and causal tendencies our expectations roughly track also help to determine which counterfactual conditionals really hold.* Thus some reliability in the assessment of counterfactuals is achieved. (Williamson 2007b, 148f.; my emphasis.)

The sentence I have emphasized suggests that although Williamson is reluctant to consider folk physics as a premise in our counterfactual suppositions, he nevertheless thinks that the simulation gets its reliability from the fact that our expectations roughly track the actual laws of nature and causal tendencies of the world. Of course, as Williamson notes, there must be some sort of correspondence between the simulation and reality because otherwise we would not catch balls. This is a crucial assumption, because it entails that the simulation must be restricted by the actual laws of physics—our expectation-forming capacities are ultimately governed by what is possible in virtue of the laws of nature and causal tendencies that actually hold in the world. Ideally, scenarios that are physically impossible will be omitted, because they would not produce reliable results.

Williamson sees a partial solution to the mentioned problem of cotenability here: our expectation-forming capacities aim to ‘roll back’ history in order to imagine alternative scenarios, but, as Williamson puts it, ‘we seem to have a prereflective tendency to minimum alteration in imagining counterfactual alternatives to actuality’ (ibid., 151). Williamson notes a connection to the Lewis-Stalnaker possible worlds semantics here; we

are primarily interested in nearby worlds. However, it should be noted that this solution to the problem of cotenability can only work if our prereflective tendency to restrict our attention to nearby worlds has some means to determine which worlds indeed *are* nearby worlds, that is, we must already possess some information that enables us to restrict our expectation-forming capacities.

Now to apply all this to (CF). We are primarily interested in how the rock would behave in nearby possible worlds where the bush is absent. Since we have a tendency for minimum alteration, we wish to rule out cases where the antecedent of (CF) is physically impossible. The purpose of the counterfactual and the whole expectation-forming process is, I take it, to predict how the sliding rock would in fact behave in an alternative scenario—we are not generally interested in scenarios which are physically impossible, even though philosophers sometimes do talk about such scenarios. Since Williamson explicitly rules out physically impossible scenarios—such as the rock floating above the lake—and emphasizes that we should proceed as ‘realistically’ as we can, with minimum alterations, it is clear that he is interested in the behaviour of the rock in scenarios that do not radically deviate from the actual world. Indeed, if the purpose of the counterfactual scenario is to predict the behaviour of the rock should we in fact proceed to remove the bush, then we better make sure that nothing else in the scenario that might be relevant for the behaviour of the rock differs from the initial conditions. Hence, the reading of ‘physically possible’ that we are using here refers to possibility in virtue of the actual laws of physics.⁴

There are of course complications concerning the metaphysical status of the laws of physics. For one thing, it is an open question whether or not the true, actual laws of physics are metaphysically necessary.⁵ For instance, we may have a counterfactual with an antecedent which violates the actual laws of physics, but whether such an antecedent is *metaphysically* impossible depends on whether alternative laws of physics are metaphysically

4. Alternatively, we could restrict our attention to only those (nearby) worlds in which the laws of physics are the same as in the actual world, or at least similar to a very high degree, in which case the notion of ‘physical possibility’ would concern possibility in virtue of the set of physical laws familiar to us while acknowledging that there may be alternative laws of physics.

5. The more common view is probably that the actual laws of physics are not metaphysically necessary (cf. Carroll 2011), but see for instance Swoyer (1982) or Shoemaker (1998) for a defence of the opposite view; Ellis (2001) and Bird (2007) have also supported the metaphysical necessity of laws.

possible. In any case, in the context of Williamson's example (CF), it is clear that we want to rule out scenarios which radically violate the true, *actual* laws of physics, as they would be irrelevant for our expectation-forming capacities.

But now the problem of cotenability strikes back. Our imagination is restricted by background knowledge, which includes a complex pattern of inference based on our simulative expectation-forming capacities. The success of the predictions that these capacities produce depends on a correspondence between the laws of nature and the causal tendencies that actually hold in the world (and, at least to a large extent, in nearby worlds). This is fine as far as it goes, but so far there has been no explanation as to how we are supposed to be able to pick out scenarios which are in fact metaphysically possible. It seems that our imagination may easily produce scenarios that are not only physically impossible, but also *metaphysically* impossible: we can conceive of cats being robots or demons, even though it is generally thought that cats are necessarily animals. The reason why this is problematic is that since we were supposed to focus on nearby worlds, scenarios where the antecedent of the counterfactual is metaphysically impossible should be ruled out. The only obvious candidate for ruling out metaphysical impossibilities is our background knowledge, which corresponds at least roughly with the true, actual laws of physics—surely anything that is physically possible is also metaphysically possible. But it is the same background knowledge that must be responsible for our prereflective tendency to choose nearby worlds when imagining counterfactual alternatives to the actual world. So, far from being a solution to the problem of cotenability, all this seems to beg the question concerning the very origin of our modal knowledge.

