ARE COUNTERPOSSIBLES EPISTEMIC?

BY

DANIEL DOHRN

Abstract: It has been suggested that intuitions supporting the nonvacuity of counterpossibles can be explained by distinguishing an epistemic and a metaphysical reading of counterfactuals. Such an explanation must answer why we tend to neglect the distinction of the two readings. By way of an answer, I offer a generalized pattern for explaining nonvacuity intuitions by a stand-and-fall relationship to certain indicative conditionals. Then, I present reasons for doubting the proposal: nonvacuists can use the epistemic reading to turn the table against vacuists, telling apart significant from spurious intuitions. Moreover, our intuitions tend to survive even if we clear-headedly intend a metaphysical reading.

In recent years, counterpossibles, that is subjunctive conditionals with metaphysically impossible antecedents, have been intensely debated. On the one hand, there is the orthodox view, established among others by Lewis (1973) and Williamson (2007):

(VACUISM): counterpossibles are vacuously true.

On the other hand, there is a growing literature of dissenters (nonvacuists). The dissenters point to our intuitions that some counterpossibles are true and some are false. Here is a standard example (Nolan 1997, p. 544, varied by Bjerring 2014):

1   If Hobbes had squared the circle, then the mathematical community at the time would have been surprised.
2   If Hobbes had squared the circle, then sick children in the mountains of Afghanistan at the time would have been thrilled.

(1) seems true, and (2) seems false.

My aim is to critically assess a new proposal that promises to reconcile the two opposing views. I call it the epistemic proposal (EP): *nonvacuous counterpossibles are epistemic*. To unpack a little bit, past subjunctives which in the usual *metaphysical* reading have a metaphysically impossible antecedent are either to be read *epistemically* (henceforth 'epistemic counterpossibles' or 'counterpossibles read epistemically') or vacuously true (Vetter 2016a, p. 2693; Williamson 2017, p. 217).[1] Antivacuity intuitions can be explained by the epistemic reading. Thus, we may have the cake and eat it, combining vacuous truth (in the metaphysical reading) with accepting the intuitions for nonvacuous truth (in the epistemic reading).

In assessing the proposal, I introduce the epistemic reading of subjunctive conditionals (Section 1). I try to spell out the proposal by a psychologically credible mild error theory (MET) (Section 2). I discuss arguments in favour of the resulting MET (Section 3), a limited argument from referential opacity (Section 3.1), and a more general argument from the stand-and-fall-relationship between indicatives and subjunctives (Section 3.2). I consider counterarguments to MET (Section 4): nonvacuists can turn the tables on the divide-and-conquer strategy, using the strategy to account for certain metaphysically weird counterpossibles (Section 4.1). The error theory cannot be mild: antivacuity intuitions persist even if we knowingly aim at a metaphysical reading (Section 4.2). I close with the general recommendation of checking intuitions for disturbing noise from the epistemic reading (Section 5).

1. Epistemic and metaphysical readings of subjunctive conditionals

Consider the usual example illustrating the distinction between past indicative and past subjunctive ('had-would') conditionals or *counterfactuals* (Adams 1970):

3 If Oswald did not kill Kennedy, someone else did.
4 If Oswald had not killed Kennedy, no one else would have.

The difference between (3) and (4) can be explained by the distinction between epistemic ('perhaps Oswald did not kill Kennedy') and 'metaphysical' possibilities ('Oswald could kill or not kill Kennedy'). Roughly, indicative conditionals trade in epistemic, whereas subjunctive conditionals trade in metaphysical possibilities.

In assessing (3), we make room for the slight epistemic possibility that Oswald did not actually do it. We consider whether the epistemically possible scenario is also one in which someone else killed Kennedy. In pondering

the scenario, we have to suspend those of our beliefs which rule it out, in particular the belief that Oswald and no one else did it. We retain the independently supported belief that someone killed Kennedy. We conclude that someone else than Oswald must have killed Kennedy.

In assessing (4), we consider some nonactual ways the world could have been. Being nonactual, they are perfectly compatible with all we know about the actual world, including the fact that Oswald actually did kill Kennedy. We only consider whether, in a relevant scenario, say one which minimally diverges from the actual world in bringing about the antecedent, someone else killed Kennedy (at the same time as Oswald). Actually, there was no backup killer. In the antecedent scenario closest to actuality, there was no backup killer either. Hence, no one else killed Kennedy. Call this the *metaphysical use* of past subjunctives.

However, it has more recently been noted that there is also an *epistemic use* of past subjunctives. A detective having worked on the Oswald case may get enmeshed in the following dialogue:

'"You had already got Oswald. Why did you continue the investigation?" "We weren't sure it was Oswald; if it hadn't been Oswald, it would have been someone else in the crowd who killed Kennedy"' (Edgington 2011, p. 239).

In this context, (4) seems false. Asked why she acted as she did, the detective remembers the past epistemic situation. When she made her decision, there were two salient epistemic possibilities. Either Oswald or someone else did it. Since she could not rule out the second alternative, the detective pursued her investigation.

From this and other examples, several authors have concluded that there is an epistemic as contrasted to the metaphysical use of subjunctive conditionals (Edgington 2011; Khoo 2015; Vetter 2016a, 2016b).[2] Again, the distinction can be put in terms of epistemic and metaphysical possibilities. The detective reports having pondered the epistemic possibility that Oswald might not have done it. I shall provide the following working characterization: introduce an *epistemic base*. The epistemic base includes all epistemic possibilities (centred worlds) that are open from a relevant vantage point (cf. Khoo 2015, p. 4). An indicative conditional assessed now is true precisely if all antecedent worlds in the epistemic base determined by our *current vantage point P@* are consequent worlds.[3] The subjunctive mood conveys *remoteness*. The epistemic reading of a subjunctive presupposes a contextually salient epistemic vantage point $P*$ that normally *differs* from our present one. The antecedent describes an epistemic possibility that is open as seen from $P*$ but not necessarily from $P@$. A counterfactual, read epistemically, is true precisely if all antecedent possibilities relative to $P*$ are possibilities such that the consequent is true.

I shall add a heuristic to be motivated in Section 3.2.2: simulate yourself at a suitably chosen point in time when it was an open question whether the antecedent was true. Add belief in the antecedent. Consider whether you

are disposed to include belief in the consequent into the resulting belief system.

