What Might Be and What Might Have Been

Benjamin Schnieder & Moritz Schulz & Alexander Steinberg

Abstract:

The article is an extended comment on Strawson's neglected paper 'Maybes and Might Have Beens', in which he suggests that both statements about what may be the case and statements about what might have been the case can be understood epistemically. We argue that Strawson is right about the first sort of statements but wrong about the second. Finally, we discuss some of Strawson's claims which are related to positions of Origin Essentialism.

Introduction

In describing and classifying things we often rely on their *modal* characteristics. We will in general not have a satisfactory account of the nature and character of an object, unless we specify at least partly how the thing might or cannot be, and also how it might have been or could not have been. In his contribution to the *Second Jerusalem Philosophical Encounter*,[1] Strawson addressed the issue of how to understand such ascriptions of modal characteristics. Although his paper is terse and provocative, and develops an intriguing account of modal predications, it has never received much attention in the philosophical literature.

Recently, the issues dealt with in Strawson's paper have become the subject of a widespread debate. Hence, we think it worthwhile to put Strawson's account under closer scrutiny. In what follows, we first discuss his account of present tense 'might'-statements, then his account of past perfect 'might'-statements, and finally some essentialist remarks that he makes in his paper. That the discussion will be rather critical for the most part (though not exclusively so) should not belie the originality and inherent value of his pioneering approach.

[1] The papers were published in A. Margalit (ed.) (1979).

1. What Might Be

1.1 Epistemic Uses of 'Might Be'

Strawson presupposes that the locutions 'may be' and 'might be' can be used to express a certain kind of epistemic possibility. It is unclear whether he thinks that they have a metaphysical use as well. For this reason, we shall focus on epistemic uses only, leaving it open whether they can also be used in a metaphysical sense.

By a present tense use of 'may' or 'might' we shall mean any sentence in which the operator 'it may be that' or the operator 'it might be that' is attached to another sentence. For instance, the sentence

(1) It might be that they will win the match.

exemplifies a present tense use of 'might'. It is *prima facie* plausible to assume that present tense uses of 'may' and 'might' together with expressions such as 'perhaps' and 'maybe' form a family of equivalent modals. For expository purposes, we shall restrict ourselves to 'might' only.

Notice that in concentrating on the described locution, we somewhat shifted the focus of Strawson's paper, for he is primarily concerned with uses of 'might' as a predicate modifier. Such uses are exemplified in sentences like

(2) They might win the match.

Since we think that a stronger case can be made for Strawson's account as it applies to the sentential operator 'it might be that', we set uses of 'might' as a predicate modifier aside.

In his paper, Strawson does not present a direct argument to the effect that present tense uses of 'might' express an epistemic possibility rather than an objective one. Fortunately, evidence for Strawson's claim is not hard to find. If a present tense use of 'might' always were to express a certain kind of non-epistemic, objective possibility, there would be two plausible candidate modalities for it to express: present objective chances or absolute metaphysical possibility. On the first suggestion, we would say with 'it might be that *p*' something equivalent to 'there is a present objective chance that *p*'. On the second proposal, 'it might be that *p*' would be equivalent to the absolute metaphysical possibility of the proposition that *p*. We shall argue that neither suggestion can account for certain present tense uses of 'might'.

Consider the following two sentences:

(3) It might be that they won last night.

(4) It might be that Goldbach's conjecture is true.

Now, present objective chances of propositions about the past are always either 1 or 0, depending on whether they are true or false. So, if (3) were to express that there is a present objective chance that they won last night, it would only be acceptable for someone who already believes that they won last night. However, that gets the acceptability conditions of (3) wrong: sentence (3) is acceptable for someone if she is uncertain about the score of last night's game. Similarly, sentence (4) can be used to show that present tense uses of 'might' cannot always be taken to express absolute metaphysical possibility either. Since Goldbach's conjecture is either necessarily true or necessarily false, it can only be metaphysically possible if it is true. Thus, sentence (4) would only be acceptable for someone who already believes Goldbach's conjecture to be true. But again, this gets the acceptability conditions of (4) wrong: it is acceptable for someone who is uncertain about whether Goldbach's conjecture is true or false.

One final remark. In our argument to the effect that some present tense uses of 'might' cannot be accounted for in terms of any kind of objective possibility, we assumed that the only two plausible candidates are present objective chances and absolute metaphysical possibility. To strengthen our argument a little further, let us point to a datum concerning the rejectability conditions of statements like (3). Someone who knows that they lost last night's game is in a position to reject that it might be that they won last night. Thus, a statement like (3) is rejectable simply on the grounds that the embedded sentence is false. However, this is compatible with the belief that the sentence is possibly true in any (non-epistemic) sense of 'possible' which does not imply actual truth. Hence, metaphysical modalities which are intermediate between present objective chances and absolute metaphysical possibility will not be able to explain certain facts about the rejectability conditions of present tense uses of 'might'. We conclude that some uses of 'might' cannot be explained in metaphysical terms.

1.2 The Details

Let us now turn to the details of Strawson's account. There is one important respect in which his account is unclear from the outset: Strawson does not decide whether it is to be seen as an account of the *acceptability* conditions of 'might'-statements or as an account of the *truth* conditions of 'might'-statements. We shall only be concerned with the acceptability conditions of present tense 'might'-statements, since we think that his approach leads to an adequate description of those conditions. On the other hand, we would like to leave it open whether it can be extended to a satisfactory account of the truth conditions of epistemic uses of 'might'.[2]

[2] We think it not unlikely that expressions of epistemic modality are not properly evaluated in the dimension of truth and falsity at all, but only in the dimension of acceptability,

Strawson expresses his view on present tense uses of 'might' in statements like the following:

> [W]e could simply conclude that to say that something may happen is merely to say that it is not certain that it will not; […] (179)
>
> When we use 'may' or 'might' to express present uncertainty about what is now future, the uncertainty is clearly relativized to a time and, more or less clearly, to persons. The time is *now*; the persons *ourselves*, the speaker and his circle and others he regards as authoritative, perhaps. […] (180)

In a first approximation, we can say that Strawson ties present tense uses of 'might' to present uncertainty with respect to the sentences it attaches to. This seems to be a good starting point. But the details of his account are less clear. In particular, two questions arise:

(i) What does Strawson's conception of *certainty* amount to?