We can expect a reply to this critique though. The reply is that our expectation-forming capacities have all of our background knowledge and beliefs available to both enable and restrict the process of imagining counterfactual scenarios (cf. Williamson 2007b, 153). By any standard, the role of background knowledge in Williamson's account is extraordinarily important. Without it, we would only have our unrestricted capability to imagine counterfactual scenarios that may or may not be metaphysically possible. Yet, unless we have some prior means to distinguish between scenarios that are merely imaginable (but not metaphysically possible) and scenarios that are metaphysically possible, our expectation-forming capacities are no good. The core of this critique is that it is possible to imagine all manner of scenarios that are perfectly compatible with

all our background knowledge, but may nevertheless be metaphysically impossible.⁶

Williamson could of course question the use of metaphysically impossible antecedents in counterfactual suppositions, but he is not immune to this critique; there are passages where he implicitly accepts metaphysically impossible antecedents. For instance, Williamson considers a scientist evaluating the claim that space-time has ten dimensions (ibid., 151). Now, he correctly notes that the scientist may ask whether the actually observed phenomena would have occurred if the hypothesis concerning ten dimensions were true. But unless the hypothesis about ten space-time dimensions is metaphysically possible, what guarantees do we have about its correspondence with the actually observed phenomena? We can easily produce a counterfactual conditional about the hypothesis that the space-time is fundamentally fairy-dust in such a way that it is perfectly compatible with all the actually observed phenomena. Of course, that would not be a particularly useful counterfactual supposition, and this is exactly because the antecedent is metaphysically impossible. There is a natural way out of this though: perhaps it is background knowledge that helps us to distinguish between metaphysically impossible and metaphysically possible scenarios as well. The problem is that if this is correct, then background knowledge cannot be modally innocent, as we set out to argue in the first place.

3. *The role of imaginability/conceivability in Williamson's account*

In this section and the next we will get into the bottom of the role of background knowledge in Williamson's account. We will first examine whether imaginability or conceivability can get off the ground without the input of background knowledge. Our interest in the relationship between background knowledge and imaginability/conceivability is two-fold. There is firstly the general problem of distinguishing between metaphysically possible and metaphysically impossible yet conceivable scenarios. Secondly, there is the more specific problem concerning counterfactual scenarios

6. See for instance Kment (2006, 248 ff.) for an analysis of why counterfactuals with metaphysically impossible antecedents need not be vacuously true. Further support for non-vacuumism can be found in Nolan (1997) and Brogaard and Salerno (2008). It may not come as a surprise that Williamson's account breaks down when metaphysically impossible antecedents are introduced, but there is more to be said about why exactly this is the case.

which are physically impossible but metaphysically possible. Williamson's account is able to accommodate nearby worlds that roughly correspond with the actual world with respect to the laws of nature and causal tendencies in the world in virtue of our background knowledge, but what could serve a role similar to background knowledge in scenarios that are merely metaphysically possible (i.e. physically *impossible*)? More generally, if our background knowledge concerns the laws of nature and causal tendencies familiar from the actual world, how can we restrict our imagination regarding scenarios that do not concern the actual world? A potential solution for both of these problems would be to argue that the actual laws of nature are metaphysically necessary, but Williamson says nothing to this effect.

Another way to go would be to show that imaginability/conceivability entails metaphysical possibility and then proceed to restrict metaphysical possibility in terms of background knowledge. However, further problems emerge if this route is taken, because conceivability does not automatically provide us with tools to distinguish between metaphysical possibilities and metaphysical impossibilities.⁷

A visual illustration (*Figure 1*) might help to clarify the situation.

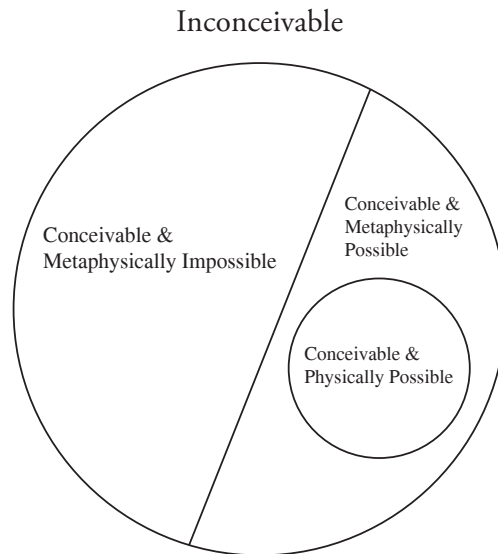


Figure 1

7. The two-dimensional account developed by Chalmers and others *may* provide us such tools, but Williamson is no two-dimensionalist. For more discussion on the differences between Chalmers' and Williamson's accounts of conceivability, see Roca-Royes (2011a).

We are primarily interested in the distinction between the left and the right half of the circle in *Figure 1*—between what is conceivable and metaphysically impossible, and what is conceivable and metaphysically possible. The latter includes both physically possible and physically impossible scenarios. An example of the former could be a scenario where water is XYZ (if water has its actual molecular structure by necessity), an example of the latter could be a scenario in which a rock is floating above a lake—this would presumably fall into the category of the physically impossible.

