

## MODAL KNOWLEDGE AND COUNTERFACTUAL KNOWLEDGE

SONIA ROCA-ROYES

### *Abstract*

The paper compares the suitability of two different epistemologies of counterfactuals — (EC) and (W) — to elucidate modal knowledge. I argue that, while both of them explain the data on our knowledge of counterfactuals, neither can subsume modal knowledge. (EC) would be available only to *extreme haecceitists*. Only (W) — Williamson’s epistemology — is compatible with all counterpossibles being true; something on which Williamson’s account relies. A first problem is that, in the absence of further data for (W) and against (EC), Williamson’s choice of (W) is objectionably biased. A second, deeper problem is that (W) cannot satisfactorily elucidate modal knowledge. Third, from a naturalistic perspective, the nature of this second problem favours (EC) against (W).

### 1. *Key elements of a counterfactual-based account*

For the sake of specificity, I focus on Williamson’s account as a reference point from where the arguments are developed, but the discussion below will illustrate what the main difficulties are for any counterfactual-based account.

Williamson assumes that we possess counterfactual knowledge (p. 141)<sup>1</sup> and sketches an epistemology of counterfactuals motivated by reflecting on the knowability conditions of *everyday counterfactuals* (p. 188). Modal claims are then argued to be logically equivalent to counterfactual claims (p. 157, my labelling):

$$\begin{aligned}(\Box) \quad \Box A &\equiv (\neg A \Box \rightarrow \perp) \\(\Diamond) \quad \Diamond A &\equiv \neg(A \Box \rightarrow \perp)\end{aligned}$$

His epistemology of counterfactuals is intended to apply not only to *everyday counterfactuals* — the motivating data — but also to the instances of the right hand sides of  $(\Box)$  and  $(\Diamond)$  (pp. 163–4). This is of vital importance to subsume

<sup>1</sup> Unless otherwise stated, page numbers are from (Williamson 2007).

the epistemology of modality under an epistemology of counterfactuals. In a naturalistic vein, the subsumption would be permitted thanks to (I):

- (I) The capacity to handle metaphysical modality is an ‘accidental’ by-product of the cognitive mechanisms which provide our capacity to handle counterfactual conditionals. [...] Our capacity for modal thinking cannot be isolated from our capacity for ordinary thinking about the natural world, which involves counterfactual thinking. (p. 162, my labelling)

Before we proceed, let me introduce some terminology. I will call ‘e-counterfactuals’ those counterfactuals that have a metaphysically possible antecedent and a logically consistent consequent. (Although Williamson does not provide a definition of *everyday counterfactuals*, the examples he uses to motivate his epistemology of counterfactuals are all e-counterfactuals.) I will call ‘m-counterfactuals’ the instances of  $(\neg A \Box \rightarrow \perp)$  and  $(A \Box \rightarrow \perp)$ . Finally, ‘counterpossibles’ standardly refers to counterfactuals with impossible antecedent.<sup>2,3</sup>

Williamson’s account has received criticisms from two quarters. First, there are arguments against the logical equivalences his account relies on. The work of Daniel Nolan on impossible worlds (1997) is especially relevant here because impossible worlds open the door to false counterpossibles and thus threaten both  $(\Box)$  and  $(\Diamond)$ . The second line of criticism (in (Jenkins 2008) and (Vaidya 2007)) grants the truth of the equivalences — for the sake of discussion — but complains of an argumentative gap: the logical equivalences do not straightforwardly support a counterfactual-based epistemology of modality.

The discussion below shows the way in which these two lines of criticism are related. I will strengthen the second line by arguing that the argumentative gap is unsolvable, and use those reasons to support Nolan, to some extent.

In §2, I introduce (EC) — a seemingly plausible epistemology of counterfactuals. In §3, I argue that (EC) cannot serve Williamson’s purposes. I then introduce (§4) Williamson’s epistemology, which I call ‘(W)’. In §5, I argue

<sup>2</sup>This is not a *mutually exclusive taxonomy*. Although e-counterfactuals can be neither m-counterfactuals nor counterpossibles, some counterpossibles are m-counterfactuals.

<sup>3</sup>For simplicity and convenience, let this terminology apply only to counterfactuals whose consequents and antecedents are *not* themselves counterfactuals.

that the key difference between (EC) and (W) lies in the evaluation of counterpossibles and object that Williamson's choice begs the question against Nolan. (EC) and (W) are extensionally equivalent as far as the data — e-counterfactuals — are concerned. In §6, I show, beyond Williamson's case, why this threatens claim (I) above. In §7, I argue that (W) cannot serve the purposes either, for independent reasons. In §8, I use those reasons to argue that the threat against (I) is stronger than I anticipate in §6. The last section (§9) concludes that the deficits of both (EC) and (W) raise a challenge for the advocates of *counterfactual-based accounts* of modal knowledge.

