ARTICLE

Williamson on Defining Knowledge

Manuel Pérez Otero*

University of Barcelona/LOGOS, Barcelona, Spain
*Email: perez.otero@ub.edu

Abstract

In his outstanding book *Knowledge and its Limits* (2000), Williamson (a) claims that we have inductive evidence for some negative theses concerning the prospects of defining knowledge, like this: knowing cannot be defined in accordance with a determinate traditional conjunctive scheme; (b) defends a theory of mental states, mental concepts and the relations between the two, from which we would obtain additional, not merely inductive, evidence for this negative thesis; and (c) presents an alternative (non-traditional-conjunctive) definition of knowledge. Here I consider these issues and extract two relevant conclusions: (i) Williamson's theory of states and concepts only supports the negative thesis because this theory would explain too much, since it imposes implausible necessary limitations on possible uses of concepts and linguistic expressions. So, there is no appropriate non-inductive evidence for the negative thesis. (ii) Williamson's own definition of knowledge is at risk.

Keywords: analysis; mental concepts; factive states; true belief; conjunctive concepts; evidence

Introduction

In his outstanding book *Knowledge and its Limits* (Williamson 2000), Williamson claims that we have inductive evidence for some negative theses concerning the prospects of defining knowledge. One of them is that knowing cannot be defined in accordance with a determinate traditional conjunctive scheme. Furthermore, he defends a theory of mental states, mental concepts and the relations between the two, from which we would obtain additional, not merely inductive, evidence for this negative thesis. Meanwhile, Williamson presents an alternative (non-traditional-conjunctive) definition of knowledge.

Here I consider these issues and extract two relevant conclusions. First, Williamson's theory of mental states and mental concepts only supports the negative thesis because this theory would explain too much, since it imposes implausible necessary limitations on possible uses of concepts and linguistic expressions. In that sense, there is no appropriate non-inductive evidence for the negative thesis. Second, Williamson's own definition of knowledge is at risk.

In section 1, to a first approximation, I present a key idea repeatedly invoked by Williamson to articulate his view: knowledge is non-compositional. I formulate different negative hypotheses about the possibility of defining knowledge in section 2, indicating which of them can be attributed to Williamson. Of those hypotheses, I will

mostly focus on what here I will call *Unanalysability of the State*. A few methodological remarks on the distinct kinds of evidence for Unanalysability of the State are important for a proper understanding of the overall dialectics (section 3). In section 4, I present Williamson's theory of mental concepts and states, called here *WTM*, as well as some of its consequences. Then, I introduce some objections to WTM (section 5). The key argument for my main thesis – WTM only supports Unanalysability of the State insofar as it also supports other claims that seem highly implausible – starts in section 6, and it involves a metaphysical-anthropological hypothesis about potential uses of concepts and languages. Some possible responses – on behalf of Williamson – and counter-responses are sketched in section 7. Finally, in section 8, I present Williamson's definition of knowledge and I comment on relevant corollaries of the previous discussion that jeopardize it.

1. The alleged non-compositionality of knowledge (first approximation)

Knowing entails believing, knowing entails truth, and true belief is insufficient for knowing; Williamson accepts all of this. So, we might think that there must be an "additional part" which conjoined with belief and truth would amount to knowledge. Let us call this "remaining part" of knowledge F. A key idea that forms part of Williamson's view is the rejection of the existence of such a third component, F, if it is conceived as a non-redundant factor in a possible definition of knowledge. We can put it in other words: if (*) is a correct definition of knowledge,

   (*) Subject S knows that P if and only if S believes that P, P is true and Q

then Q is a redundant condition, in the sense that Q – by itself – entails the truth of P. Traditional approaches to defining knowledge have proposed characterizations that fit the template (*). Furthermore, according to Williamson, such traditional approaches have looked for a Q that is a non-redundant factor (so that neither of the other two factors – P is true and S believes that P – is entailed by Q). But, Williamson takes that tradition to be doomed to failure. In that sense, knowing could not be analysed by discomposing it into different independent components.

   Williamson's main negative thesis on the prospects of defining knowledge has to do with this idea. Things are somewhat more complicated, however, which makes the task of identifying that negative thesis a hard one. Some other theses of his, closely related to the previous remarks, are also relevant to this discussion. For instance, suppose that (*) is a correct definition, where Q is a redundant factor (Williamson does not reject this possibility). Even in this case, Williamson claims, the concept expressed by 'knows that P' (the concept of knowing that P) will not be identical to the conjunctive concept expressed by 'believes that P, P is true and Q'. This other negative result is a consequence of his theory about mental concepts and states (cf. section 4 below). In fact, Williamson defends not just one, but several interrelated negative theses that, when combined, form a general pessimistic picture of our chances of finding an appropriate characterization of knowledge. Since I am going to examine some of these theses I first need to identify them.

2. Alternative versions of Williamson's negative thesis

Williamson relies on a standard distinction between concepts, on the one hand, and states or conditions, on the other.[1] He illustrates the contrast by means of a classical

---

[1]Williamson distinguishes states from conditions (cf. Williamson 2009*b*: 330). However, the difference is irrelevant to the discussion herein.

example. The concept *gold* – i.e., the concept expressed by 'gold' – and the concept *the element with atomic number 79* are two different concepts of the same state, or condition: being gold. And, of course, the metaphysical state, being gold, is not identical to either of those two concepts, or to any other concept of it that may exist. (Cf. Williamson 2000: 29.) Now, the concept *knows* (or the concept *knows that P*, for a given proposition P) is not identical to the metaphysical state of knowing (or to the metaphysical state of knowing that P), which it is a concept of. It is an open possibility that some other concept – different from the concept *knows* – is also a concept of the same state, a concept of the state of knowing.