There is an overlap between what is conceivable and what is metaphysically possible, but given that there is also an overlap between what is conceivable and what is metaphysically impossible, the previous overlap does not entail a reliable epistemic link between conceivability and metaphysical possibility. Indeed, there are infinitely many imaginable metaphysically possible scenarios (albeit in the case of non-ideal conceivers this is subject to limitations in cognitive capacity), but there are also infinitely many imaginable metaphysically *impossible* scenarios. If we hope to use conceivability as a reliable guide to metaphysical possibility, we should somehow be able to distinguish between metaphysically possible and metaphysically impossible scenarios.

So, what would be needed for us to be able to distinguish between the left and the right half of the circle, between metaphysical possibilities and metaphysical impossibilities that are nevertheless conceivable? Williamson is not in the same boat with those who have developed sophisticated but controversial accounts to defend the connection between conceivability and possibility (e.g., Chalmers 2002). Chalmers for instance relies on two-dimensional modal semantics to salvage a sense of conceivability that has a bearing on metaphysical modality⁸, but Williamson does no such thing. Indeed, he is wary of some of these accounts himself, although he does wish to underline the use of imagination in evaluating counterfactuals. The general understanding of conceivability is similar in Chalmers' and Williamson's accounts though: they both treat conceivability in terms of an absence of contradictions, that is, a proposition is conceivable insofar as it is not contradictory. Chalmers (*ibid.*) gives a corresponding negative definition of conceivability: everything not ruled out by *a priori* reason-

8. However, Chalmers treats metaphysical and conceptual possibility as co-extensive, which implies that the line dividing the left and the right halves of the circle in *Figure 1* can be removed altogether. I cannot do justice to the work of Chalmers and other two-dimensionalists here, but Williamson's approach clearly differs from Chalmers'. In what follows the principal differences between these accounts will be discussed briefly.

ing is conceivable. Plausibly, *a priori* considerations rule out contradictory propositions, so this is one crucial qualification for conceivability. However, *a priori* considerations do not rule out metaphysically impossible antecedents—in fact, Williamson acknowledges that there are scenarios which are conceivable and metaphysically impossible, such as water not containing oxygen:

Discussions of the epistemology of modality often focus on imaginability or conceivability as a test of possibility while ignoring the role of the imagination in the assessment of mundane counterfactuals. In doing so, they omit the appropriate context for understanding the relation between modality and the imagination. For instance, scorn is easily poured on imagination as a test of possibility: it is imaginable but not possible that water does not contain oxygen, except in artificial senses of “imaginable” that come apart from possibility in other ways, and so on. Imagination can be made to look cognitively worthless. Once we recall its fallible but vital role in evaluating counterfactual conditionals, we should be more open to the idea that it plays such a role in evaluating claims of possibility and necessity. (Williamson 2007b, 163.)

Williamson appears to be using ‘imaginability’ and ‘conceivability’ interchangeably here; I have opted to use ‘conceivable’, but nothing crucial hangs on this choice of terminology. The most important difference between Chalmers’ and Williamson’s understanding of conceivability is that Williamson is undecided about the epistemic status of our modal judgements; where Chalmers clearly thinks that conceivability and hence modal epistemology is an *a priori* matter, Williamson is unwilling to label it either *a priori* or *a posteriori* (cf. 2007b, 165–169).⁹ So, Williamson acknowledges the problems associated with conceivability and possibility, although he does not explicitly discuss the problem at hand, namely how to distinguish between the left and the right halves of the circle in *Figure 1*. However, he does hint towards an answer, at least regarding cases such as ‘gold is the element with atomic number 79’. Once again he relies on background knowledge:

If we know enough chemistry, our counterfactual development of the supposition that gold is the element with atomic number 79 will generate a

9. Again, see Roca-Royes (2011a) for further discussion on Williamson vs. Chalmers (and Yablo) on conceivability—she emphasizes in particular that whereas Chalmers and Yablo try to make room for modal rationalism, Williamson has no such interests. Accordingly, it is clear that Williamson will not have the same tools as Chalmers has to accommodate the distinctions in *Figure 1*.

contradiction. The reason is not simply that we know that gold is the element with atomic number 79, for we can and must vary some items of our knowledge under counterfactual suppositions. Rather, part of the general way we develop counterfactual suppositions is to hold such *constitutive facts* fixed. (Williamson 2007b, 164; my emphasis.)