(ii) *Whose* uncertainties are relevant for the evaluation of present tense 'might'-statements?

Re (i): Let us first introduce some common and useful epistemological background: beliefs come in degrees; degrees of belief can be understood as subjective probabilities. Certainty amounts to subjective probability 1, so an epistemic subject is certain that *p* if it assigns probability 0 to the proposition that not *p*.

This way of defining certainty does not always coincide with our pretheoretical conception of it, for we sometimes say that some things are *more (or less)* certain than others. Within the probabilistic framework, this manner of speaking is reconstructed in terms of degrees of belief, reserving the title 'certainty' for a single kind of epistemic state which is not further determinable. This conception of certainty may at first seem to serve Strawson's purposes well: if there is some subjective chance that *p,* one should accept that it might be that *p*.

But actually, things are more complicated. Strawson wants to assimilate cases in which there is a slight but (in the relevant context) negligible chance to those cases in which there is *no* chance at all:

> [S]omeone says: 'The tree may (might) fall on the house.' […] The […] remark says: 'There is a non-negligible chance that the tree will fall on the house.' (180f.)

---

sincerity etc. Although we lack a conclusive argument for this view, it is at least supported by the observation that in the current discussion about the truth conditions of epistemic modals, highly complex accounts have become popular, such as contextualist and relativist ones; see for instance DeRose (1991) and Egan & Hawthorne & Weatherson (2005).

From the quote it can be seen that Strawson would take the sentence 'it might be that *p*' to be acceptable for a speaker only if she assigns a non-negligible chance to '*p*'. Accordingly, Strawson would equate the certainty required for an assertion of 'it cannot be the case that *p*' only with a *very high* degree of belief which need not, however, equal 1.[3]

Before we discuss whether Strawson is right on these points, some terminological remarks are in order. We contrast acceptability (and rejectability) conditions with assertability conditions: by the *acceptability conditions* of a sentence *S* we mean the epistemic conditions under which a subject is justified to assign a high subjective probability to the belief expressed by *S*, and we take the *rejectability conditions* of a sentence to be the acceptability conditions of its negations. By the *assertability conditions* of *S*, we mean the conditions under which an utterance of *S* is appropriate in a certain context. Thus, acceptability and rejectability are epistemic properties whereas assertability is a pragmatic one.

Now let us see in which ways the negligibility of subjective chances is relevant for the evaluation of a 'might'-statement. On the one hand, it seems clear that the assigned chance will often have to be non-negligible for the utterance of a 'might'-statement to be *conversationally appropriate*. Suppose that Mary's husband Paul is ten minutes late and John says to her

(5) It might be that Paul is dead.

However, John asserts this simply on the basis that sometimes people die unexpectedly. Of course, John should not make this assertion without qualification, since it is strongly misleading for Mary. Hence, it seems clear that in certain cases there need to be a non-negligible chance to make a present tense 'might'-statement assertable. But it seems open to us to say that John should still *accept* that her husband might be dead. After all, he cannot exclude that Paul suddenly died of a heart attack. So, it might be that Paul is dead.

On the other hand, we often *tolerate* an utterances of 'it cannot be that *p*' even though the speaker assigns *some* subjective chance to '*p*', if only it is negligibly low. It is not absolutely clear what to make of this observation within a proper theory: *either* it reflects a part of the acceptability conditions of 'might'-statements, *or* it shows that we sometimes treat such conditions rather loosely and let pass some utterances unsanctioned that are strictly not acceptable on the basis of their acceptability conditions.

---

[3] We use 'cannot' and 'could not have been' as the contradictory negations of 'might' and 'might have been', respectively. Those formulations avoid the syntactic ambiguities of the direct negations 'might not' and 'might not have been'.

There are two points that speak in favour of the latter option: firstly, when someone says 'it cannot be that *p*' while *not* being absolutely certain that ¬*p*, she would usually retreat from the statement on being criticised: 'Okay, it is not *really* impossible, but it is *so* unlikely.' Secondly, it is generally not inconsistent to say:

(6) It might be that *p*, but the chances are negligible.

But then, the acceptability conditions of a 'might'-statement should not involve the (non)negligibility of the relevant chances. Since we take these points seriously, we settle for the second option here. Although it might not be forced on us, our choice has the pragmatic advantage of producing a less complicated theory.[4]

In summary, we think that non-negligibility only plays a pragmatic role for 'might'-statements and should not be mentioned in their acceptability conditions.

Re (ii): Let us finally focus on the question of *whose* certainties are relevant for the evaluation of 'might'. Concerning present tense statements, Strawson thinks that the relevant subjects are we '*ourselves*, the speaker and his circle and others he regards as authoritative, perhaps'. We find the vagueness of this statement hardly tolerable. What is worse, in whatever way we remove the vagueness, the epistemic state relevant for the evaluation of a 'might'-statement should, according to Strawson, not be that of a single epistemic subject. But this idea is dangerous for the whole account.

What is indeed plausible is that *my* uncertainty as to whether *p* legitimates my utterance of

(7) It might be that *p* (and it also might be that ¬ *p*).

Now assume the epistemic states that are responsible for the uncertainty are not only *mine* but include that of another person. In general, I am quite uncertain about what other people believe. So, in general, I should be quite uncertain about whether the beliefs of another person make it certain that *p*, or whether it makes it certain that ¬ *p*. Thus, *if* the beliefs of another person were relevant, I should in general refrain from asserting 'it might be that *p*', and rather say 'I do not know whether it might be that *p*; to know this, I would have to consult *NN* first'. (I should not even say 'it might be that it might be that *p*', because again, I lack access to a significant proportion of the beliefs that are relevant

[4] A contextualist could argue that the acceptability conditions of 'might'-statements indeed require non-negligible subjective chances, but that the relevant threshold varies with the context of utterance. That alone does not suffice to accommodate the acceptability of sentences such as (6), however. The contextualist would additionally have to appeal to intra-sentential shifts of context, or allow that one and the same context assigns different thresholds to explicit and implicit talk about negligibility. Although we are sceptical about such an approach, we cannot discuss it here in more detail.

for the outer 'it might be that'-operator.) But as we usually feel free to make 'might'-statements without consulting either "our circle" or whatever authorities there may be, Strawson's proposal is apparently on the wrong track. The relevant epistemic subject should be the speaker alone.