Epistemic readings of past subjunctives are not confined to remote past epistemic possibilities. Consider the following, originally due to van Fraassen:

'the conjuror holds up a penny and claims he got it from the boy's pocket. "That didn't come from my pocket", says the boy. "All the coins in my pocket are silver.

> 5 If that had come from my pocket, it would have been a silver coin."'
>   (Edgington 2011, p. 239)

Read epistemically, the boy simply considers the epistemic possibility that the coin has come from his pocket in order to rule it out by stating his knowledge that all coins in his pocket are silver. In contrast, a metaphysical reading is problematic. Consider a *literal* metaphysical reading ('*that* particular coin' as a rigid designator): assuming that material constitution is necessary, the consequent is impossible, but the antecedent is possible. The coin *could* have been slipped into the boy's pocket before, but that would not have made it a silver coin. Hence, (5) is false. An alternative would be a non-literal metaphysical reading: 'that coin' is paraphrased roughly by the description 'the coin the conjuror holds up'. We consider a situation in which the conjuror holds up a silver coin, which has been taken from the boy's pocket. However, this move seems *ad hoc* without a systematic paraphrase strategy (cf. Brogaard and Salerno 2013, p. 645). In sum, an epistemic reading is a salient alternative.

The coin example provides some evidence that we do not have to dwell on remote past epistemic possibilities to get an epistemic reading. It does not take a lot of context either. There may even be past subjunctives that are read epistemically 'out of the blue'. Yablo (2002) argues that the following conditionals are nonvacuous:

> 6 If Hesperus had turned out not to be Phosphorus, Hesperus would not be Phosphorus.
> 7 If Hesperus had turned out not to be Phosphorus, Hesperus would still be Phosphorus.

To Yablo, (6) is true, and (7) is false. If this is right, one explanation is that (6) and (7) are to be read epistemically. We do not even need a minimal context, the epistemic locution 'turn out' in the antecedent seems sufficient to trigger the epistemic reading. We consider the epistemic possibility that Hesperus was not Phosphorus from some relevant vantage point. This epistemic possibility is compatible with the metaphysical necessity that Phosphorus is Hesperus. I admit that nonvacuists may explain Yablo's intuitions by the

nonvacuity of (6) and (7). Still, (6) and (7) may lend some support to an epistemic reading 'out of the blue' if we follow Yablo in assuming that 'would have turned out' – conditionals are not simply normal subjunctive conditionals, but something in between indicatives and (metaphysically read) subjunctives (Yablo 2002, p. 453). This assumption does not beg the question against nonvacuism.

I have focused my attention on past subjunctives to avoid uncertainties about the counterfactual status of present-tense subjunctives. Still, it seems highly plausible that, whatever else we may have to say on present-tense subjunctives, the arguments for distinguishing metaphysical and epistemic readings apply to them as well. We can devise present-tense versions of (3)–(7), which can be read epistemically. Back in the 60s, the detective may have reasoned:

8 If Oswald were not the killer, someone else would be.

The boy may reason:

9 If that came from my pocket, it would be a silver coin.

The same for Yablo's 'turn out':

10 If Hesperus turned out not to be Phosphorus, Hesperus would not be Phosphorus.
11 If Hesperus turned out not to be Phosphorus, Hesperus would still be Phosphorus.

For the sake of simplicity, I shall mainly concentrate on past subjunctives and leave present-directed and future-directed subjunctives to future debate. EP might depend on the possibility of reading them epistemically.

In sum, there are good reasons to assume that some subjunctive conditionals have to be read epistemically. Of course, this assumption leads to many questions: what triggers the two readings? Is their difference semantic or pragmatic? Are they derivative of each other or of some common 'root meaning'? Such questions go beyond this article. Instead, I shall take the intuitive distinction for granted and consider what follows for the debate on counterpossibles.

2. *Fleshing out the epistemic proposal: Mild error theory*

I shall now try to flesh out EP. Since nonvacuists normally do not accept that nonvacuous counterpossibles are epistemic, EP is only to a certain

extent reconciliatory. It involves an error theory, explaining away vacuist intuitions as based on neglecting the distinction of epistemic and metaphysical readings. Error theories are uncharitable. Hence, they should be subject to standards of credibility.

Here is an idea for such a standard: Yablo (2006, p. 334) suggests a 'psychoanalytical standard' for illusions of possibility. Those who are committing a modal error should come to accept the explanation of why they go wrong once it is on the table. Yablo's example of an error theory that does not live up to the psychoanalytical standard is Kripke's explanation of the intuition that heat, instead of being high mean molecular motion, could be low mean molecular motion (LME, cf. Kripke 1980, pp. 150–151). According to Kripke, the intuition can be explained by confusing two possibilities. On the one hand, there is the possibility that some creature is in qualitatively the same epistemic situation with regard to LME that we were in when we first experienced heat. On the other hand, there is the possibility of heat being LME. Yablo finds it unlikely that anyone would accept this explanation of her intuition. The qualitatively identical situation would have to be subject to stronger restrictions: it does not suffice that some creature is wired such as to experience LME as heat; it would take a creature with a perceptual apparatus like ours in its normal environment.

I am not sure about Yablo's psychoanalytical standard, but we surely should avoid attributions of overly gross errors. In particular, it counts against an error theory if the erroneous intuitions persist once the purported source of error is exposed. I shall ponder the prospects of deriving an error theory meeting a standard of plausibility of the sort Yablo has in mind from Vetter's hypothesis. I shall not survey alternative error theories (e.g. Williamson 2017).

My proposed error theory can be contrasted to a strong version of error theory: Vetter says '… there is little reason to think that we … have a good intuitive grasp of the distinction between epistemic and metaphysical readings in the case of counterfactuals...' (Vetter 2016b, p. 785, slightly adapted). Vetter may be read as questioning our very ability to disambiguate which reading is the right one in a context. The consequence would be a general scepticism about our competence with counterfactuals. Such a scepticism is implausible. Indeed, the two readings are introduced into the debate by examples, presupposing our competence of distinguishing them, and the very hypothesis that counterpossibles are to be read epistemically or vacuous as endorsed by Vetter depends on our competence to discern and assess epistemic readings of counterpossibles.