Williamson provides a partial answer to the problem, as supposedly we can dismiss the conceivable but metaphysically impossible scenario of gold having some other atomic number than 79 with the help of ‘constitutive facts’, i.e., background knowledge that we do *not* vary when considering counterfactual scenarios. The manner in which the method reveals that we must hold constitutive facts fixed is based on the idea that any counterfactual supposition which fails to hold such a fact fixed will generate a contradiction. This enables us to distinguish between conceivable metaphysical impossibilities and metaphysical possibilities. However, this solution will not work in all cases: it only works in cases where we are dealing with metaphysical necessities, such as ‘gold is the element with atomic number 79’.¹⁰ This is because in cases where we are dealing with mere metaphysical possibilities we can vary many more items of our knowledge under counterfactual suppositions: in cases such as the rock sliding down a slope the possibilities for variation are much greater—even the laws of physics may be varied unless it is assumed that they are metaphysically necessary, which, as has been noted, is something that Williamson does not explicitly argue for. So, only metaphysically necessary constitutive facts could help in addressing the problem of metaphysically impossible yet conceivable scenarios, but many counterfactual suppositions do not involve *any* metaphysically necessary constitutive facts, that is, many counterfactual antecedents are metaphysically contingent propositions. Indeed, even Williamson’s own example concerning the sliding rock (CF) lacks metaphysically necessary constitutive facts.

10. It should be noted here that although ‘gold is the element with atomic number 79’ is indeed generally considered to be a metaphysical necessity, this is by no means uncontroversial, at least unless the metaphysical necessity of the true, actual laws of physics is assumed. See Lowe (2007), Hendry (2006), and Tahko (2009) for further discussion; the problem has also been discussed by Bealer (1987) and Salmon (2005).

4. *Constitutive facts and physical possibility*

The upshot of the previous section is that unless there are some metaphysically necessary conditions that serve as clear candidates for constitutive facts that should be held fixed, such as the fact that elements have their atomic number by necessity (if this is indeed the case), then there are no obvious criteria which would help us to decide which items of background knowledge should be held fixed. Thus, it is not clear how we can constrain our imagination so that metaphysically impossible variations are ruled out, except in cases where we are dealing with metaphysical necessities. But even in cases such as ‘gold is the element with atomic number 79’ we ought to have some good reasons to think that the atomic number *is* a constitutive fact. It is not obvious that Williamson would agree with this though: Williamson does not say that we know *which* facts are constitutive, only that we *know* some constitutive facts. He suggests that if we know enough chemistry, any counterfactual supposition which varies the constitutive fact that gold is the element with atomic number 79 will generate a contradiction. However, even in this case it is not straightforward that we know it to be a constitutive fact, or even that it *is* a constitutive fact.

For instance, Robin Hendry (2006) has recently suggested that rather than atomic weight, we should use nuclear charge as the defining characteristic of elementhood. Drawing on Lavoisier, Hendry suggests that nuclear charge appears to be the most crucial attribute for the chemical properties of substances and hence the most likely to track their essential features. Now, Williamson’s account could of course accommodate Hendry’s scenario as well: if Hendry is correct, then any counterfactual supposition which varies the constitutive fact that nuclear charge is the defining characteristic of elementhood will generate a contradiction. So now we have two competing hypotheses of what the constitutive fact is.

Say that Hendry is correct about elementhood and that atomic weight is not a constitutive fact in this case, rather, it is nuclear charge. If we do not know this and mistakenly keep atomic weight fixed while failing to keep nuclear charge fixed in our counterfactual suppositions, we are going to get unexpected results. Specifically, on one hand we will get a contradiction where there should be none if we mistakenly hold atomic weight fixed, and on the other hand we will not get a contradiction where there should be one if we fail to keep nuclear charge fixed. We must be able to know what we should hold fixed if we hope to get useful results with our counterfactual suppositions. Furthermore, since these constitutive facts are

supposed to be metaphysically necessary, deciding which one is correct will surely require some modal knowledge. Yet, Williamson's account remains completely neutral about which hypothesis is the correct one, it can only tell us that *if* a certain fact is metaphysically necessary, then varying it in a counterfactual supposition will generate a contradiction. Interesting as this may be, it seems to omit a rather important part of the story, namely, the part in which we come to know metaphysical necessities. The method can produce reliable results only if we know which facts to hold fixed. The only way out is that we must have some modal knowledge of metaphysical necessities that precedes the counterfactual assessment.

At times Williamson suggests that we come to know constitutive facts empirically, even though observation alone may not be enough to determine that they are indeed constitutive (cf. 2007b, 165–169). He remains non-committal about whether modal judgements are *a priori* or *a posteriori*, but I do not think that anything very important in terms of the present discussion depends on this choice. It is at least clear that Williamson considers experience to act in an enabling role for our knowledge of modal judgements. A further issue may emerge here: even if we come to know constitutive facts empirically, many metaphysically possible scenarios do not have any bearing on the actual world since they are much too far removed from the world as we know it. What this means is that 'our sense of how nature works' (2007b, 143), which is central for Williamson's account of constitutive facts, is based on actual observations of how nature works rather than mere conceivability. But metaphysical modality is surely not restricted by how nature works in the *actual* world: for instance, a different set of laws of physics may be possible. At any rate, this is the typical Humean conviction about the laws of nature: there is no necessity in nature (cf. Swyer 1982). The idea is that the laws of nature concern only actual (accidental) regularities and hence further arguments are needed to justify the (metaphysical) necessity of these regularities. On the face of it there seems to be nothing impossible about superluminal motion, for instance—or at least the notion does not appear to be inconsistent. Recent work in scientific/dispositional essentialism (e.g. Ellis 2001, Bird 2007) in particular has made necessitarianism about the laws of nature popular, but the debate remains open.¹¹ In any case, Hendry's suggestion concerning elementhood appears to be a consistent, metaphysically possible scenario.