1.3 A Modified Proposal

If we modify Strawson's account as suggested in the last section, we arrive at the following hypothesis:

(Acceptability of 'Might be')

> A sentence of the form 'It might be that *p*' is acceptable for an epistemic subject *s* at time *t* just in case, at *t*, *s* gives some subjective chance to the proposition that *p*.

This way of specifying the acceptability conditions seems plausible: accept that it might be that *p* just in case you have some evidence for the proposition that *p*, be it as small as it may.

The thesis about the acceptability conditions of present tense 'might'-statements has a natural counterpart:

(Rejectability of 'Might be')

> A sentence of the form 'It might be that *p*' is rejectable for an epistemic subject *s* at time *t* just in case *s* is at *t* certain that not *p*.

It seems that the specification of the rejectability conditions is equally plausible as the thesis about the acceptability conditions. If one is certain that they are away, one is in a position to reject that they might be at home. Conversely, if one rejects that they might be at home, it seems to be rationally required that one is certain that they are away.

Interestingly, the rejectability conditions come very close to the negation of the acceptability conditions. This is surprising, for it is not true *in general* that a statement is rejectable if it is not acceptable. After all, there are many statements about which we should suspend judgement in certain epistemic situations. For instance, I am now in an epistemic state in which it is rationally required to suspend judgment about Goldbach's conjecture. Thus, two interesting questions arise:

(i) Are present tense 'might'-statements always either acceptable or rejectable?

(ii) Does it follow from our specification of the acceptability and rejectability conditions that 'might'-statements are always either acceptable or rejectable?

There is evidence that present tense 'might'-statements are *almost always* either acceptable or rejectable. However, we shall show that it does not *follow* from our specification of the acceptability and rejectability conditions that this must *always* be the case. The reason for this will point to a class of cases which may generate instances of epistemic situations in which a present tense 'might'-statement is neither acceptable nor rejectable.

The evidence for the claim that present tense 'might'-statements are in many cases either acceptable or rejectable derives from the observation that acceptance of them tends to be an all-or-nothing matter. We are rarely or never inclined to qualify a present tense 'might'-statement with an epistemic use of 'probably'. Also, we do not apply 'might' to other 'might'-statements. There is something odd about the following sentences:

(8) Probably, it might be that they are at home.

(9) It might be that it might be that they are at home.

If acceptance of present tense 'might'-statements were a matter of degree, we should expect relatively frequent uses of sentences like (8) and (9). But these forms are virtually uninstantiated. So, there is evidence for the claim that present tense uses of 'might' tend to be either acceptable or rejectable.

However, the general claim that present tense uses of 'might' are always either acceptable or rejectable does not follow from the acceptability and rejectability conditions as specified above. This is because there is no good reason to suppose that we are always in a position to assign a definite credence to every proposition. Rather, there may well be situations in which rationality leaves it open whether we should assign a positive credence to a certain proposition or not. Sometimes we just do not have any evidence for or against a proposition. And in such cases it may be that we do not need to assign any credence to that proposition, or at least not a definite one. But then there would be cases in which we are neither certain of the proposition nor do we assign a positive credence to its negation. Application of our theses would yield that in such a situation we should suspend judgment about the corresponding present tense 'might'-statement, since we neither give it a positive credence nor are we certain about its negation.

In conclusion, present tense 'might'-statements tend to be either acceptable or rejectable, but it is consistent with their acceptability and rejectability conditions that there are cases in which we should suspend judgement about them.

2. What Might Have Been

2.1 The Proposal(s)

Let us now come to Strawson's account of 'might'-statements in the past tense, statements about what might have been the case.

Strawson did not develop a wholly general analysis of 'might have been'-statements; instead, he only treated statements of the form '*a* might have φ-d'. We should certainly expect a good analysis of such statements to be applicable, *mutatis mutandis*, to statements of other forms, such as 'someone might have φ-d', 'there might have been no *F*s' etc. But let us put aside the question of how a generalized version of his account would look like and focus on the material he did present, the core of which is stated in the following passage:

> Some proposition to the effect that *a* might have φd is true (acceptable) if and only if there was some point in the history of the individual concerned such that presently available knowledge regarding that point does not permit the rational inference that *a* did not φ (or, in other words, the facts as we know them left open a chance that *a* would φ). (183)

With '<' as an expression of the parthood relation, and '⊨' as an expression for *weak* implication (such that given the premises, the conclusion is *beyond reasonable doubt*), we can construe Strawson's twofold proposal as follows:

(Truth) '*a* might have φ-d' is *true* ↔
∃t (t < the history of *a* & present knowledge about *t* ⊭ *a* did not φ)

(Acceptability) '*a* might have φ-d' is *acceptable* ↔
∃t (t < the history of *a* & present knowledge about *t* ⊭ *a* did not φ)

For the moment, we will concentrate on the claim about truth conditions (we come back to the acceptability conditions at the end of section *c.*).[5] To come to a judgement about the proposal, two issues have to be addressed. First issue: *if* the general idea behind the proposal is correct and 'might have been'-statements express epistemic possibilities, the question remains whether Strawson got the details right. As will be seen, there are some problems with his account. Second issue: *is* the general idea behind the proposal correct? do we talk about epistemic possibilities when we talk about what might have been? There are reasons for doubt.

[5] We will later argue that 'might have been'-statements do *not* express epistemic possibilities but rather objective ones. This is why we are *not* sceptical as to whether they possess truth conditions, contrary to the case of 'might'-statements.