I thus suggest to cash out EP by a mild error theory *MET*: although we generally get the different readings right in their context, nonvacuists fail to observe the distinction epistemic-metaphysical in the general linguistic consequences they draw from their intuitions. They draw the consequence that counterpossibles are nonvacuous full stop, but in fact they should say

that they are nonvacuous when read epistemically. MET stands a good chance of meeting the psychoanalytical standard. Nonvacuists do not get their counterfactuals wrong, they just fail to appreciate a classification that has only recently gained attention from linguists and philosophers.[4]

3. *Support for a mild error theory*

3.1. THE ARGUMENT FROM OPACITY

I shall now consider arguments for MET. The *main argument* in the literature why nonvacuous counterpossibles are epistemic is that metaphysical modals are referentially transparent, whereas epistemic modals may be opaque:

'… when faced with a choice between an epistemic and a metaphysical reading, a modal's giving rise to referential opacity is conclusive evidence for the epistemic reading. The reason is simply that metaphysical modality concerns the objects, properties, and relations that a given modal claim is about, not any representational or cognitive features of the terms we use to refer to them' (Vetter 2016a, p. 2698).

Whenever a counterpossible is treated as referentially opaque, we automatically get an epistemic reading. Antivacuity intuitions treating counterpossibles as referentially opaque ipso facto are due to an epistemic reading.

Consider

12 Hesperus might have been different from Phosphorus.

If 'might' is read metaphysically and enforces a referentially transparent context, (12) simply says that the object that is Hesperus could have been different from itself, and this seems an absurd thing to say. Contrast this with an epistemic reading: if (12) is read as saying that, from some salient epistemic vantage point it was open whether Hesperus was Phosphorus, (12) is true and can tell us something new and relevant.

There is a parallel to:

13 If Hesperus had been different from Phosphorus, Hesperus would not have been Phosphorus.
14 If Hesperus had been different from Phosphorus, Hesperus would not have been Hesperus.

In the epistemic reading, we use an epistemic vantage point from which it is yet open whether Hesperus was Phosphorus. In the resulting epistemically

possible scenarios, Hesperus was not Phosphorus, but Hesperus was still Hesperus. (13) seems true, and (14) false.

Consider Williamson's discussion of counterpossibles like (13) and (14). Their antecedent and consequent

'… must have the same truth value because their antecedents and consequents concern the same objects, properties, and relations: it matters not that different names are used, because the counterfactuals are not about such representational features (if the substitution of coreferential names in propositional attitude ascriptions does not preserve truth value, the reason is that such ascriptions are about representational features)' (Williamson 2007, p. 175).

If the metaphysical reading permits to exchange coreferential expressions salva veritate, the truth-value of (13) and (14) is the same.

Moreover, consider:

15 If Hesperus had been different from Phosphorus, Hesperus would still have been Phosphorus.

If we follow Williamson, the truth-value of (15) in the metaphysical reading is the same as the truth-value of

16 If Hesperus had been different from Hesperus, Hesperus would still have been Hesperus.

While (13) intuitively seems true and (14) and (15) seem false, (16) does not elicit the same clear intuitive reaction. If (13) is equivalent to (14) and (15) is equivalent to (16), the divergence in our intuitions indicates that they are either incoherent or are due to an epistemic reading.[5] Since the latter has to do with how we represent things (Williamson's 'representational features'), referential opacity as an epistemic phenomenon should be expected. Read epistemically, (13) seems true, (14) and (15) false, and (16) seems somewhat misplaced. It is not clear how to make room for the antecedent as an epistemic possibility. This squares with our intuitions.

The argument from opacity supports only that *some* intuitions about counterpossibles must be explained by an epistemic reading. They all somehow have to do with questions of identity. The impossibility of the antecedent results from violating constraints on referring to the very same entity. EP concerns *nonvacuity intuitions in general*. It seems doubtful that the argument from referential opacity generalizes to all of our nonvacuity intuitions, though (cf. Locke forthcoming). In particular, it seems highly important to extend MET to mathematical counterpossibles like (1)–(2). For it is argued that the vacuity of counterpossibles is crucial to their most cherished use, mathematical (and logical) *reductio ad absurdum* arguments:

'it is hard to make sense of the practice of counterfactual *reductio* except by attributing to mathematicians a tacit commitment both to the necessity of all mathematical truths and a logic according to which all counterpossibles are true' (Yli-Vakkuri and Hawthorne 2020, cf. Williamson 2017, 214).

Other things being equal, it would be good for all mathematical counterpossibles to be true.[6] Thus, we have to examine whether EP can make good on this requirement.

3.2.   A MORE GENERAL ARGUMENT

3.3.   *The Epistemic Reading as Default?*

Many antivacuity intuitions like those about (1) and (2) 'come out of the blue':

1 If Hobbes had squared the circle, then the mathematical community at the time would have been surprised.
2 If Hobbes had squared the circle, then sick children in the mountains of Afghanistan at the time would have been thrilled.

(1) seems true, and (2) seems false. How can the epistemic reading be used to explain these intuitions? I shall develop an explanation. Only since 1882, we know that the circle cannot be squared. We may apply Edgington's 'why did you investigate …?' – template to conjure up a dialogue between pre-1882 historians of mathematics:

Historian1: 'Why didn't you investigate the possibility of Hobbes having squared the circle?'

Historian2: '(1) If Hobbes had squared the circle, then the mathematical community at the time would have been surprised. There should have been traces of that surprise, but there were none. Not everyone would have been surprised, though. It is simply false that (2) if Hobbes had squared the circle, then sick children in the mountains of Afghanistan would have been thrilled.'

In this case, we presumably get an epistemic reading.

One concern is that the special epistemic vantage point in the dialogue does not explain why we endorse (1) and reject (2) *out of the blue*. There are many fairly standard situations in which (1) comes out false. Consider another pre-1882 dialogue:

Historian1: 'In pondering your hypothesis that Hobbes squared the circle, why didn't you take into account the records of the mathematical community at the time?'

Historian2: 'Well, I went through the records. There is no mentioning of Hobbes having contributed to mathematics. Hence I could rule out that,

(1) if Hobbes had squared the circle, the mathematical community would have been surprised. It would not have taken any notice.'