11. In fact, Roberts (2010) has argued that there may be reasons to think that some laws of nature are metaphysically contingent even if the scientific essentialism advocated by Ellis and others is correct about the metaphysics of natural kinds and properties.

This alternative conception of elementhood and the conception of elementhood in terms of atomic weight cannot both ‘carve nature at its joints’, but they both appear to be metaphysically possible. This would mean that we have two metaphysically possible scenarios, but only one of them can be physically possible (where physical possibility is understood as possibility in virtue of the actual laws of physics). Yet, Williamson’s method is unable to distinguish these scenarios without prior modal input.

The crucial shortcoming here is that insofar as our background knowledge is based on empirical information, it will be of no help in analysing physically impossible yet metaphysically possible scenarios and hence a great deal of the scope of metaphysical modality is completely omitted. It *will* of course help in determining what is *physically* possible, but unless the laws of physics are metaphysically necessary, there are also metaphysical possibilities that are physically impossible. Examples such as ‘gold is the element with atomic number 79’ may create the illusion that we have a working theory because metaphysical necessities are an obvious candidate for constitutive facts, but cases where no such constitutive facts are forthcoming will be more challenging, and we have just seen that our knowledge of constitutive facts is problematic even in the case of elementhood.

It seems then that the only reason we hold constitutive facts fixed is that they are metaphysically necessary. This leaves entirely open how we are supposed to know which items of our background knowledge *are* metaphysically necessary and hence constitutive facts that should be held fixed. Our epistemic access to these modalities cannot be in terms of counterfactuals—that would render Williamson’s account irredeemably circular, as constitutive facts are supposed to help us to get the account started in the first place. Perhaps a full analysis of how counterfactual reasoning produces modal knowledge is too much to ask, but at the very least Williamson should, on pain of circularity, provide us with some reasons to think that constitutive facts can be grounded independently.

5. *Williamson’s reply*

We should give Williamson a chance to reply, because he is certainly not ignorant about the possibility of the sort of objection that has been raised in this paper. The objection that Williamson specifically replies to goes as follows:

Knowledge of counterfactuals cannot explain modal knowledge, because the former depends on the latter. More specifically, in developing a counterfactual supposition, we make free use of what we take to be necessary truths, but not of what we take to be contingent truths. Thus we rely on prior or at least independent stock of modal knowledge or belief. (Williamson 2007b, 169f.)

While this objection is not exactly the one developed above, it nevertheless underlines an important point, namely the apparent need for modal knowledge that is prior to, or independent of, counterfactual knowledge. Williamson insists that the use of necessary truths in developing further counterfactual suppositions does not imply that we have some further cognitive ability to handle modality independently of counterfactuals (*ibid.*). As an example he mentions the case of the atomic number of gold and claims that the judgement that ‘gold is the element with atomic number 79’ is metaphysically necessary is not required before we can use the proposition for the purposes of developing a counterfactual supposition. According to Williamson, the metaphysical necessity of the proposition is revealed as the output of the method of analysing counterfactual suppositions. But it will not do to insist that metaphysical necessities are recognized as necessary because they are held fixed by the method and thus need no independent grounding. This only amounts to admitting that the method has no means to distinguish between *genuine* metaphysical necessities and mere metaphysical possibilities that are held fixed due to a flaw in the method, or a mistaken intuition—we already saw this in the case of elementhood in the previous section. The problem, to put it briefly, is that our capacity to develop counterfactual suppositions is entirely unreliable unless we have independent epistemic access to metaphysical modality. The method *might* be reliable if we happened to hold fixed the correct facts, but we have no means to determine whether this is indeed the case. The upshot is that there is no way to test the method or to calibrate it in the first place. We would need prior access to metaphysical necessities to enable this, but that is exactly what Williamson denies.

Equally problematic is the fact that Williamson’s method of analysing counterfactual suppositions cannot be applied to cases where there are no metaphysical necessities at hand, such as the very first example that was considered, a rock sliding down a slope (CF). The method can address this scenario only by postulating that the laws of nature and causal tendencies that we actually observe are metaphysically necessary. Indeed, perhaps we are supposed to recognize them as necessary because they are held fixed by the method, for supposedly *only* constitutive facts are held fixed by

the method.¹² But if this is the case then Williamson's method is rather controversial, as the modal status of the laws of nature is far from settled.