## 2. *An epistemology of counterfactuals: (EC)*

Here I shall introduce (EC) — a seemingly plausible epistemology of counterfactuals.<sup>4</sup> (EC) differs from Williamson's (introduced in §4) in just one — crucial — respect, made explicit in due course.<sup>5</sup> Given the vast overlap between (EC) and (W), most of what I use here to sketch (EC) is borrowed from (Williamson 2007).

Suppose that in our world a rock falls down a slope and, instead of ending in the lake at the bottom, it rolls into a bush. Suppose further that (1) is true and that, after evaluation, we come to know that it is so:

(1) If the bush had not been there, the rock would have ended in the lake.

An assumption — granted here and shared by Williamson (p. 141) — is that:

(II) We have non-trivial knowledge of counterfactuals.

The question arises, therefore, as to how we know counterfactuals. To a first approximation, we know counterfactuals by using our imagination. Roughly, to evaluate (1), we imagine a scenario like this:

w<sub>1</sub>: There is the lake, the slope, nothing on the slope. The rock ends up in the lake.

However, this approximate answer is inaccurate. Imagination is typically unconstrained, so an unqualified appeal to imagination does not explain why, in fact, we do not imagine situations like these:

<sup>4</sup> I will restrict myself to counterfactuals whose evaluation uses the "rolling-back" method (p. 150–151). This will be enough for present purposes.

<sup>5</sup> Strictly speaking, '(EC)' and '(W)' are names for *families of epistemologies*. I can skip the details and focus instead on sketches because what interests us here — and what I will exploit in my arguments — is *the difference* between any (EC)-style epistemology and any (W)-style epistemology. For easiness of exposition, I shall ignore this qualification in the remaining of the paper.

w<sub>2</sub>: There is the lake, the slope, no bush, but a massive block of ice instead. The rock rolls into the ice.

w<sub>3</sub>: There is the lake, the slope, no bush, but the laws of nature are different. The rock keeps on floating and never reaches the lake.

Scenarios w<sub>2</sub> and w<sub>3</sub> accommodate the antecedent of (1) but not its consequent. If we considered them when evaluating (1), we would judge it false, but we believe it to be true. Consequently, in counterfactual evaluation, something is constraining our imagination. What exactly these constraints are is something that deserves further investigation, as Goodman's *cotenability* problem shows (p. 143). Yet, in agreement with Williamson, the following appears to be an important way in which imagination is constrained in selecting relevant scenarios:

The default for the imagination in its primary function may be to proceed as realistically as it can, subject to whatever deviation the thinker imposes by brute force: [there], the absence of the bush. (p. 143)

Imagination is also constrained in the way it develops counterfactual suppositions:

Your imaginative exercise is radically informed by your perception of the rock, the slope, and your sense of how nature works. [...] imagination can in principle exploit all our background knowledge in evaluating counterfactuals. (p. 143)

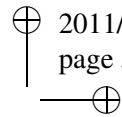
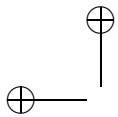
The first constraint parallels *the minimality requirement* on truth-conditions.<sup>6</sup> The second places constraints on imagination that should parallel other constraints on truth-conditions: in terms of worlds, the ones governing the *similarity (or closeness) relation*.<sup>7</sup> In line with this, (EC) offers these (rough) constraints upon imagination:

- (EC-i) Imagination proceeds as realistically as it can. We add the antecedent of the counterfactual at hand and, keeping it, minimally amend our background knowledge in order to preserve consistency.
- (EC-ii) Imagination can exploit all our background knowledge (except for what has been imagined away): e.g., relevant knowledge of the actual scenario and our sense of how nature works.<sup>8</sup>

<sup>6</sup> See (Bigaj 2006, 73) for a statement of this constraint.

<sup>7</sup> Requiring that imagination be constrained by "our sense of how nature works", for instance, amounts to saying that nomically possible worlds are closer to the actual world than counter-nomic worlds.

<sup>8</sup> The epistemic role of (EC-ii) is, therefore, to "discover" the character of the *close* antecedent-worlds. One might want to explore the similarities between (EC-i)–(EC-ii) and Kment's analysis of the closeness relation in (Kment 2006b, 296).



Understood as a minimality constraint, (EC-i) explains why imagination sanctions  $w_2$  and  $w_3$  as irrelevant for the evaluation of (1), whereas (EC-ii) explains why, in all relevant scenarios, we develop the supposition in the way we do — i.e., why we are led to add its consequent.<sup>9</sup> The idea is that exploitation of our background knowledge is somehow responsible for our “various propensities to form expectations about what happens next” (p. 148) and that this tracks counterfactual *truth*.