We can conjure up contexts in which an epistemic reading of (1) seems false as well as contexts in which it seems true. The same for (2). All these contexts can be expected to be a bit recherché, but so is the very supposition of Hobbes having squared the circle.

We need an explanation of our out-of-the-blue intuitions. I shall present an explanation in two steps. The explanation comes with certain assumptions on how the epistemic reading works. I offer them as the best version of EP I can think of. There may be alternative understandings, which I cannot cover in this article. The first step dissolves the competition between the epistemic and the metaphysical reading. When we encounter a past subjunctive out of the blue, we have to decide whether to read it epistemically or metaphysically. Khoo (2015, pp. 16–17) claims that the metaphysical reading is our default choice as we normally care about past and present metaphysical but only about current epistemic possibilities. I find Khoo's arguments unconvincing. Past epistemic possibilities can explain past actions, as in Edgington's detective case, and they may be used to guide future actions, for instance when the detective works on similar cases. Moreover, there are doubts that such pragmatic issues apply to far-fetched conditionals like counterpossibles.

One feature of counterpossibles that could be exploited to break the tie in favour of the epistemic reading is precisely the peculiarity that makes counterpossibles far-fetched: the antecedent is a metaphysical impossibility. As Williamson says, '… the objective impossibility of the antecedent might pragmatically trigger an epistemic reading …' (Williamson 2017, p. 217). When a subjunctive is read epistemically, the antecedent concerns just the epistemic possibility of some p which as a matter of fact is metaphysically impossible. We encounter such possibilities all the time, for instance when we do not know whether A is identical to B. Since past subjunctives, read epistemically, are mainly in the business of assessing epistemic vantage points differing from our current one, they can be expected to routinely deal with metaphysically impossible antecedents.

In contrast, metaphysical impossibilities are quite surprising in the case of past subjunctives, read metaphysically. The metaphysical reading is often taken to come with the presupposition that the antecedent is metaphysically possible (e.g. Khoo 2015, p. 9). Moreover, the subjunctive mood normally conveys that the antecedent is genuinely contrary-to-fact; it is actually false. When we consider the supposition that A is (not) identical to B as genuinely contrary-to-fact, we normally deal with a *manifest* metaphysical impossibility. Such considerations rarely make sense outside of a philosophical context.[7] This lends some plausibility to the hypothesis that we implicitly use the manifest impossibility of the antecedent encountered out of the blue as a cue how to dissolve the competition between an

epistemic and a metaphysical reading in favour of the former.[8] This hypothesis would explain the default status of the epistemic reading for counterpossibles. I do not incur a commitment to it, offering it only as a potential support for EP.

3.4. Stand and Fall

On behalf of MET, I have offered an explanation of why we by default tend towards the epistemic reading for past subjunctives with an impossible antecedent. This does not yet account for their intuitive truth values. My next step concerns finding a suitable vantage point to be used when reading a past subjunctive epistemically out of the blue. It has long been observed that past subjunctive conditionals, uttered at a later time t2, often 'stand and fall' with certain present-directed or future-directed indicative conditionals, uttered at an earlier time t1. For any future-or-present-directed indicative conditional that could be truly uttered from viewpoint Pi1 at t1, there is a corresponding past subjunctive that could be truly uttered at t2 to simulate Pi1 in retrospect, which stands and falls with the former, *provided the past subjunctive is read epistemically*. In turn, we can simulate any hypothetical past epistemic vantage point by a past subjunctive as uttered now.

I thus supplement my initial characterization of the epistemic reading by a systematic connection to indicative conditionals:

A past subjunctive conditional, read epistemically, is true in the current context precisely if the corresponding indicative conditional is true as assessed from any past epistemic vantage point that is contextually maximally salient.

Of course, this condition is only of limited avail as long as we do not have a general semantics of indicative conditionals.

I do not claim that the view generalizes to metaphysical readings of past subjunctives (cf. Bennett 2003, pp. 242–244). But for epistemic readings, it is highly plausible. A past subjunctive invites us to detach from our present vantage point: we vicariously take a remote past vantage point at which some epistemic possibility was open and consider whether the possibility was one in which the consequent was true. The latter question is also answered by a future-or-present-directed indicative conditional as it would have been assessed by someone occupying that vantage point.

Any past vantage point can be raised to salience by contextual cues, whether it actually has been occupied or not. Consider Edgington's detective who has reason to reject:

4 If Oswald had not killed Kennedy, no one else would have.

But from a viewpoint before Kennedy's death, the corresponding indicative does not seem false:

17 If Oswald will not kill Kennedy, no one else will.

In order to get the right candidate to stand and fall with the detective's denial of (4), the contextually salient vantage point must be one at which she already knew that Kennedy was shot and had to decide between Oswald and other candidates for the killer. Edgington conjures up a *special* context.

To explain our *out-of-the-blue* intuitions about (1) and (2), we have to posit a *default* vantage point. I have no general recipe to offer, but the following assumptions are tempting: firstly, the presuppositions of the antecedent and the consequent have to be satisfied: Hobbes should be alive. Secondly, it should be an open question whether the antecedent is true. It should be open whether Hobbes succeeds in squaring the circle. Thirdly, we should not go much farther into the past than necessary to accommodate the first two conditions. The Hobbes counterpossibles can be related to future-directed or present-directed indicative conditionals as uttered from a vantage point around 1600:

18 If Hobbes squares the circle, then the mathematical community will be surprised.
19 If Hobbes squares the circle, then sick children in the mountains of Afghanistan will be thrilled.

(18) seems true, and (19) seems false, as desired. This explains our intuitions about (1) and (2).

The general stand-and-fall template can be exploited to get epistemic readings for many subjunctive conditionals in a suitable context. Consider again

13 If Hesperus had been different from Phosphorus, Hesperus would not have been Phosphorus.
15 If Hesperus had been different from Phosphorus, Hesperus would still have been Phosphorus.

(13) seems true, and (15) seems false. The same goes for the following as uttered at any time at which it was still an open question whether Hesperus was Phosphorus:

20 If Hesperus is different from Phosphorus, Hesperus is not Phosphorus.
21 If Hesperus is different from Phosphorus, Hesperus is still Phosphorus.