2.2 Struggling With the Details

Let us start with the first issue. The basic idea of Strawson's proposal is that a 'might have been'-statement expresses uncertainty based on insufficient knowledge. The two crucial questions then are:

(i) Why should it be uncertainty based on insufficient *knowledge* (instead of some other mental states)?

(ii) *Which* knowledge should be relevant for a given 'might have been'-statement.

The latter question can be divided into two more specific questions:

(ii.a) *Whose* knowledge should count?

(ii.b) What *kind* of knowledge should count?

We will address these questions in turn.

Re (i): Why should it only be *knowledge* that is relevant for the uncertainty expressed by a 'might have been'-statement? Imagine a speaker, Fred, is the subject of a Gettier case and fancies to know that *p*, while in fact he only justifiedly and correctly believes that *p*. Assume further that the belief that *p* concerns a certain time *t* and makes it certain for Fred that *a* would not φ at any time after *t*. Asked whether *a* might have φ-d (judged from time *t*), Fred would presumably answer in the negative; given his epistemic background it is certain that *a* would not φ. But on Strawson's analysis, Fred would be wrong; while his *beliefs* make it certain that *p*, his *knowledge* does not. This suggests that Strawson's talk about knowledge is too narrowly focussed. If 'might have been'-statements express epistemic possibilities it should rather be replaced by 'belief'. What would then count is that the speaker can rationally infer from whatever she accepts concerning the time in question that *a* did not φ.[6]

Re (ii.a): For the evaluation of a 'might'-statement in the present tense, we evidently rely on *present* epistemic states (of whatever subjects are relevant). Analogously, when it comes to past statements, one might think that *past* epistemic states (of then existing subjects) are relevant. Not so, says Strawson (181ff.). For assume someone says about an election, 'Given the actual composition of the committee, candidate *X* couldn't have won.' This does not imply that it was certain for anyone during the elections that *X* would not win. So, the relevant epistemic subjects should not be drawn from those that

[6] Notice that a possible motivation to speak only about knowledge may stem from Strawson's wish (which we criticise below) to count some epistemic subjects apart from the speaker as relevant for the evaluation of a 'might have been'-statement. For while it is safe to pool the knowledge of different subjects, pooling their beliefs will often lead to inconsistent systems.

were around at the time to which the statement is pointing. Rather, what counts are again *present* epistemic states. There is, by the way, a much more straightforward argument to the same conclusion: we can make true and false 'might'-statements with respect to times at which no sentient beings existed at all. Whoever wants to interpret such statements as expressions of epistemic (un)certainty should better rely on present subjects and their epistemic states, then.[7]

But which subjects exactly? As already seen, Strawson thinks that the subjects relevant to the evaluation of present tense statements are we '*ourselves*, the speaker and his circle and others he regards as authoritative, perhaps' (180). But we argued in section 1.2. that other people's epistemic states do not play the role Strawson thinks they do. In a similar fashion, one can argue that other people's epistemic states are not relevant for the evaluation of 'might have been'-statements either. If we are to find out whether *a* might have φ-d, we do not typically investigate into the epistemic states of other people. But if the truth conditions of 'might have been'-sentences were to involve epistemic states of others, we would expect such behaviour. So, the relevant epistemic subject should again be the speaker alone.

The account would then run as follows:

(Truth*) '*a* might have φ-d' as uttered by *x* is *true* ↔
∃t (t < the history of *a* & *x*'s present beliefs about *t* ⊭ *a* did not φ)

Unfortunately, this result will create a much more serious problem for Strawson; we will come back to it in a minute. But first, we will take a look at some more details.

Re (ii.b): Can the account work if *all* present knowledge about *t* is relevant to the evaluation of '*a* might have φ-d', or does some knowledge have to be filtered out? Strawson's answer is as follows:

> I think we go yet further in our verdicts on past possibilities, taking into account not only the evidence, the particular facts, collectively available at the time but at the time uncollected, but also general truths now known but then unknown, and even particular truths relating to that time now known but then unknown. (182)

But by not limiting the body of relevant knowledge any further, Strawson apparently threatens the whole account. For assume a speaker knows that *a* φ-d at a given time *t*. Then the speaker also knows about any earlier time *t** that it was such that, at it, *a* was going to φ in some more or less distant future. So, the speaker knows enough about any time before *t* to make it certain that at that time, *a* was going to φ. Therefore, it turns out that on Strawson's analysis, any speaker who knows that *a* φ-d will say something wrong when he utters '*a* might have not φ-d'.

---

[7] Although Strawson mentions such statements in a footnote (182, n. 1), he does not put much weight on them.

This is certainly an unintended result. It shows that for the proposal to get off the ground we have to sort out a portion of a speaker's knowledge which will then count as relevant for the uncertainty expressed by 'might'-statements. A straightforward idea is that for a statement concerning a time *t* only such knowledge about *t* (or earlier times) is relevant that does not characterise that time in relation to times later than *t* or whatever happened at them. What does not count, for instance, is knowledge that characterises *t* as the last day before the war, or that characterises some events that occur at *t* in terms of their effects, etc.

However, Strawson himself mentions another problem of his account which is somewhat intertwined with the currently discussed one. Assume Fred is a convinced determinist. He thinks that for every event *e* which happens at some time *t*, there are truths about any earlier time *t** that imply, in conjunction with the laws of nature, that *e* happens at *t*.[8] Now Strawson notes that if Fred knows that *a* φ-d, then he will believe that at any earlier time there were some facts which ensured that *a* would φ. So, with respect to any time Fred will be certain that *a* would sooner or later φ, and thus, on Strawson's analysis, it seems that Fred should never utter '*a* might have refrained from φ-ing'. Strawson does not welcome this apparent outcome of his account, but he thinks it can be avoided:

> The objection construes 'now available knowledge' too widely. What is required […] is some more specific knowledge of conditions obtaining at the relevant time than can be derived from the premises that *a* did not in fact φ […] coupled with the conviction that *a*'s φing or not φing is subject to deterministic laws. (184, n.)