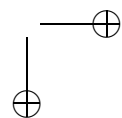
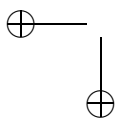
(EC) is *seemingly plausible*. Following constraints (EC-i) and (EC-ii) is likely to give extensionally adequate results — at least as far as the data are concerned. First, if we did not *preserve consistency* (e.g., if we did not imagine away the presence of the bush when evaluating (1)), anything would counterfactually follow from any supposition, and many counterfactuals that we think are false — e.g., *if the bush had not been there, the rock would have exploded* — would come out true. (Similarly, counterfactuals like (1) would also come out true, but for the same wrong reasons.) Second, if we did not *minimally* amend the scenario,  $w_2$  and  $w_3$  would be considered relevant for the evaluation of (1), and this would mislead us into thinking that (1) is false. In both cases, we would lose the connection between *counterfactual evaluation* and *counterfactual knowledge*. Given assumption (II) — that we *do* achieve counterfactual knowledge —, therefore, (EC-i) is sufficiently motivated as an explanation of how we select the relevant scenarios and (EC-ii) as an explanation of how we discover what (else) goes on in them.

With (EC) in mind — as an answer to what is involved in counterfactual evaluation — we can grasp the full content of (A) and (D):<sup>10</sup>

- (A) We assert  $A \Box \rightarrow B$ , when our counterfactual development of the supposition A robustly yields B.
- (D) We deny  $A \Box \rightarrow B$ , when our counterfactual development of A does not robustly yield B (and we do not attribute the failure to a defect in our search).

<sup>9</sup>(EC-i) and (EC-ii) do not work as independently as I might be suggesting. When developing the antecedent of (1), for instance, and before reaching its consequent, we might be led to add that the rock passed *through* the point where (we know) it actually stopped. To preserve consistency, therefore, we will also need to imagine away *that the rock stopped there*. Therefore, (EC-i) and (EC-ii) work together to discover what is/are the relevant scenario(s). What is non-negotiable according to (EC) is that the antecedent should hold, and that *no* inconsistency can.

<sup>10</sup>I borrow also this from Williamson (p. 163), except for the labelling.



### 2.1. Subsuming modal epistemology under counterfactual epistemology

So far, I have introduced (EC) as an epistemology of counterfactuals. To elucidate modal knowledge *via* (EC), we need a "subsumption strategy". Let me then borrow Williamson's and apply it to (EC). (I will then show, in §3, that (EC) cannot serve Williamson's purposes.)

Let me first note that the imaginative exercise involved in the evaluation of *m*-counterfactuals is a *conceivability exercise*.<sup>11</sup> (EC) has been motivated using *e*-counterfactuals. For the subsumption to be successful, (EC) must apply to *m*-counterfactuals too:

- (III) Such conceivability and inconceivability [m-counterfactual evaluation] *will be subject to the same constraints, whatever they are, as counterfactual conditionals in general*, concerning which parts of our background information are held fixed. (pp. 163–4; my labelling and emphasis)

With the same (EC)-constraints in mind, therefore, (A) and (D), together with ( $\Box$ ) and ( $\Diamond$ ), provide a counterfactual-based epistemology of modality. As an illustration:

- (A $\Box$ ) We assert  $\Box A$  when our counterfactual development of the supposition  $\neg A$  robustly yields a contradiction.

(I leave (A $\Diamond$ ), (D $\Box$ ) and (D $\Diamond$ ) to the reader.)

(III) is crucial to support (I) — the claim that our capacity for modal thinking cannot be isolated from our capacity for counterfactual thinking. The agreed data are all *e*-counterfactuals. These data, therefore, even when under assumption (II) — that we have counterfactual knowledge — only imply that we have a capacity for *e*-counterfactual knowledge. However, modal claims are equivalent to *m*-counterfactuals. A prerequisite to securing (I), therefore, is that the capacity for m-counterfactual knowledge cannot be isolated from our capacity for e-counterfactual knowledge, and this is (III)'s crucial role.

If *different set of constraints* about imagination were to apply in the cases of e- and m-counterfactuals, nothing would guarantee that whoever has the cognitive capacity to follow — or even internalize — one of these two sets also has the cognitive capacity to follow the other. As a result, our capacity for m-counterfactual knowledge *could* be isolated from our capacity for e-counterfactual knowledge and, even when granting the logical equivalences ( $\Box$ ) and ( $\Diamond$ ), this would jeopardize the project of subsuming modal epistemology under (*e*-)counterfactual epistemology.

<sup>11</sup> For more on this, see (Roca-Royes, 2011). This explains why Williamson thinks of *conceivability* and *inconceivability* as providing, respectively, tests for possibility and impossibility (p. 163).

3. (EC) applied to *m*-counterfactuals

Despite the fact that (EC) is, for all we know, a seemingly plausible epistemology of counterfactuals, it cannot be Williamson's epistemology for it does not provide *general* constraints on imagination *that Williamson could endorse*. I shall show here that, if *m*-counterfactuals are equivalent to modal claims — as Williamson believes — then, by Williamson's lights, (EC) over-generates possibility claims.