(20) seems true, and (21) seems false. We get the same effect for analogues of Yablo's 'turn out':

22 If Hesperus turns out to be different from Phosphorus, Hesperus is not Phosphorus.
23 If Hesperus turns out to be different from Phosphorus, Hesperus is still Phosphorus.

The strategy works most straightforwardly in cases in which there was a time when we did not yet know the antecedent proposition. One may suspect that it won't work as well in cases of very basic logical or mathematical truths like $5 + 7 = 13$, which might be knowable *a priori*. Consider (Williamson 2007, p. 172)

24 If it had been the case that $5 + 7 = 13$, $5 + 7$ would be 13.
25 If it had been the case that $5 + 7 = 13$, it would still be the case that $5 + 7 = 12$.

Whatever our intuitions about (24) and (25) are, I surmise that we have parallel intuitions about:

26 If $5 + 7 = 13$, then $5 + 7 = 13$.
27 If $5 + 7 = 13$, then $5 + 7 = 12$.

Applying the stand-and-fall template, we would have to imagine a vantage point at which we did not yet know that $5 + 7 = 12$. This raises important questions about aprioricity, which I cannot address here.[9] Yet I observe that many of our intuitions about *indicatives* with metaphysically impossible antecedents as considered above come out of the blue. We do not need to imagine them uttered from a temporally remote epistemic vantage point. Here is an explanation: an utterance tends to change the context set of epistemic possibilities that are common ground in *current* conversation. Indicative conditionals come with the presupposition that the antecedent is a live possibility. For instance, we normally rule out Oswald not having killed Kennedy from the live epistemic possibilities. When (3) is uttered, we are willing to make room for the supposition that Oswald did not kill Kennedy by including corresponding epistemic possibilities among the live possibilities (Stalnaker 2014, p. 177). In a similar vein, Putnam (1975, p. 61) grants that pencils may eventually turn out to be animals, although we surely know them to be artefacts. The same goes for the antecedent of (26)–(27). When (26) and (27) are uttered, we accommodate epistemic possibilities in which $5 + 7 = 13$. Throughout these possibilities, it seems that $5 + 7 = 13$ and $5 + 7 \neq 12$.

I venture some tentative remarks on why we may be inclined to presently accept (26) and reject (27): the epistemic possibilities in which $5 + 7 = 13$ trivially seem possibilities in which $5 + 7 = 13$ by virtue of the way the

antecedent possibility has been introduced. The question whether the relevant possibilities are possibilities in which $5 + 7 \neq 12$ is more intricate. I surmise that certain mathematical heuristics are in play: Names like '12' and '13' are known to uniquely identify numbers, which therefore are different. This knowledge trumps calculating the sum and makes (27) seem false. Since I offer this explanation only to support the error theory MET, I incur no commitment as to whether our intuitions about (26)–(27) are sound.

Coming to *subjunctives* like (24)–(25), unlike indicatives they do not come with the presupposition that the antecedent is a live possibility. Read epistemically, they refer us to a remote epistemic vantage point. To account for them by an epistemic reading, we in principle need the stand-and-fall template. Still for our assessment of (24)–(25), the observations about presuppositions I used in the Hobbes examples do not seem to apply. Instead, we may simply simulate a very recent utterance of (26)–(27). I am not fully satisfied with this solution, but we should keep in mind that I only try to cash out MET. If my attempts fail, the worse for MET.

The patterns outlined can be applied to further examples. The following subjunctive conditionals are intuitively false:

28 If intuitionistic logic had been the correct logic, then the law of excluded middle would still have been unrestrictedly valid.

It is not raining:

29 If it had actually been raining (here and now), then it would not have been raining (here and now).

These counterpossibles arguably stand and fall with the following innocuous indicative conditionals, which deal with epistemic possibilities:

30 If intuitionistic logic is the correct logic, then the law of excluded middle is still valid.
31 If it is actually raining (there and then), then it will not be raining (there and then).[10]

Taking stock, we apparently can get an epistemic reading for any past subjunctive in a suitable context. We can even explain out-of-the-blue intuitions regarding counterpossibles like (1)–(2) by an epistemic reading. The argument is still limited, though. It does not show that *any* antivacuity intuition can be explained in this way. In Section 4.2, I shall present cases in which such an explanation is rather unlikely.

4. Counterarguments

4.1. DIVIDE-AND-CONQUER FOR NONVACUISTS

Having tried my best to support MET, I shall now present a reply to EP on behalf of nonvacuists. They may hijack the divide-and-conquer strategy of EP. Take the argument from referential opacity. The argument is only of limited avail in the dialectics against nonvacuists. They are often ready to bite bullets like referential opacity and lack of deductive closure for the metaphysical reading (Brogaard and Salerno 2013, p. 650). They tend to reject the claim that the hyperintensionality of counterpossibles is only due to 'representational features': 'There are hyperintensional contexts that are not in any way "about representational features"' (Berto *et al.* 2018, p. 701). An example may be counterpossibles tracking hyperintensional metaphysical relationships (cf. Nolan 1997; Dorr 2008; Brogaard and Salerno 2013; Kment 2014; Bernstein 2016; Wilson 2018). Metaphysical explanations are among the main contexts in which counterpossibles seem useful to nonvacuists. And the use they have in mind requires a metaphysical reading. In many cases, metaphysical features of the world seem to account for the hyperintensionality of counterpossibles.

However, this reply harbours a substantial challenge for nonvacuists: what features account for our intuitions about the Hesperus-Phosphorus cases (13)–(16)? Assume there are no metaphysical features of identity which support the differences in our intuitions. If nonvacuists bring in nonmetaphysical features to account for them, for example 'representational features', they may jeopardize their metaphysical ambitions.[11] Surprisingly, at this point, epistemic readings of counterpossibles may come in handy to nonvacuists.