Unfortunately, Strawson's reply is hardly helpful: what exactly counts as 'more specific knowledge of conditions obtaining at the relevant time'? Let us see whether the filter suggested above may do the deed. To see this, we have to get clearer about what the determinist is supposed to infer about the relevant time. Apparently it is the knowledge that *at t, some conditions obtained which inevitably bring it about that sooner or later* a *will* φ. Now this is relational knowledge about *t*, which characterises conditions obtaining at *t* in terms of their future effects. So, the above filter sorts out this kind of inferred knowledge and the determinist may still say that *a* might have φ-d, even if *a* in fact did not φ. The suggested filter, therefore, seems to be faithful to Strawson's intentions.

However, there may be something wrong with his intentions in the first place. He ends his discussion of the determinist by saying that

> There is no reason to think that convinced universal determinists would (or do) eschew the idiom in the sense expounded or confine themselves to denying others' uses of it. (184, n.)

---

[8] For such a notion of determinism see van Inwagen (1983: 3).

The present proposal manages to make the idiom available to the determinist. But whether the determinist would want to make the relevant 'might'-statements seems much more doubtful to us than to Strawson. Take a determinist and a closed physical system *S*, about which the determinist knows the internal states at some time *t*. He moreover knows some deterministic laws which, together with the knowledge about the state of *S*, imply that *S* will φ at some later time. Then, according to the present proposal, he is not justified in saying that *S* might have refrained from φ-ing (after *t*); and it seems a determinist would in fact not want to say that. But Strawson wants to grant him such 'might'-statements if he has less specific knowledge. However, would not a determinist typically just say that, given some prior state of the world, whatever happened afterwards, had to happen, and that nothing might have happened that in fact did not happen? If so, the suggested filter is too strong; but it is hard to see what a filter might look like that avoids that we generally have to accept the inference '*a* did not φ, therefore *a* could not have φ-d' but which grants the determinist the inference '*a* did not φ and *a*'s φ-ing or not φ-ing is subject to deterministic laws, therefore *a* could not have φ-d'.

Since we have nothing better to offer, we must leave this problem of Strawson's account unsolved: it either requires the determinist to make a hardly reasonable distinction between certain 'might have been'-statements, or it needs a modification other than the suggested one, in order to avoid collapse. As we will now argue, it is actually not that bad that we cannot solve the problem – for we think that Strawson's account is flawed at a more general level. But then, there is no need to straighten out the details any more.

2.3 Against the General Idea

So we now come to the second issue we promised to address: is the general idea behind Strawson's proposal correct, i.e. do we express uncertainty of any sort when we talk about what might have been? Our answer will be based on two observations:

(i) 'might have been'-statements are generally *corrigible*, and

(ii) one may, in general, be *agnostic* with respect to a given 'might have been'-statement.

As we will now argue, an epistemic account of 'might have been'-statements cannot do proper justice to these observations and should accordingly be rejected. In the end of the section we will explain how 'might have been'-statements should be understood if not epistemically. But first, let us describe the two features of 'might have been'-statements in more detail.

(i) *Corrigibility*: 'might have been'-statements are in general subject to *correction*, which means in the light of counterevidence they will be *revised* and counted as *mistaken*. An example for illustration: Assume Fred says that

(10) Last night, Ann might have won the race.

What Fred does not know is that poor Ann had an accident in the morning, rendering her virtually immobile and excluding any possibility of her participating in the race, let alone of winning it. Once we inform Fred of these circumstances, he both revises his statement *and* counts his former statement as mistaken: 'Oh, now I see that I was wrong; last night, Ann couldn't have won the race.' We take it that corrigibility, thus understood, is a general feature of 'might have been'-statements.

(ii) *Possibility of an agnostic stance*: we do not always accept or reject a 'might have been'-statement right away, but sometimes withhold judgement. Four weeks after the described accident, there is another race, and on the following day Fred is again confronted by an utterance of (10). As Fred is uncertain whether Ann's legs have already healed or not, he is uncertain whether Ann was in a condition to participate in the race. He does think that if she was healthy, she had a good chance to win. In this situation Fred will – at least if he is rational – neither accept nor reject the 'might have been'-statement. Rather he will reply: 'Perhaps you are right, but, actually, I do not know whether she might have won or not.'

We take it that 'might have been'-statements generally possess the said features (corrigibility and possibility of an agnostic stance). But if they possess them, they cannot be understood epistemically. For, *if* such a statement expressed uncertainty on the part of the speaker, it would – recherché cases aside – be incorrigible. Of course, new evidence may give the speaker new certainties or remove old ones and thereby still lead to a *revision* of a statement about the speaker's certainties – but it will not lead to a rejection of the former statement as *mistaken*. This can be seen from explicit statements about certainties: assume in our first scenario (after the first race) Fred said

(11) In light of all my present knowledge about yesterday, I am not certain that Ann did not win the race.

Being informed of Ann's accident, he revises his statement:

(11*) In light of all my present knowledge about yesterday, I am *now* certain that Ann did not win the race.

But Fred does *not* count his former statement as mistaken; why should he? He only said that the knowledge he then had did not make it certain that Ann did not win, and that was quite correct.

The possibility of an agnostic stance equally counts against the epistemic interpretation of 'might have been'-statements, for a speaker generally knows – borderline cases set aside – whether he is certain about some fact, or not. So if a 'might have been'-statement expressed uncertainty on the part of the speaker, it should not be open for an agnostic stance.

Admittedly, our arguments so far are directed only against a certain kind of epistemic interpretation of 'might have been'-statements, i.e. those that take them to express *personal* (un)certainties, based only on the knowledge of the *speaker*. However, we mentioned before that Strawson thinks the relevant knowledge is not merely that of the speaker, but that it additionally includes knowledge of his circle and/or of those he accepts as authorities. Such an inclusion would make room for corrigibility and agnostic judgements; for Fred's utterance of (10) would then amount to:

(10*) The knowledge available to me, or you, or other authorities, does not make it certain that Ann did not win the race.

After being told that the addressee of Fred's statement knows that Ann was incapacitated, Fred can correctly count his former statement as false. Since speakers do not generally know what knowledge a hearer or an alleged authority has, he may (and even should) be agnostic about certain 'might have been'-statements; namely about those concerning which there might be knowledge of others that would provide him with a new certainty or undermine the one he has.