Gold is the element with atomic number 79. Now consider (2):

(2)  $\diamond$  [Gold has an atomic number other than 79]

Assume that (2) is false — nothing essential depends on this. By  $\diamond A \equiv \neg(A \Box \rightarrow \perp)$ , (2) is logically equivalent to:

(2 $\equiv$ )  $\neg$ (Gold has an atomic number other than 79  $\Box \rightarrow \perp$ )

Yet, according to (EC), (2 $\equiv$ ) is true. Let us see this by seeing that its negation is false:

(3) (Gold has an atomic number other than 79  $\Box \rightarrow \perp$ )

According to (EC-i), we have to add the antecedent of (3) to our background knowledge and, keeping it, minimally amend the background knowledge *in order to preserve consistency*. Trivially, therefore, no contradiction can possibly follow counterfactually from the supposition. So (3) is false. Since (3) is false, (2 $\equiv$ ) is true, and so is — under the assumption of ( $\diamond$ ) — its logical equivalent (2).

As a result, the outcome of (EC) is that *gold could have an atomic number other than 79*. This example generalizes to almost any *A*. The only exceptions are those *A*'s that are themselves logically contradictory — for, in those cases, there is no way of both keeping *A* and preserving consistency (since we do not have consistency to begin with). The generalization is therefore this:

(Gr) If (EC) is correct as a *general* epistemology of counterfactuals, then, *if* ( $\Box$ ) *and* ( $\diamond$ ) *hold*, then, anything that is not logically contradictory is metaphysically possible — that is, *extreme haecceitism* is true.

Since denying ( $\Box$ ) and ( $\diamond$ ) is a non-starter for Williamson, the dialectically relevant generalization is that if (EC) is correct, anything that is not logically contradictory is metaphysically possible. Whether this is a bad result depends on what our essentialist views are. For Williamson, who believes in (substantial) essential properties<sup>12</sup> and, therefore, rejects *extreme haecceitism*, it is an unwelcome result, and this is why (EC) cannot serve his

<sup>12</sup>Without further ado, Williamson believes that having atomic number 79 is essential to gold (164).

purposes. The generalization shows — as anticipated — that nothing essential depends on the essentialist assumption above. The dialectical impact of (Gr) is this: it shows that endorsement of both (EC) and  $(\Box)\text{--}(\Diamond)$  would commit us to the unpopular *extreme haecceitism*.

#### 4. Williamson's Epistemology of Counterfactuals: (W)

What are, then, the *general* constraints Williamson has in mind?

If we know enough chemistry, our counterfactual development of the supposition that gold is the element with atomic number [other than] 79 will generate a 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. (p. 164)<sup>13</sup>

We can and must vary some items of our knowledge. However, by the argument for (Gr) above, we cannot vary *everything* that is incompatible with the truth of the antecedent of the counterfactual at hand. Constitutive facts *must* be held fixed, even if at the expense of consistency. In the argument above, therefore, we were wrong, according to Williamson, in imagining away the fact that gold is the element with atomic number 79. For, if this is a constitutive fact, we should hold it fixed. If we hold it fixed, a contradiction *does* follow counterfactually from the antecedent, for the antecedent is the negation of this constitutive fact that, by holding fixed, and according to (EC-ii), we can exploit in counterfactual evaluation.

These are, roughly, Williamson's *general constraints*:<sup>14</sup>

- (W-i) Imagination should proceed as realistically as it can. We should add the antecedent of the counterfactual at hand and, keeping it, minimally amend our background knowledge, *aiming to preserve consistency, but never imagining away constitutive facts*.<sup>15</sup>

<sup>13</sup> The square brackets are my addition. Reading the quote without their content makes it clear that the omission in (Williamson 2007) is a typo.

<sup>14</sup> Footnote 9 applies to (W-i) and (W-ii) too.

<sup>15</sup> One might want to explore the similarities between (W-i)–(W-ii) and Kment's analysis of the closeness relation in (Kment 2006a, 287).

In relation to footnote 8, one should note the tension between (Kment 2006a, 287) and (Kment 2006b, 296), especially in the context of comparing (EC) and (W).



(W-ii) Same as (EC-ii).

(W) is (stipulated to be) exactly like (EC) except for the — crucial — fact that (W-i) replaces (EC-i). It is because of this difference that, in developing counterfactual suppositions, we will arrive at contradictions more often by following (W) than by following (EC), thereby escaping (Gr). Only (W), therefore, has a chance to serve Williamson's purposes. However, we know that — and *why* — Williamson is aiming at *general* constraints. Therefore, we need to see now whether (W) behaves appropriately also with respect to *e*-counterfactuals. I will show next that we have no reason to be less optimistic here than we were, in §2, about (EC).