6. *Possible rejoinders for Williamson*

Williamson has two apparent escape routes. Firstly, he could acknowledge that we need prior epistemic access to metaphysical modality after all, and attempt to address the nature of this epistemic access in a number of ways. One option here would be to adopt the view familiar from Kit Fine's (1994) work, namely that metaphysical modality is a special case of essence. According to Fine's account, the traditional manner of analysing essence in terms of modality is mistaken, and the order of explanation is in fact the opposite: essence is a more fundamental notion than modality, and metaphysical modality can be reduced to essence. Essentialist facts themselves are thus non-modal, although they do have modal implications. Fine's account of essence is inspired by the Aristotelian view according to which the essence of a thing is its *real definition*. For current purposes, it is not necessary to consider the details of Fine's interpretation of essence, as our main concern is that we need to explain the modal implications of constitutive facts without relying on further modal facts. The reason why this approach might be of use to Williamson is that according to the Finean account, facts about essences can plausibly be understood as non-modal, constitutive facts. Now, if we have some prior epistemic access to constitutive, essentialist facts, then we can address the concern that was previously raised against Williamson's account. For instance, if we consider atomic number to be essential for elements, then the atomic number of gold being 79 is a constitutive fact that we know independently of Williamson's counterfactual analysis.

12. If we held non-constitutive facts fixed, we would end up ruling out certain possibilities that we wish to include, say, if we hold it fixed that Earth is the third planet from the sun, we would end up ruling out the possibility that the planet Venus never formed, which is surely (physically) possible. Similarly, if we fail to hold some constitutive fact as fixed, say that elements are defined by their atomic number (if that is indeed the case), then we would erroneously include metaphysical impossibilities, such as gold failing to be the element with atomic number 79. Note also that although it may appear that in (CF) we hold it fixed that there is a lake towards which the rock is sliding, this is not strictly correct: one metaphysically possible scenario is that the lake evaporates before the rock reaches it. Be that as it may, we are once again faced with the problem concerning the scope of physical possibility that was discussed earlier. See also Roca-Royes (2011a).

Admittedly, adopting the Finean approach would undermine much of the explanatory power of Williamson's account, since one initial attraction of the counterfactual analysis of modality is that it appears to explain modal epistemology without the need for further work. Consequently, the essentialist route is perhaps not the most likely route for Williamson, even though he does express some sympathy towards this line of thought (2007, 170). Furthermore, even if Williamson were to take this line, the apparent difficulty that remains is the question concerning the epistemology of essentialist facts—Fine's account of essence leaves almost entirely open how we come to *know* essentialist facts, we might just have to take them as brute facts. The purpose of the present analysis is to explain modal knowledge, but if this is done with reference to further—non-modal—essentialist facts, then the pressing question concerns the epistemology of essence. We will return to the approach inspired by Fine's work in the final section and consider whether it has any independent appeal.

Given the problems associated with the first possible rejoinder, it may be that Williamson would be more likely to go for a different approach. A second possible rejoinder for Williamson would be to argue that the method of developing counterfactual suppositions is reliable on some other grounds, but it is not clear what these grounds could be; at the very least this would require a thorough analysis of the nature of constitutive facts, and especially whether we can know which facts are constitutive before applying the method. In general, the greatest challenge for any possible rejoinder is to come up with plausible criteria for distinguishing constitutive and non-constitutive facts. Admittedly, Williamson does not claim to be able to do this, but, as we have seen, without such criteria it will be difficult to give a complete analysis of modal knowledge.

Accordingly, despite the problems it entails, the first rejoinder seems much more promising, as it would enable us to give a complete analysis of modality provided that we can give a plausible account of the epistemology of essence. However, even if Williamson were to take this line, the order of explanation in his account would have to be revised. For if it is acknowledged that modal truths are grounded in essences, then the most important part of the story concerns the link between the two, and the nature of the epistemic problem becomes more complex due to the added requirement of explaining our knowledge of essence.

Williamson might resist this approach by arguing that our capacity to handle metaphysical modality cannot be based on anything except our

capacity to think about the physical world in general—including our capacity to handle counterfactuals—as postulating a further cognitive capacity which is responsible for our ability to handle modal knowledge would bear heavy ontological costs (Williamson 2007b, 162). But given that Williamson’s own account clearly requires something further than our cognitive capacity to handle counterfactuals if it is to be reliable, namely knowledge of constitutive facts, it is plausible that the more ontologically parsimonious solution would be to treat *both* our ability to handle metaphysical modality and our ability to handle counterfactuals concerning the physical world as special cases of our cognitive capacity to handle essences, drawing on Fine’s account. This would further motivate the first possible rejoinder.

7. *Prospects for an alternative account of modal epistemology*

Given the discussion in the previous section, it might be interesting to conclude the paper with a brief survey about the prospects of an alternative account of modal epistemology, facing the problem that emerged in our discussion of Williamson’s account head on. The central task of such an account is to give an analysis of our epistemic access to constitutive, essentialist facts. Since these facts are supposed to explain modal truths, they should be understood in non-modal terms. Kit Fine’s (1994) account is hence a promising starting point, since his account of essentialist facts is built around the idea that the notion of essence is prior to that of modality.¹³ Specifically, Fine considers propositions that express necessary truths to be true in virtue of the natures of the objects they concern. But what is our epistemic access to these essentialist facts?