But still the account is no good, for a reason we already stated earlier: an account which takes knowledge of other people as relevant will require speakers to be agnostic about *way too many* 'might have been'-statements. So, while integrating other people's knowledge into the account avoids our counterarguments from corrigibility and possibility of an agnostic stance, it does so only at the cost of falling prone to the argument of too much agnosticism. Moreover, although Strawson's original proposal allows for corrigibility and agnosticism, it makes false predictions about the *reasons* we accept as relevant: in general, we neither revise a 'might have been'-statement nor do we take an agnostic stance towards them *because of* other people's epistemic states.

Therefore, we regard Strawson's proposal as mistaken: 'might have been'-statements do not express present uncertainties of any kind, in the sense that their truth conditions cannot be spelled out in terms of someone's being (un)certain about something.

We may now come back to the second part of Strawson's proposal, which concerns the acceptability conditions of 'might have been'-statements. As far as we see, none of the worries we articulated speaks directly against *that* claim. So, what if it were correct? Would this be good news for Strawson? Could 'might have been'-statements then still be said to express uncertainties? No! For notice that quite generally the acceptability

conditions of *any* kind of statement may be spelled out with respect to the beliefs of the speaker and their degrees of firmness. What makes

(12) 4 + 4 = 8

acceptable for some speaker *S* is that *S* believes that four plus four equals eight. That does not make arithmetical statements in *any* good sense epistemic statements or expressions of beliefs; it only shows that in a sincere assertoric utterance of an arithmetical sentence (or any other sentence), a speaker gives voice to one of his beliefs. So, while we think that Strawson's account of the acceptability conditions of 'might have been'-statements has a fairly good standing, it does not qualify as an epistemic account of such statements.

Let us finally make some remarks on how we do think that 'might have been'-statements are to be understood. They assert that there was a real possibility of something's happening, or that there was an objective chance for some event:

(Truth†) '*a* might have φ-d' is *true* ↔
∃t (t < the history of *a* & at *t*, there was an objective chance that *a* would φ).

If this correctly describes their truth conditions (a small modification is yet to come), Strawson's account of their acceptability conditions may at least point in the right direction: assume a speaker has (sufficiently substantial) knowledge about some time *t*, which does not suffice to make it certain for her that, from *t* onwards, *a* would not φ. This will, *ceteris paribus*, be a good basis for the speaker to think that, at *t*, there was a genuine possibility that *a* would φ. So, if the truth conditions of 'might have been'-statements are spelled out via *objective* chances, it may be promising to approach their acceptability conditions via *subjective* chances which correspond to uncertainties.

3. What Could Not Have Been

By focusing on an alleged non-metaphysical reading of 'might have been'-statements Strawson made an attempt to de-mystify them:

> My subject in this paper is particular possibilities: the may-bes and the might-have-beens that relate essentially to particular individuals or situations. […] Some detect, or think they detect, an intoxicating scent of something more metaphysically interesting than either merely epistemic possibilities on the one hand or merely *de dicto* possibilities or necessities on the other. My remarks will not give much satisfaction to them. (179)

We have already seen that Strawson's hypothetical antagonists had a better nose than he himself was willing to concede: contrary to what he thought, 'might have been'-statements express something metaphysically more heavyweight than epistemic and *de dicto* possibilities (namely objective chances). Not only was he wrong on that account, but he also underestimated the potential significance of his own considerations to friends of metaphysical modality. For, at the end of his paper (see 185ff.), we find a discussion of why we should accept certain negated 'might have been'-statements which, on the face of it, are apt to express instances of *essentialist* theses, e.g.

(13) Aristotle couldn't have had different parents; and

(14) This table couldn't have been made of marble instead of wood.

(13) and (14) are clearly not expressions of *de dicto* impossibilities.[9] If we are right, they do not express epistemic impossibilities either, but rather (the absence of) objective chances. But the justification Strawson offers for accepting such sentences as (13) and (14) is independent of what kind of chances – subjective or objective – 'might have been'-statements express. Consequently, if it goes through, Strawson has offered nothing less than a justification for something as metaphysically interesting as the objective chance readings of (13) and (14).

In what follows we will introduce another modification to the truth conditions of 'might have been'-statements that is necessary in order to deal with possible-non-existence claims. Then, we will outline a Strawsonian justification of Origin Essentialism that becomes available through this modification. Finally, we shall criticise the proposal.

[9] For a standard definition of modality *de dicto* and *de re* see e.g. Forbes (1985: 48).

3.1 Another Modification

What we salvaged from Strawson's discussion of 'might have been'-statements were the following truth conditions for them:

(Truth†) '*a* might have φ-d' is *true* ↔
∃t (t < the history of *a* & at *t*, there was an objective chance that *a* would φ).

We retained from Strawson's original proposal to explicitly restrict the relevant times to those during the history of the object in question. But while it is plausible that

(15) Aristotle might have become a soldier;

is true iff there was a time in Aristotle's life at which there was a chance that he would become a soldier, there was never a time in Aristotle's life at which there was a chance that he had never existed. Nevertheless, it is true that

(16) Aristotle might have never existed.

The same holds for any ordinary object: animals, tables and libraries could have failed to exist. But while animals could have failed to be born, and tables and libraries could have failed to be built, there has never been a time during any animal's (table's, library's) life at which there was any chance that it failed to have been born (built). Hence, as Strawson realises himself (cf. 184f.), no account that restricts its purview to times during the life of an object can adequately deal with non-existence 'might have been'-statements and those that entail them:[10] according to (Truth†), all of them should be false.[11] However, many of them, those that deal with common or garden material objects, are straightforwardly true. Something needs to be done.

Strawson's quite intuitive idea is to appeal to the *pre-history* of the objects in question: though there was no time during Aristotle's life at which there was a chance that he had never existed, there was a time *before* he was born at which it was not settled that he would exist – his mother might have miscarried, or his parents may even have never met at all. In such a case, Aristotle would have never existed. Consequently, it is reasonable to maintain that chances concerning relevant *other* things in the pre-history of Aristotle are responsible for the truth of (16). The same holds, *mutatis mutandis*, for non-existence 'might have been'-statements concerning other animals, tables and

[10] Strawson's assumption that 'Aristotle has never been born' entails 'Aristotle has never existed' is wrong (see *Macbeth*), but nothing hangs on the particular example.