The only thing (EC) and (W) differ on is whether we can imagine constitutive facts away. Since all *e*-counterfactuals have possible antecedents, none of their antecedents will challenge (contradict) constitutive facts. Therefore, even when following (EC-i), constitutive facts are held fixed when evaluating *e*-counterfactuals. This is also true — and trivially so — when following (W-i). Therefore, the background knowledge left available to develop *e*-counterfactual suppositions will be the same in both cases. (EC) and (W) do not differ in how suppositions are developed ((EC-ii) *is* (W-ii)). Consequently, *other things being equal*, these two epistemologies will not differ in outcome when it comes to *e*-counterfactuals.

In sum, for someone who believes in ( $\Box$ ) and ( $\Diamond$ ) — e.g., Williamson — (W) might get extensionally adequate results throughout, but (EC) is at most extensionally adequate in relation to *e*-counterfactuals. Since Williamson relies on ( $\Box$ ) and ( $\Diamond$ ), he needs an epistemology along the lines of (W). Mere extensional adequacy, however, is not sufficient for epistemic adequacy and in §§7–8 I will present some concerns about (W)'s epistemic adequacy. Before this, however, it is convenient to see why, given the available data, endorsing (W) begs the question against Nolan.

5. (EC) vs. (W): Counterpossibles

(EC) and (W) disagree on (3) — an *m*-counterfactual. However, there are *m*-counterfactuals on which they agree. Looking at (EC-i) and (W-i), it is easy to see that, other things being equal, (EC) and (W) agree on the evaluation of those counterfactuals *whose antecedents do not challenge constitutive facts*, and some *m*-counterfactuals are of this kind. Under the assumption that the presence of the bush is not a constitutive fact, they agree, for instance, on sanctioning (4) as false:

(4) (The bush is not there  $\Box \rightarrow \perp$ )

It is not the case either that all counterfactuals they disagree on are *m*-counterfactuals. They disagree on those counterfactuals *whose antecedents challenge constitutive facts*, and some counterfactuals that satisfy this condition are not *m*-counterfactuals. Assuming — for the sake of the example — that *biological origins* are essential, they disagree on (5), which is not an *m*-counterfactual:

- (5) If the parents of Queen Elizabeth II had been Mr. and Mrs. Truman, she would have been a dinosaur.

*Reason:* Assume Essentiality of Origins. Assume further that the parents of Queen Elizabeth II are George VI and Elizabeth Bowes-Lyon. Jointly, these assumptions imply that the antecedent of (5) contradicts a *constitutive fact*.<sup>16</sup> According to (EC), when evaluating (5) we imagine away our knowledge about the parents of Queen Elizabeth II. If — as we might well assume — our sense of how nature works includes that a dinosaur cannot result from two-way-human reproduction, then, by (EC-ii), (5) will come out false. By contrast, according to (W), we cannot imagine away our knowledge about the parents of Queen Elizabeth II (for, by assumption, this is a constitutive fact). This contradicts the supposition that the parents of Elizabeth II are Mr. and Mrs. Truman. Therefore, the antecedent in (5) counterfactually implies everything and (5) is (vacuously) true.

The key difference between (EC) and (W) hinges, therefore, on *counterpossibles*: only according to (EC) are some counterpossibles false. (Indeed, on any understanding of what a constitutive fact is, a counterpossible is a counterfactual whose antecedent challenges some constitutive fact.)

The equivalences ( $\Box$ ) and ( $\Diamond$ ) require that all counterpossibles are true. Of the two options — (EC) *vs.* (W) — therefore, Williamson made the convenient choice. However, it might well be *objectionably* biased. The data Williamson uses to motivate (W) are all *e*-counterfactuals, and (W) and (EC) agree on *e*-counterfactuals. The data are, therefore, neutral with respect to these two epistemologies. A debate is then required as to whether (W) or (EC) is correct. Parallel to this, there is the (old) Nolan-Williamson debate on counterpossibles. That Williamson should win this second debate is essential to the success of his project, for false counterpossibles would invalidate the logical equivalences he relies on. However, now that we know what the key difference between (W) and (EC) is, we see that the Nolan-Williamson debate on counterpossibles is no different from the new debate on (W) *vs.* (EC). Agreement on the truth-value of counterfactuals like (3) or (5) would help here. Unfortunately, neither (3) nor (5) are amongst the agreed data. In the absence of new, disambiguating data, Williamson's

<sup>16</sup>This example is adapted from one in (Kripke 1972).

choice is objectionable because it begs the question against someone like Nolan: the data are compatible with (EC-i).

Williamson (2007, Ch. 5, §6) deals with the objection that there might be false counterpossibles. There, he examines some potential counterexamples and concludes that "*the case for false counterpossibles looks feeble*" (175). I am closer to Nolan on this. (I think that (5) is false and that (EC) broadly-adequately describes my evaluation procedure.) However, we have enough evidence that this *counterexample/rebuttal* route leads nowhere, so I will try a different route. I argue (§6) that the mere fact that (EC) is, for all we know, a live alternative, threatens counterfactual-based epistemologies. In §7, I argue that, independently of this threat, (W) cannot ground an epistemology of modality. The reasons in §7 are then used, in §8, to favour (EC) and, as a result, make the threat in §6 stronger.