Perhaps one of the most interesting accounts that would enable us to overcome the epistemic problem that Williamson also faces is due to Lowe (2008). He suggests that we have previous *a priori* access to essentialist facts—the account builds on his more general metaphysical commitment according to which, in general, essence precedes existence both ontologically and epistemologically. On the epistemic side, the idea is that we can—indeed we must—generally know the essence of a thing before we can know whether that thing exists. In other words, we have to

13. Several philosophers have been inspired by Fine’s account, see for instance Shalkowski (1997), Hale (2002), Oderberg (2007), Lowe (2008), Tahko (2009), and Vaidya (2010).

know that it is possible for a thing of a particular kind to exist before we can determine whether a thing of that kind actually exists. Lowe suggests that transuranic elements constitute a good example that would appear to support this idea, as many of them were only synthesized after their *possible* existence was determined by non-empirical means. In fact, Lowe (2008, 35) thinks that we cannot even ‘talk or think comprehendingly’ about things whose essences we do not know.

Consider the classic case of a bronze statue and the lump of bronze that it is composed of: if they are indeed two different things, as Lowe thinks they are (2008, 46), then no amount of empirical information will help us to determine this. Empirically, the statue and the lump are indistinguishable. It is only because we know what kind of things bronze statues and lumps of bronze are—and specifically knowing what they essentially depend on for their existence and identity—that we can distinguish them. Here we can take advantage of an idea which is related to Williamson’s discussion as well, namely that when we consider cases such as this, we should only make minimal changes to the scenario. We can for instance consider a change in the shape of the bronze statue and the lump of bronze: it seems plausible that the lump could survive such a change without any radical implications, whereas the statue could not. The general idea would then be that as long as the changes we make to the scenario do not involve radical change in the characteristic functions of the things involved, we have succeeded in holding constitutive facts fixed. But since changing the shape of a bronze statue *does* entail radical changes in its characteristic functions, in this case we have failed to hold fixed all constitutive facts. This method should thus reveal to us that the shape of a statue is a part of its essence—a constitutive fact.

Similarly, we could apply this type of method to the case of elements, which is already familiar to us from the discussion concerning Hendry’s work on the nature of elements. Recall Hendry’s suggestion: elements are characterized by their nuclear charge rather than their atomic weight. This has a number of implications on how we understand elements. For instance, hydrogen and its isotope, deuterium, are instances of the same element, since they have the same nuclear charge. This is consistent with the current definition of elements and their isotopes, even though Hendry’s suggestion deviates from the usual line. The core of Hendry’s account concerns a natural defining characteristic of chemical substances, namely whatever it is that gives chemical substances the distinct chemical properties that they have. This fits the idea that there are certain constitutive facts

that impose modal constraints which govern the behaviour of elements, and Hendry identifies nuclear charge to be responsible for the majority of these constraints (for reasons that we need not get into here). We could of course still decide to use the notion of 'element' in such a way that deuterium and hydrogen are distinct elements, but we better have some concept that tracks the similarity between different isotopes of the same element, as the similar chemical properties that these different isotopes manifest suggests that we have discovered a constitutive fact that should be held fixed.

Nathan Salmon's (2005, 258f.) discussion concerning the defining characteristics of elements is also relevant here. Salmon asks whether the following scenario would constitute an empirical discovery: we discover some further defining characteristic for elements apart from their atomic number, and we then encounter two samples of substances which share this new defining characteristic, but differ with regard to their atomic number, that is, have a different number of protons. Well, to a certain extent, Hendry's suggestion *is* such a defining characteristic; if we define elements in terms of nuclear charge, a difference in the number of protons in two sample substances would certainly suggest that they are samples of two distinct substances. Salmon's conclusion is that we cannot determine this strictly by empirical means, or by linguistic analysis, rather, we need some further explanation about what makes certain attributes constitutive, and how we are able to know this.

Both Salmon's and Hendry's discussions follow the familiar pattern that emerged from the previous examples. Hendry argues that nuclear charge is the most promising *possible* candidate for defining elements, and although he is quick to offer historical and empirical considerations in support of his claim, there must be some previous reasons to think that nuclear charge indeed is a constitutive fact. It is in Lowe's work (e.g. 2008) where we find one of the most promising accounts concerning our epistemic access to these facts: we determine the possible candidates (e.g. for the defining characteristics of elements) *a priori*, paying close attention to the modal constraints that they would impose. Then we proceed to evaluate each candidate by taking into account empirical considerations. In the case of elements, as in the case of the bronze statue and the lump of bronze, we should consider the importance and role of each attribute of the thing in question when determining which of its characteristic functions to hold fixed. As Hendry suggests, in the case of elements these (constitutive) characteristic functions plausibly involve the chemical properties of

the elements in question, and nuclear charge would appear to be a good candidate for such a constitutive property.