One radical thesis suggested to nonvacuists by Kocurek (2020) is that all counterfactuals, whether counterpossibles or not, should be read epistemically as even normal counterfactuals display referential opacity.[12] The hypothesis that all counterfactuals are epistemic has to be researched more carefully, but the evidence so far is not compelling.[13]

A less radical move for nonvacuists is the following: they may agree that in cases like (13)–(16), a modal's giving rise to referential opacity triggers an epistemic reading. The status of our intuitions has to be clarified accordingly. But nonvacuists may insist that, in other cases where there is no referential opacity, our antivacuity intuitions are due to metaphysical features and thus require a metaphysical reading.[14] Counterlogicals and countermathematicals may be treated as vacuous or as nonvacuous.[15] The consequence is that we have to carefully distinguish the different readings at stake, but there is no reason to be sceptical about our competence of doing

so. Hence, the divide-and-conquer-strategy of EP may be hijacked by nonvacuists.

MET and the counterstrategy I offer to nonvacuists can be tested by their predictions. As a preliminary test, take a standard example of a metaphysical relationship. It is widely assumed that Socrates is essentially human, but it is not essential to Socrates that he is an element of singleton Socrates (Fine 1994). In turn, it is essential to singleton Socrates that Socrates is an element of it. Consider

32 If Socrates had existed but not been human, he would still have been the same entity he actually was.

33 If Socrates had existed but not been human, he would have been a substantially different entity than he actually was.

34 If Socrates had existed but Socrates had not been an element of singleton Socrates, he would still have been the same entity he actually was.

35 If Socrates had existed but Socrates had not been an element of singleton Socrates, he would have been a substantially different entity than the one he actually was.

I assume that, when we weigh (32) against (33), keeping in mind that Socrates was human, there is a certain pull to reject (32) and accept (33); in contrast, (34) sounds true, and (35) false. A nonvacuist explanation for these intuitions is that essence grounds identity. Socrates was essentially human but not essentially a member of singleton Socrates. He was the entity he was in virtue of being human but not in virtue of being an element of singleton Socrates. Counterpossibles are sensitive to the pertinent hyperintensional relationships.

I admit that (32)–(35) may be problematic as 'same entity' and 'substantially different entity' are used deliberately vague. Our intuitions may be dismissed as impaired by this vagueness. But assume our intuitions are sufficiently clear and strong. How does MET fare as an explanation? Applying stand-and-fall, a confusion with the epistemic reading seems rather unlikely. Presumably we can be a priori certain that, if Socrates existed, he was the same entity he actually was, but not that he was human. Making room for the epistemic possibility that Socrates was not human, we still accept that Socrates was the very same entity he actually was. To be sure, our intuitions about (34) and (35) can be explained by the epistemic reading, but if we consider (32)–(35) in one breath, the contrast in our intuitions may be better explained by the metaphysical bond between Socrates and his being human as contrasted to his being an element of singleton Socrates than by MET. This would give the nonvacuist counterstrategy a slight advantage compared with MET.

4.2. NOT ALL ERRORS ARE MILD

According to MET, nonvacuists neglect the fact that their intuitions are due to reading counterfactuals epistemically. I shall now argue that MET does not work for EP; this reduces the latter's plausibility. The idea of MET was that we normally get the different readings right when assessing utterance tokens, but that we then fail to correctly classify the epistemic readings giving rise to nonvacuist intuitions as such. Yet there are cases in which a classification error is unlikely; our *enlighted* nonvacuist intuitions target a metaphysical reading even if we attend to the distinction between the two readings. I shall present two kinds of cases, which may be called *coupling* and *error about the modal status of the antecedent*. I note that they pose a problem for EP and not for the nonvacuist counterstrategy outlined in the last section.

Beginning with the argument from coupling, I consider an uncontentious metaphysical reading of a normal counterfactual. The counterfactual has a *possible* antecedent, but it is uttered 'in one breath' (cf. Wright 1983, p. 138) with another past subjunctive that is closely related, the difference being that the antecedent of the latter is *impossible*. I shall also avail myself of the commonplace that metaphysical readings often deal with *causal* relationships. When no specific epistemic vantage point is salient and the relationship between antecedent and consequent corresponds to one of cause and effect, we would normally expect a metaphysical reading.

I shall use the Hobbes counterpossibles to conjure up such a situation.[16] Consider two contemporary historians of mathematics discussing the causal-social impact of mathematical innovations among learned communities in early modern Europe.

Historian3: 'Mathematical innovations like Fermat's Last Theorem had a strong impact on society in general and mathematical communities in particular. But Hobbes did not substantially contribute to mathematics. In *Live of Hobbes*, Aubrey says of his attempts at squaring the circle (cf. Jesseph 1999): "I have heard Sir Jonas Moore (and others) say that'twas a great pity he had not began the study of the mathematics sooner, for such a working head would have made great advancement in it. So had he donne, he would not have layn so open to his learned mathematicall antagonists …."(Aubrey 1898, pp. 332–333) To be sure:

36 If Hobbes had been the first to write down Fermat's Last Theorem, he would have made a lasting impact on the history of mathematics.

In the same vein,

1 if Hobbes had squared the circle, then the mathematical community at the time would have been surprised.

But of course, he failed, as it is impossible to square the circle. In sum, Hobbes did not make a lasting impact on mathematics.'

Historian4: 'The impact of mathematics on humanity should not be exaggerated, though. Take your examples: it is clearly both false that,

37 If Hobbes had been the first to write down Fermat's Last Theorem, then sick children in the mountains of Afghanistan at the time would have been thrilled; and

  2 If Hobbes had squared the circle, then sick children in the mountains of Afghanistan at the time would have been thrilled.'

There is little reason to read the past subjunctives in this dialogue epistemically. The historians are not interested in assessing a remote epistemic vantage point. They are musing about the causal impact of a certain scientific discovery, measured by the counterfactual difference it would have made.

There is no reason whatsoever to read the Fermat counterfactual (36) epistemically. An epistemic reading even faces substantial difficulties. (36) seems somewhat plausible, even considered out of the blue, but the epistemic reading does not support (36). Applying my general stand-and-fall-template, try to imagine a *future-directed* viewpoint in which the epistemic possibility of Hobbes writing down what became later known as Fermat's Theorem is assessed avant la lèttre. This possibility is not obviously one in which the mathematical community reacts. After all, the impact of Fermat's Last Theorem became only visible later. We can get (36) for example by imagining a later epistemic vantage point in which someone discusses the epistemic alternative that Hobbes and not Fermat first wrote down Fermat's Theorem ('Why didn't you investigate …?'). But this would take a different context, which is not supported by my scenario or my stand-and-fall template for out-of-the-blue epistemic readings.