#### 6. *The threat from (EC): A dilemma*

For all we know, there is no problem *with* (EC). A problem arises, however, *from* the theoretical possibility that (EC) is an adequate epistemology of e-counterfactuals.

The data that supports (II) — the assumption that we have non-trivial knowledge of counterfactuals — are ambiguous between (EC) and (W). Consequently, (EC) is no less a live option than (W). This threatens counterfactual-based epistemologies — beyond (Gr) from §4 — because the (potential) adequacy of (EC) opens the door to a dilemma which would ultimately undermine (I) — the claim that our capacity for modal thinking is an accidental by-product of our capacity for counterfactual thinking.

For, *if (EC) turns out to be the correct epistemology for e-counterfactuals* (the ones that motivate (II)), *then*, either it generalizes or it does not.

- (IV) If (EC) generalized to m-counterfactuals and counterpossibles (as Nolan, but not Williamson, could grant), then ( $\Box$ ) and ( $\Diamond$ ) would be false.<sup>17</sup> In the absence of these equivalences, knowledge of m-counterfactuals need not amount to, even less be, modal knowledge. Therefore, our capacity to handle m-counterfactuals need not amount to a capacity to handle metaphysical modality.
- (V) If (EC) did not generalize to m-counterfactuals and counterpossibles, then (III) — the claim about there being general constraints — would

<sup>17</sup> Given (Gr), the qualification here should be *unless extreme haecceitism (EH) is true*. Because (EH) is highly unpopular, we can ignore this qualification without significant dialectical harm.

be false. For the reasons in §2.1, therefore, our capacity to handle *e*-counterfactuals would not necessarily bring, as an accidental by-product, the capacity to handle *m*-counterfactuals and counterpossibles and, consequently, the capacity to handle metaphysical modality.

In neither case would the capacity to handle metaphysical modality follow from the capacity to handle *e*-counterfactuals.

This dilemma is conditional upon (*EC*) turning out to be the adequate epistemology for *e*-counterfactuals. What we need to ask now is whether there are further considerations that, by favouring (*EC*) against (*W*), could allow us to draw stronger, non-conditional conclusions. The problems with (*W*) that I will present in §7 will provide some of these considerations (on which I elaborate in §8).

### 7. The problems with (*W*)

We should evaluate (*W*) against two different, yet related, goals. First, there is the goal of elucidating modal knowledge *somehow*. Second, there is the aim of subsuming modal knowledge under (*W*). Achieving the latter goal might be intended as a way of achieving the former, and this is Williamson's strategy. So let me first evaluate (*W*) in relation to the second goal and, if the conclusion is that this goal is achievable, I shall evaluate the extent to which achieving it would amount to achieving the first one.<sup>18</sup>

One might be tempted to think that (*W*) cannot subsume modal knowledge. The distinctive feature of (*W*) is that it requires us to hold fixed *constitutive facts*. Furthermore, for our counterfactual judgements to amount to counterfactual *knowledge*, it is not enough that we merely happen to hold fixed the right things — our counterfactual judgements would be (extensionally) correct in this case, but hardly knowledge. We need to hold them fixed *knowledgeably*. This seems to require knowledge of what the constitutive facts are.<sup>19</sup> Given this, the argument against the possibility of a subsumption — the second goal — would come from the fact that, for some people (those who endorse the *modal account* of the notion of essence<sup>20</sup>),

<sup>18</sup> In (Roca-Royes, 2011), I present a criticism of Williamson's account similar to the one in this section. There, however, the emphasis is on Williamson's account as a *conceivability-based* one. Here, instead, I focus on the problems that derive from the account being a *counterfactual-based* account.

<sup>19</sup> For an objection to this and my reply, see (Roca-Royes, 2011, §9).

<sup>20</sup> See (Fine 1994) for its characterization and Fine's objections to it.

constitutive facts *are* modal facts. If this is so, (W) implies that to obtain modal knowledge through counterfactual evaluation we must have prior modal knowledge. This prior modal knowledge would be a *pre-condition* for counterfactual knowledge and, as such, it could not be the outcome of counterfactual evaluation. Consequently, (W) could not subsume this kind of modal knowledge in a non-circular manner.

Williamson (2007, Ch. 5, §6) deals with this potential charge of circularity and, appealing to Fine's arguments against the *modal account* of the notion of *essence* (1994), he argues, rightly, that he can escape it. Roughly, Fine argues that modal facts are ontologically consequential upon, but not the same as, constitutive facts. This opens the door for a parallel analysis of modal knowledge as consequential upon, but not equivalent to, constitutive knowledge, and Williamson strongly suggests that this is the analysis he favours (p. 170).