The alternative account of modal epistemology and the interpretation of constitutive facts sketched in this section does leave a lot of details open, and there are some difficult questions to settle in terms of the epistemology of essence before it is even viable. However, if we accept the Finean account of a non-modal interpretation of essence and Lowe's idea that essence precedes existence, the merits of this account in explaining the nature of constitutive facts are certainly superior to Williamson's, as in his account it is not even clear that we know *which* facts are constitutive, only that we know some constitutive facts. Hence, we can conclude that there are some promising competitors to Williamson's account of modal epistemology, which are arguably also capable of addressing the difficult questions concerning constitutive facts.

REFERENCES

- Bealer, George 1987: "The Philosophical Limits of Scientific Essentialism". In: James E. Tomberlin (ed.), *Philosophical Perspectives 1: Metaphysics*. Atascadero, CA: Ridgeview, 289–365.
- Bird, Alexander 2007: *Nature's Metaphysics*. Oxford: Clarendon Press.
- Brogaard, Berit & Salerno, Joe 2008: "Counterfactuals and Context". *Analysis* 68 (297), 39–46.
- Carroll, John W. 2011: "Laws of Nature". *The Stanford Encyclopedia of Philosophy (Spring 2011 Edition)*, ed. E. N. Zalta, URL = <<http://plato.stanford.edu/archives/spr2011/entries/laws-of-nature/>>.
- Chalmers, David 2002: "Does Conceivability Entail Possibility?". In: Tamar S. Gendler & John Hawthorne (eds.), *Conceivability and Possibility*. Oxford: Oxford University Press, 145–200.
- 2011: "The Nature of Epistemic Space". In: Andy Egan & Brian Weatherson (eds.), *Epistemic Modality*. Oxford: Oxford University Press, 60–107. URL = <<http://consc.net/papers/espace.pdf>>.
- Ellis, Brian 2001: *Scientific Essentialism*. Cambridge: Cambridge University Press.
- Fine, Kit 1994: "Essence and Modality". In: James E. Tomberlin (ed.), *Philosophical Perspectives 8: Logic and Language*. Atascadero, CA: Ridgeview, 1–16.
- Hale, Bob 2002: "The Source of Necessity". In: James E. Tomberlin (ed.), *Philosophical Perspectives 16: Language and Mind*. Atascadero, CA: Ridgeview, 299–319.

- Hendry, Robin F. 2006: "Elements, Compounds, and Other Chemical Kinds". *Philosophy of Science* 73, 864–75.
- Jenkins, Carrie S. 2008: "Modal Knowledge, Counterfactual Knowledge and the Role of Experience". *Philosophical Quarterly* 58, 693–701.
- Lowe, Jonathan E. 2007: "A Problem for a Posteriori Essentialism Concerning Natural Kinds". *Analysis* 67 (4), 286–92.
- 2008: "Two Notions of Being: Entity and Essence". *Royal Institute of Philosophy Supplements* 83 (62), 23–48.
- Nolan, Daniel 1997: "Impossible Worlds: A Modest Approach". *Notre Dame Journal of Formal Logic* 38 (4), 535–72.
- Oderberg, David 2007: *Real Essentialism*. London and New York: Routledge.
- Roberts, John 2010: "Some Laws of Nature Are Metaphysically Contingent". *Australasian Journal of Philosophy* 88 (3), 445–57.
- Roca-Royes, Sonia 2011a: "Conceivability and *De Re* Modal Knowledge". *Noûs* 45 (1), 22–49.
- 2011b: "Modal Knowledge and Counterfactual Knowledge". *Logique et Analyse* 54 (216): 537–52.
- Salmon, Nathan U. 2005: *Reference and Essence*, 2nd ed. New York: Prometheus Books.
- Shalkowski, Scott 1997: "Essentialism and Absolute Necessity". *Acta Analytica* 12 (2), 41–56.
- Shoemaker, Sydney 1998: "Causal and Metaphysical Necessity". *Pacific Philosophical Quarterly* 79 (1), 59–77.
- Swoyer, Chris 1982: "The Nature of Natural Laws". *Australasian Journal of Philosophy* 60 (3), 203–23.
- Tahko, Tuomas E. 2009: "On the Modal Content of A Posteriori Necessities". *Theoria: A Swedish Journal of Philosophy* 75, 344–57.
- Vaidya, Anand J. 2010: "Understanding and Essence". *Philosophia* 38, 811–33.
- Williamson, Timothy 2005: "Armchair Philosophy, Metaphysical Modality and Counterfactual Thinking". *Proceedings of the Aristotelian Society* 105, 1–23.
- 2007a: "Philosophical Knowledge and Knowledge of Counterfactuals". In: Christian Beyer & Alex Burri (eds.), *Philosophical Knowledge—Its Possibility and Scope*. (Grazer Philosophische Studien 74). Amsterdam: Rodopi, 89–123.
- 2007b: *The Philosophy of Philosophy*. Oxford: Blackwell.