[11] The problem is even more pervasive than Strawson himself realises. It is not only non-existence 'might have been'-statements that force us to look into the prehistory of an object. 'Socrates might have been conceived in a different country', for instance, equally requires such a move.

libraries. Hence, to cover these 'might have been'-statements the restriction to times during the life of the object concerned should be lifted:

(Truth‡) '*a* might have φ-d' is *true* ↔
∃t (at *t*, there was an objective chance that *a* would φ).[12]

But what *are* the other objects and chances involving them that are relevant to the question of whether something might have never existed? Strawson discusses the case of objects that come into existence by a process of natural generation (e.g. plants, animals) and artefacts. According to him, it is chances involving an animal's (Aristotle's) progenitors, the material of which the artefact 'of a fairly standard kind' (a table) is made, and the design of the 'more elaborate' artefact (a library) that are relevant to the question of possible non-existence.[13] Aristotle might have never existed because his parents might have never met. This table might have never existed because the piece of wood it is made of could have been used to build a chest of drawers instead, or, indeed, nothing at all. The Old Bodleian may have failed to exist because the plans for it might have never been realised.

3.2 Strawson's Case for Origin Essentialism

Strawson is certainly right that these are the kinds of considerations that convince us of the truth of a non-existence 'might have been'-statement. What is interesting about them is that they are closely linked to essentialist theses, more specifically: Origin Essentialism. The orthodox view about essentialist theses is that they are analysable in terms of necessity and existence:[14] Aristotle is essentially human just in case it is necessary that he is human provided he exists at all. Aristotle essentially has the parents he actually has just in case he must have them if he is to exist.

This last thesis is an instance of a kind of Origin Essentialism, defended by e.g. Kripke (1980: 112) and Forbes (1985: ch. 6):

(OE1) Things that come into existence by a process of natural generation are necessarily such that if they exist, they have the progenitors they actually have.

Other varieties of origin essentialism, also defended by Kripke (113f.) and Forbes (loc. cit.), concern artefacts like tables:

---

[12] It might be desirable to allow the quantifier ranging over times to receive some restrictions from the context of utterance.

[13] Cf. 186f.

[14] Cf. e.g. Mackie (2006: 3). For the heterodoxy, see Fine (1994).

(OE2) Tables are necessarily such that if they exist, they are initially made of the material they are actually initially made of,[15]

and Strawson's 'more elaborate constructions' like buildings:

(OE3) Buildings are necessarily such that if they exist, they are made according to the plan they are actually made according to.[16]

The truth of (OE1) to (OE3) would explain why considerations like the ones alluded to are relevant for evaluating non-existence 'might have been'-statements. If Aristotle couldn't have existed unless as the child of his actual parents, what his parents used to do made all the difference to whether Socrates would come into existence. If they had not fulfilled their part, nothing else could have happened that would have made it the case that Socrates existed nonetheless. If this table couldn't have existed unless it was initially made of the piece of wood it was actually initially made of, what happened to that piece of wood made all the difference to whether the table would come into existence. If the piece of wood had not fulfilled its part, nothing else could have happened that would have made it the case that the table existed nonetheless. And finally, if the Old Bodleian couldn't have existed unless it was made according to Sir Giles Gilbert Scott's plan, what used to happen to that plan made all the difference. If it had not done its share, nothing else could have happened that would have seen to it that the Old Bodleian existed nevertheless. If Origin Essentialism is true, our reasons for accepting the reasons we do accept for non-existence 'might have been'-statements are fully vindicated.

But maybe things also work the other way. It is sufficiently clear that, given how our 'might have been'-statements work, the reasons we accept as pertinent for evaluating the non-existence 'might have been'-statements are indeed pertinent. Maybe this part of our practice can *explain* why Origin Essentialism has the solid philosophical standing it enjoys. If we keep parentage, initial composition and plan fixed whenever we are forced to consider the pre-history of something in order to evaluate a 'might have been'-statement, what better explanation is there for the pertinence of these considerations than that Origin Essentialism is indeed true? If this is so, the justificatory detours through the pre-histories of the subjects of non-existence 'might have been'-statements afford a novel justification for (some varieties of) Origin Essentialism. Its truth could be seen to be presupposed by our successful handling of an important class of 'might have been'-

---

[15] This is not quite right. It is usually held that these artefacts must be initially made of *most* of the material they are actually made of - but not necessarily all – if they are to exist. This is more plausible than (OE2) but leads to considerable complications; see e.g. Chandler (1976).

[16] Cf. also Salmon (1979: 757f.).

statements. Our modification of (Truth†), anticipated by Strawson, would then shed considerable light on those features of our everyday modal idioms that account for the intuitive plausibility of an important variety of essentialist theses.

3.3 Against Strawson's case for Origin Essentialism

However, neat as this justification would be if only it were true, the truth of Origin Essentialism is *not* the best explanation for the pertinence of our reasons for accepting non-existence 'might have been'-statements.

What should make us wary to begin with is that the Strawsonian proposal overgenerates justified essentialist theses. This table might have failed to exist because the piece of wood it is made of might have been thrown away instead. True. But surely this is not the only possible reason for accepting the non-existence 'might have been'-statement. Here are a few more:

(a) The carpenter who built it might have opted to switch profession shortly before the time he actually built the table,

(b) He might have been killed,

(c) He might have been ill the day he actually built the table,

(d) His workshop might have burned down.

But if the pertinence of considerations involving the piece of wood showed that (OE2) was true, the pertinence of (a) to (c) would show that it is essential to the table that it be built by whoever actually built it. While that thesis may still be remotely plausible, the more specific theses also in play are not: it is certainly not essential to the table to have been built by a carpenter (a), to have been built the very same day it was actually built (c), to have been built in the workshop it was actually built in, or, indeed, in any workshop at all (d). But again, if the pertinence of the claim that the piece of wood might have been thrown out instead of going into the construction of a table is evidence for the truth of (OE2), the pertinence of (a) to (d) should be evidence for far more eccentric essentialist theses. Since it is not, we have reason to assume that something has gone wrong in the proposed justification of Origin Essentialism.