Fine's arguments, therefore, leave room for a non-circular subsumption of modal knowledge by (W). I consequently grant that the second goal is achievable. The question is whether the achievement of this second goal is a way of achieving the first — elucidating modal knowledge — and I shall answer this in the negative.

If counterfactual evaluation is to be knowledge-conducive, (W) requires, as suspected above, prior constitutive knowledge (which, I am granting, is not modal). For us to *project* "constitutive matters such as atomic numbers into counterfactual suppositions" (p. 170) — i.e., for us to follow the rule implied by (W) — we need to know — if counterfactual evaluation is to be *knowledge-conducive* — that facts about atomic numbers are constitutive matters. This must generalize to any constitutive fact, or we will get extensionally wrong results by *over-generating* possibilities. Conversely, if Humphrey having five fingers on his left hand is *not* constitutive of him, we must (knowledgably) *not* hold it fixed when evaluating "*Humphrey has six fingers on his left hand*  $\Box \rightarrow \perp$ ", or we will get extensionally wrong results by *under-generating* possibilities.

A first problem is that, while knowledge of atomic numbers may not be problematic, knowledge of atomic numbers is not, by itself, knowledge that atomic numbers are constitutive, and it is not obvious that we possess constitutive knowledge. More importantly, if we do possess constitutive knowledge, it is not obvious what are the *cognitive mechanisms* — as they are referred to in (I) — by means of which we are able to, more or less reliably, tell apart constitutive and non-constitutive facts. All we have been told is that, according to (W), these cognitive mechanisms are amongst the "cognitive mechanisms that provide our capacity to handle counterfactual conditionals" (p. 162). To spell out completely the knowability conditions of counterfactuals, therefore, (W)-endorsers should elucidate constitutive knowledge; especially because some are sceptical about it.

The second, deeper problem is that the explanatory deficit just identified is not something one can repair without collateral damage. Its very nature jeopardizes the strategy of achieving the first goal via the achievement of the second. The ultimate aim of the project is to elucidate modal knowledge. With respect to this goal, the manoeuvres made above to avoid the charge of circularity will be of no use. For Fine's ontological divorce between the modal and the constitutive does not amount to a divorce in epistemological worries. Far from it: the ontological divorce shows that we should be more careful in how we describe things. The so-called 'epistemic challenge in modality' is and always has been a challenge concerning both *modal* and *constitutive* knowledge. However, (W) is not even the beginning of an answer to the question about the knowability conditions of constitutive facts. This means that (W) does not address — let alone satisfactorily — a substantial part of our original epistemic challenge. Furthermore, and this is the crucial, jeopardizing part, we know that constitutive knowledge is not explainable in terms of counterfactual knowledge. *That* would be inescapably circular, since the former is a pre-condition for the latter. (W), therefore, cannot ground a counterfactual-based account of modal knowledge — 'modal' understood theory-neutrally: including constitutive knowledge.

#### 8. *Strengthening the threat of the dilemma*

I will now exploit the problem in §7 to argue that, as a *general* epistemology, (W) is less plausible than (EC) — especially from a naturalistic perspective, which Williamson shares (§§1–2).

As a general epistemology (as Williamson intends it), (W) implies that, even when evaluating *e*-counterfactuals like (1), we (knowledgeably) hold fixed constitutive facts. However, is this *what it takes to know e*-counterfactuals? Compare: According to (EC), our reason for imagining away the presence of the bush is that *it contradicts the antecedent*. By contrast, according to (W), our reason must be that *it contradicts the antecedent and is not a constitutive fact*.

We might well complain that (W) requires too much of the *folk* counterfactual-evaluator for it to be plausible as a naturalistic epistemology of *e*-counterfactuals. Specially because "humans evolved under no pressure to do philosophy" (p. 136), and the less demanding (EC) already gives us everything we are under pressure to obtain; namely, true judgements — and arguably knowledge — of *e*-counterfactuals.<sup>21</sup>

<sup>21</sup> When sketching his epistemology of counterfactuals (Ch. 5, §3) — which he does by reflecting on *e*-counterfactuals — Williamson does not mention constitutive facts. This raises the suspicion that (the postulation of) a capacity for judgements about the constitutive is not justified by the data.

I suggest, therefore, that, from a naturalistic perspective, (EC) is more plausible for *e*-counterfactuals than (W). Let us call the cognitive mechanisms implied by (EC), 'CM<sub>EC</sub>', and let us call the ones implied by (W) 'CM<sub>W</sub>'. CM<sub>EC</sub>, *unlike* CM<sub>W</sub>, do not imply a capacity for constitutive knowledge. If, as Williamson suggests, all that evolution *requires* is that we are (non-accidentally) able to handle *e*-counterfactuals, the naturalistic threat is that the emergence of CM<sub>W</sub> (or of CM<sub>W</sub> *minus* CM<sub>EC</sub>, in case of overlap) is quite mysterious.