The two historians harbour no doubts about Fermat's authorship. They consider a genuine counterfactual alternative from their own present-day viewpoint. The Fermat counterfactuals should be read metaphysically. The same goes for the circle counterfactuals as there is no change in context between the Fermat counterfactuals (36) and (37) and the circle counterfactuals (1) and (2) as uttered in the dialogue. The circle counterfactuals are salient as they deal with Hobbes's documented attempts at squaring the circle. Again, the two historians are not interested in any epistemic vantage point at which Hobbes squaring the circle was a salient epistemic possibility. They intend the circle counterfactuals (1), (2) to be read in one breath with the Fermat counterfactuals (36), (37), although they are well-aware that the circle cannot be squared. In sum, there are strong reasons to read (1)–(2) metaphysically in my example, and still there is a strong inclination to deem them nonvacuous (one true and one false).

Now one may argue that my historians in spite of all the contextual cues for a metaphysical reading are misled by surreptitiously replacing it with an epistemic reading.[17] However, things are different if *we*, heeding the contextual cues for a metaphysical reading, share the mathematicians' intuitions. I confess that I share these intuitions to the same extent to which I share other nonvacuity intuitions. To me, this is evidence that the error theory cannot be mild.

My cases can be seen as test cases for MET. If the test turns out against MET, there is still the option of a strong error theory SET: according to that theory, we may be lured into nonvacuity intuitions by the epistemic reading even where we attentively target the metaphysical reading. I think that SET considerably reduces the original plausibility of EP. In particular, there are doubts that it satisfies Yablo's psychoanalytical standard: why should we, attentive as we are to the metaphysical reading, be deluded by confusing it with the epistemic one?

My second case, *error about the modal status of the antecedent*, leads to the same result. Take someone who is convinced of the metaphysical possibility that the circle can be squared, say Aubrey musing about his late contemporary Hobbes. Knowing that Hobbes's efforts in mathematics were futile, Aubrey might have said:

'Mathematical discoveries cause great delights to the life of the learned. It is surely false that,

38 if Hobbes had made a great mathematical discovery, the mathematical community would not have cared.

In this vein, it is surely false that,

39 if Hobbes had squared the circle, the mathematical community would not have cared.'

Given his belief that the circle can be squared, Aubrey has no reason to read (39) epistemically. It would also be inappropriate for his audience to read his utterance of (39) in a way that differs from Aubrey's intentions. Even if the audience knows that the circle cannot be squared, the utterance should be interpreted metaphysically, taking into account Aubrey's false belief that the circle can be squared. My point is not that *Aubrey's* intuition about (39) was right. Since Aubrey's modal beliefs were formed under conditions of modal error, his intuition may be mistaken. But things are different if *we*, forced to consider a metaphysical reading by our intention to correctly interpret Aubrey, are inclined to agree. Although we find fault with his modal belief that the circle could be squared, we tend to share his intuitions about (39) as uttered by him.

5. *Conclusion: The epistemic reading as a relevant alternative*

While EP is appealing at first glance, upon closer inspection, it comes with high burdens. My results are limited. I do not take any move I have discussed to decide the dispute between vacuists and nonvacuists. A message to take home for both vacuists and nonvacuists is that, in judging antivacuity intuitions, we should weigh in how far they can be explained by an epistemic reading.

6. *Acknowledgments*

Dipartimento di Filosofia
Università degli Studi di Milano

NOTES

[1]   Vetter fully endorses the hypothesis; Williamson only accepts that the impossibility of the antecedent might trigger an epistemic reading.

[2]   Veltman (2005, p. 174) disagrees, arguing: 'Notice that only people who have gone through the same epistemic process as you did in your role of detective, will be able to appreciate this epistemic reading.' Yet it may be useful to use a subjunctive to simulate epistemic processes undergone by other people and thus to 'appreciate the epistemic reading'.

[3]   We may add an ordering source selecting the 'best' worlds. Then, only the best antecedent worlds have to be consequent worlds.

[4]   An anonymous reviewer has doubted that my MET cuts much ice as long as it does not explain why nonvacuists take their intuitions to support their position even if they are explicitly aware of the semantic distinction of epistemic and metaphysical readings. To a certain extent I agree. Indeed, it is one of my main criticisms of EP that *enlighted* nonvacuist intuitions cannot be completely explained away by MET. Yet the debate on an epistemic reading of subjunctive conditionals is not widely known. I do not find it so implausible that many authors fail to appreciate the epistemic reading as such. Thus, a good part of the antivacuist intuitions voiced in the literature may indeed be explained away by such a failure.

[5]   An anonymous reviewer has voiced doubts that these intuitions are shared by all nonvacuists. I for one share them (being undecided on the vacuism–nonvacuism debate), and

I think that they feature in Williamson's reconstruction of nonvacuism. Still, I acknowledge that there is room for disagreement here.

[6]   Though see Jenny 2018.

[7]   Nonvacuists like Tan (2019) might demur. Yet I think it would be overdemanding to insist that EP has to rely exclusively on assumptions shared by nonvacuists.

[8]   An anonymous reviewer has suggested that our tendency towards an epistemic reading may be explained by the Gricean maxims of quantity and quality. I am open to an interpretation in terms of pragmatic maxims of conversation, perhaps considerations of conversational relevance.

[9]   One may already feel concerned about aprioricity in the case of the Hobbes counterpossibles, but here we have a clear date at which people arguably came to know that the circle cannot be squared.

[10]   'Here and now' in (29) refers to the present point in spacetime t2, 'there and then' to this point as seen from an earlier point t1.

[11]   Cf. Williamson's criticism of Brogaard and Salerno's 2013 semantics as '…patched together from epistemic and non-epistemic pieces.'(2017, 211).

[12]   Kocurek's main argument is that we only sometimes find arguments based on substituting coreferential terms in counterfactuals valid. He suggests that, whenever we find such an argument valid, we add a hidden premise: an identity statement featuring the coferential terms remains true under the pertinent counterfactual suppositions. For instance, Hesperus = Phosphorus would have to remain true under the counterfactual supposition that Hesperus is not Phosphorus.