When we reconsider (a) to (d) it is relatively straightforward to see *why* they are pertinent to the question of whether the table might have never existed. The carpenter who built the table could have opted to switch profession before building the table. Had he done so, he himself would not have built the table and neither would have anyone else. So, if he had left the profession, the table would not have been built, i.e. it would not have existed. Similar things can be said about (b) to (d). Quite clearly, (a) is a reason

for affirming the non-existence 'might have been'-statement because the counterfactual 'if the carpenter had switched profession, the table would not have existed' holds. Now, a counterfactual *may* hold because a corresponding strict conditional holds. To put it in terms of possible worlds: 'p' may hold at the closest possible worlds at which 'q' holds because 'p' holds at *all* worlds at which 'q' holds. But this is the limiting case. Typically, whether 'p' would hold, if 'q' did is *not* independent of other things that are the case. This is certainly so in the case of the true counterfactual

(17) If the carpenter who actually built the table had switched profession, no one would have built *a* table just like the one in question.

As things are, (17) is true. But if our carpenter had had a co-worker with a tendency towards plagiarism, the co-worker would have built a table just like the one in question. Hence, (17) *would have been* false, and the corresponding strict conditional – 'necessarily, if the carpenter left the profession, no one else built a table just like this one' – *is*. Moreover, it is not clear that in the case in which there is a plagiarising co-worker present (a) would count as good evidence for the non-existence 'might have been'-statement: even if the carpenter who built the table had left the profession, his co-worker would have built a table of the same design from the very same piece of wood. Would it have been *this very table*? At the very least: it is not clear that it would not have been.

The problem with the proposed justification of Origin Essentialism is now in full view. That chances involving Aristotle's parents (the piece of wood the table is made of, the design of the Old Bodleian) are reasons for affirming the non-existence 'might have been'-statements about Aristotle (the table, the Old Bodleian) may be *merely* due to the truth of the corresponding counterfactual. Just as in the case of (a), their pertinence may *not* be underwritten by the truth of the corresponding *strict* conditional, i.e. by the truth of (an instance of the relevant variety of) Origin Essentialism.

Let us sum up. The phenomenon that for the truth of certain 'might have been'-statements chances in the pre-history of their subject are pertinent promised a neat justification for Origin Essentialism: its truth, we hypothesized, is the best explanation for the pertinence of these chances. However, we saw that by the same token some essentialist theses can be justified that are clearly false: the pertinent chances are pertinent not because their actualisation *necessitates* the non-existence of the thing in question, but because the corresponding counterfactual holds. Hence, the Strawsonian considerations *by themselves* are not able to justify Origin Essentialism: just because chances involving the piece of wood it is actually made of are pertinent to the question of whether the table might have not existed, does not mean that the table *must* be made of that piece of wood if it is to exist at all – just as the pertinence of chances involving

the profession of the man who built it does not ensure that it is essential that the table was built by a carpenter.

However, we still feel that there is sufficient reason to end on a happy note. For, it seems to us that there may be ways of exploiting the differences between the cases. To conclude, let us sketch how this might be done. As we have seen, that the carpenter who built the table could have switched profession would not have been a reason for affirming the non-existence 'might have been'-statement had there been a plagiarising co-worker around. Similarly, for (b) to (d), given that no essentialist thesis underpins their pertinence, there are other possible circumstances such that, had they obtained, (b) to (d) would not have been good reasons for affirming the non-existence 'might have been'-statements – e.g. if the carpenter had not typically worked in the workshop the table was actually built in. If, on the other hand, for all possible circumstances, some pertinent consideration is still pertinent in them, the best explanation for *that* is still that the corresponding essentialist thesis holds.[17] Take Aristotle and consider the possibility that his parents had never met: no matter what else might or might not have been the case, that his parents might have never met will always be pertinent to whether Aristotle might have failed to exist.[18] If this is so, non-existence 'might have been'-statements may still have a role to play in our justification of Origin Essentialism – it just is not as straightforward as we might have thought on a first reading of Strawson's paper.

---

[17] This is a reflection of the fact that '$\forall r\,[(\sim q \,\&\, r) \,\Box\!\!\rightarrow \sim p]$' entails '$\Box\,(p \rightarrow q)$'.

[18] How do we know? By considering relevant possible circumstances. If we cannot come up with possible circumstances in which some pertinent consideration becomes non-pertinent, even though we have tried long and hard enough, this is evidence for thinking that it is underpinned by the truth of the corresponding essentialist thesis.

Literature

Chandler, H. (1976): 'Plantinga and the Contingently Possible', *Analysis* **36**, 106–9.

DeRose, K. (1991): 'Epistemic Possibility', *Philosophical Review* **100**, 581–605.

Egan, A. & Hawthorne, J. & Weatherson, B. (2005): 'Epistemic Modals in Context', in: G. Preyer & G. Peter (eds) (2005): *Contextualism in Philosophy*, Oxford: Clarendon Press, 131–68.

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

Forbes, G. (1985): *The Metaphysics of Modality*, Oxford: Clarendon Press.

Krikpe, S. (1980): *Naming and Necessity*, Oxford: Blackwell.

Margalit, A. (ed.) (1979): *Meaning and Use*, Dordrecht: Reidel.

Salmon, S. U. (1979): 'How Not to Derive Essentialism from the Theory of Reference', *The Journal of Philosophy* **76**, 703 –25.

Strawson, P.F. (1979) 'May Bes and Might Have Beens', in: A. Margalit (ed.) (1979): 229–38. Page numbers refer to the reprint in: Strawson, P. F. (1997): *Entity and Identity*, Oxford: Clarendon Press, 179–90.

Van Inwagen, Peter (1983): *An Essay on Free Will*, Oxford: Clarendon Press.