We are now closer to being able to formulate non-conditionally the dilemma in §6. (Not 'we are able' because I am not claiming that (EC) is correct. Yet, the reason why (EC) is more plausible sufficiently motivates the claim that the correct epistemology will be relevantly similar to (EC) in this respect: it will require no knowledge of constitutive facts. Therefore, conditionals analogous to (IV) and (V) will still hold for such an epistemology.) As a result, the threat against (I) — the claim that our capacity for modal thinking is an accidental by-product of our capacity for counterfactual thinking — is now more pressing.

### 9. Concluding remarks

I have compared the suitability of (EC) and (W) to elucidate modal knowledge. Because of the dilemma, (EC) cannot. Because of the problem in §7, (W) cannot either. The problems with (EC) and (W) raise a challenge for those who endorse *counterfactual-based accounts*: that of finding an epistemology for counterfactuals relevantly dissimilar to both (EC) and (W).<sup>22</sup>

As an epistemology of *e*-counterfactuals, the paper favours (EC). Does this contribute to the Williamson-Nolan debate on counterpossibles in favour of Nolan? Not *necessarily*. Williamson needed *general constraints* upon imagination — much of his argument depends on (I). However, everything developed here is compatible with (EC) being correct for *e*-counterfactuals and all counterpossibles being vacuously true. To endorse both things consistently we simply need not to extend (EC) to counterpossibles, and nothing here commits us to doing so.<sup>23</sup> Still, if *theoretical unity* is given any weight,

<sup>22</sup>I cannot extrapolate here to other counterfactual-based accounts, like Hill's (2006), or Kment's (2006a). However, we find in their accounts elements closely related to Williamson's (III) and (W-i), and these similarities offer the grounds for the extrapolation. (See (Hill 2006, 224 and 230) and (Kment 2006a, 284 and §7).)

<sup>23</sup>This might strengthen the objection against his account that Williamson considers in p. 171. We know (from §5), that the generalization at issue is not a generalization to *m*-counterfactuals, but to *counterpossibles*; and what is special with counterpossibles are their *antecedents*.

what has emerged here *might* offer a route — different from the *counterexample/rebuttal* one — in support of Nolan and against ( $\square$ ) and ( $\diamond$ ).<sup>24</sup>

School of Arts and Humanities  
University of Stirling  
Stirling FK9 4LA  
Scotland  
UK

E-mail: [sonia.rocaroyes@stir.ac.uk](mailto:sonia.rocaroyes@stir.ac.uk)

#### REFERENCES

- Bigaj, Tomas (2006). *Non-Locality and Possible Worlds. A Counterfactual Perspective on Quantum Entanglement*. Heusenstamm: Ontos Verlag.
- Fine, Kit (1994). Essence and Modality. In J. Tomberlin (Ed.), *Philosophical Perspectives 8: Logic and Language*. Atascadero, CA: Ridgeview.
- Jenkins, Carrie (2008). Modal Knowledge, Counterfactual Knowledge and The Role of Experience. *Philosophical Quarterly*, 88(233), 693–701.
- Hill, Christopher (2006). Modality, Modal Epistemology, and the Metaphysics of Consciousness. In S. Nichols (Ed.), *The Architecture of Imagination* (pp. 205–236). Oxford University Press.
- Kment, Boris (2006a). Counterfactuals and the Analysis of Necessity. *Philosophical Perspectives*, 20, 237–302.
- Kment, Boris (2006b). Counterfactuals and Explanation. *Mind* 115/458: 261–309.
- Kripke, Saul (1972). *Naming and Necessity*. Harvard University Press, Cambridge, Massachusetts.
- Lewis, David (1973). *Counterfactuals*. Blackwell Press.
- Nolan, Daniel (1997). Impossible Worlds: a modest approach. *Notre Dame Journal of Formal Logic*, 38, 535–72.
- Roca-Royes, Sonia (2011). Conceivability and *de re* modal knowledge. *Noûs*, 45/1, 22–49.
- Vaidya, Anand (2007). The epistemology of modality. *Stanford Encyclopedia of Philosophy*, Edward N. Zalta (Ed.), URL: <http://plato.stanford.edu/entries/modality-epistemology>.
- Williamson, Timothy (2007). *The Philosophy of Philosophy*. Oxford, Blackwell.

<sup>24</sup> Many thanks to an anonymous referee for this journal, my colleagues in Stirling, the members of the LOGOS Group (Barcelona), and the audiences of the *SPA Winter Meeting 2008* (Stirling), the *VAF Conference 2009* (Tilburg), the *SOPHA 2009* (Geneva) and a Logos Colloquium at the University of Girona. Many thanks also to Thomas Kroedel, Andrea Sauchelli, Margot Strohminger, Tuomas Tahko and Anand Vaidya.