[13]   The only example of a referentially opaque counterfactual with a possible antecedent provided by Kocurek goes as follows: a rocket is programmed to continue its course to Hesperus only if it finds it different from Phosphorus upon checking its database. The rocket correctly aborts it mission. The following seems true: 'if the rocket had continued on its course, Hesperus would not be Phosphorus.' I see that this can be true in an epistemic reading, but I doubt that we would prefer it in the metaphysical reading to 'if the rocket had continued on its course, Hesperus would still be Phosphorus, but the rocket would not have followed its programming'.

[14]   Williamson criticizes a parallel divide-and-conquer move for counterlogicals and countermathematicals, arguing that the 'motivating features' are the same in the purportedly different cases (Williamson 2017, p. 214). But nonvacuists may insist that the motivating features are epistemic in one case and metaphysical in the other.

[15]   Kment (2014, p. 73), who champions the metaphysical use of nonvacuous counterpossibles, reports intuitions in favour of the vacuity of counterlogicals.

[16]   Bernstein (2016) argues that counterpossibles like (1) have a function in causal explanation. The pertinent causal relationships have to be construed as causation by omission ('Had Hobbes not failed to…', cf. Bernstein 2016, p. 2576).

[17]   An anonymous referee raised the objection that 'the inclination might come from a mixture of vestiges of the understanding before we realized the impossibility, together with the unconsciously substituted epistemic reading.' I think the diagnosis of a reading that is substituted unconsciously and contrary to one's explicit interpretation would be quite uncharitable here. Again, the error theory would not be mild.

REFERENCES

Adams, E. (1970). 'Subjunctive and Indicative Conditionals,' *Foundations of Language* 6, pp. 89–94.

Aubrey, J. (1898). *Brief Lives: Chiefly of Contemporaries, Set Down by John Aubrey, Between the Years 1669 & 1696*, Volume 1. Oxford: Clarendon Press, pp. 332–333.

Bennett, J. (2003). *A Philosophical Guide to Conditionals.* Oxford: Oxford University Press, pp. 242–244.

Bernstein, S. (2016). 'Omission Impossible,' *Philosophical Studies* 173, pp. 2575–2589.

Berto, F. et al. (2018). 'Williamson on Counterpossibles,' *Journal of Philosophical Logic* 47, pp. 693–713.

Bjerring, J.-C. (2014). 'On Counterpossibles,' *Philosophical Studies* 168, pp. 1–27.

Brogaard, B. and Salerno, J. (2013). 'Remarks on Counterpossibles,' *Synthese* 190, pp. 639–660.

Dorr, C. (2008). 'There Are No Abstract Objects,' in T. Sider et al. (eds) *Contemporary Debates in Metaphysics*. Oxford: Blackwell, pp. 32–63.

Edgington, D. (2011). 'Causation First: Why Causation Is Prior to Counterfactuals,' in C. Hoerl et al. (eds) *Understanding Counterfactuals, Understanding Causation: Issues in Philosophy and Psychology*. Oxford: Oxford University Press, pp. 230–241.

Fine, K. (1994). 'Essence and Modality,' *Philosophical Perspectives* 8, pp. 1–16.

Jenny, M. (2018). 'Counterpossibles in Science. The Case of Relative Computability,' *Noûs* 52, pp. 530–560.

Jesseph, D. M. (1999). *Squaring the Circle: The War Between Hobbes and Wallis*. Chicago: University of Chicago Press, pp. 1–433.

Khoo, J. (2015). 'On Indicative and Subjunctive Conditionals,' *Philosopher's Imprint* 15, pp. 1–40.

Kment, B. (2014). *Modality and Explanatory Reasoning*. Oxford: Oxford University Press, pp. 3–73.

Kocurek, A. (2020). 'On the Substitution of Identicals in Counterfactual Reasoning,' *Noûs* 54, 600–631.

Kripke, S. (1980). *Naming and Necessity*. Cambridge, Mass: Harvard University Press, pp. 150–151.

Lewis, D. (1973). *Counterfactuals*. Cambridge, Mass: Harvard University Press, pp. 24–25.

Locke, T. (forthcoming). 'Counterpossibles for Modal Normativists,' *Synthese*.

Nolan, D. (1997). 'Impossible Worlds: A Modest Approach,' *Notre Dame Journal of Formal Logic* 38, pp. 535–572.

Putnam, H. (1975). 'The Meaning of "Meaning",' *Minnesota Studies in the Philosophy of Science* 7, pp. 131–193.

Stalnaker, R. (2014). *Context*. Oxford University Press: Oxford, p. 177

Tan, P. (2019). 'Counterpossible Non-Vacuity in Scientific Practice,' *The Journal of Philosophy* 116, pp. 32–60.

Veltman, F. (2005). 'Making Counterfactual Assumptions,' *Journal of Semantics* 22, pp. 159–180.

Vetter, B. (2016a). 'Counterpossibles (Not Only) for Dispositionalists,' *Philosophical Studies* 173, pp. 2681–2700.

Vetter, B. (2016b). 'Williamsonian Modal Epistemology, Possibility-Based,' *Canadian Journal of Philosophy* 46, pp. 766–795.

Williamson, T. (2007). *The Philosophy of Philosophy*. Blackwell: Oxford, pp. 171–175.

Williamson, T. (2017). 'Counterpossibles in Semantics and Metaphysics,' *Argumenta* 2, pp. 195–226.

Wilson, A. (2018). 'Grounding Entails Counterpossible Non-Triviality,' *Philosophy and Phenomenological Research* 96, pp. 716–728.

Wright, C. (1983). 'Keeping Track of Nozick,' *Analysis* 43, pp. 134–140.

Yablo, S. (2002). 'Coulda, Woulda, Shoulda,' in T. S. Gendler and J. Hawthorne (eds) *Conceivability and Possibility*. Oxford: Clarendon Press, pp. 441–492.

Yablo, S. (2006). 'No Fool's Cold,' in M. Garcia-Carpintero and J. Macia (eds) *Two-Dimensional Semantics*. Oxford: Oxford University Press, pp. 327–346.

Yli-Vakkuri, J. and Hawthorne, J. (2020). 'The Necessity of Mathematics,' *Noûs* 54, pp. 549–577.