Suppose a given definition has the following form: S knows that P if and only if F. Let us introduce some terminology. The definition is *reductive* just in case all the concepts in the definiens (the concepts constituting the concept expressed by 'F') are more basic than the concept *knows that P*. It is a *conjunctive* definition just in case the concept F is a conjunctive concept, with the form: C_1 *and* C_2 *and* … C_n (for n > 1). It is a *traditional-conjunctive* definition just in case it is a non-circular conjunctive definition and one of the conjuncts, C_i, is *P is true*. We can now formulate some negative hypotheses about the chances of defining knowledge, or the chances of defining the concept of knowledge (i.e., the concept *knows*). As we will see, Williamson holds all of them, save one.

> *Undefinability*: No informative and non-circular definition will provide correct necessary and sufficient conditions for knowledge. (This is the negative claim that Williamson does not hold.)
> *Unreductibility*: No reductive definition will provide correct necessary and sufficient conditions for knowledge.
> *Unanalysability of the State*: No traditional-conjunctive definition will provide correct necessary and sufficient conditions for knowledge; i.e., there will be no necessarily true and non-circular definition of knowledge of the form "S knows that P if and only if C_1 and C_2 and … C_n", where one of the conjuncts, C_i, is *P is true*.[2]
> *Unanalysability of the Concept*: No traditional-conjunctive definition where definiendum and definiens are the same concept will provide correct necessary and sufficient conditions for knowledge; i.e., there will be no necessarily true and non-circular definition of knowledge of the form "S knows that P if and only if C_1 and C_2 and … C_n", where one of the conjuncts, C_i, is *P is true*, and such that *S knows that P* and C_1 *and* C_2 *and* … C_n are the same concept.

In subsequent sections, I will comment on each of these negative hypotheses. For reasons that will become clear later, I think that Unanalysability of the State is the most important of them, and the best candidate to be considered as Williamson's main negative thesis on the prospects of defining knowledge. In any case, my discussion will mostly focus on Unanalysability of the State. Before entering into that discussion, however, it is convenient to dedicate the next section to some methodological remarks.

3. The evidence against traditional definitions of knowledge

Unreductibility, Unanalysability of the State and Unanalysability of the Concept are supported by Williamson's claims in different ways. On the one hand, he thinks that we have inductive evidence against the three theses: all reductive and traditional-

---

[2]In fact, Williamson also maintains the following stronger thesis (which entails Unanalysability of the State): No non-circular conjunctive definition will provide correct necessary and sufficient conditions for knowledge.

conjunctive definitions of knowledge proposed to date have counterexamples, according to Williamson. In this article, I will not go into the discussion of this particular claim. For the sake of the argument, I grant this history of failures of previous traditional attempts to analyse knowledge. This alleged experience certainly provides support for the three negative theses.

On the other hand, Williamson advances a theory about the mental nature of some concepts, the mental nature of some states, and the links between the two issues. I will refer to his theory as *WTM* (for Williamson's Theory of the Mental). Unanalysability of the Concept is entailed by WTM. Unanalysability of the State is not strictly a consequence of WTM, but if we accept WTM it would be very reasonable – in Williamson's view, at least – to accept Unanalysability of the State (cf. Williamson 2000: 30; I will return to this point in section 4).

Williamson develops his proposals using different argumentative strategies that have not always proved to be sufficiently clear. So much so that Pritchard (2012: note 1), for instance, has interpreted Williamson as attacking the task of searching for an adequate (informative and non-circular) definition of knowledge. That is not Williamson's intention at all, since – as we will see in section 8 – Williamson himself offers and defends his own definition; in other words, Williamson rejects Undefinability.

Not so wide of the mark, Cassam (2009) partially misidentified the kind of evidence Williamson offers in favour of some of his negative theses. Apparently, Cassam focuses on a claim that he calls the *Unanalysability Hypothesis* (UH): the concept *knows* cannot be analysed into more basic concepts (cf. Williamson 2000: 33; Cassam 2009: 12). At first sight, UH would amount to – what I have called – Unreductibility. However, when Cassam examines the issue, the argument of Williamson that he discusses most is not evidence for Unreductibility, but rather for something closer to Unanalysability of the Concept. Cassam calls it the *Distinct Concepts Argument*. This argument takes as premises some theses from WTM, in particular the crucial claim that *knows* is a mental concept, and concludes that every standard analysis of the concept *knows* is incorrect.[3] Cassam suggests that the Distinct Concepts Argument is independent of the inductive argument also put forward by Williamson in support of the same conclusion; and he rejects the former argument, mainly on the basis of a number of objections to some of its premises. I mostly agree with those objections (cf. section 5 below). However, I do not think they are as strong as Cassam seems to assume, precisely because the Distinct Concepts Argument is not a completely independent argument.

Williamson clarifies the dialectics in his reply to Cassam (cf. Williamson 2009a: 291–2). The claim that *knows* is a mental concept – which, in light of Cassam's criticisms, might seem to have some stipulative flavour – is not supported in isolation with independence of the inductive evidence for the conclusion of the Distinct Concepts Argument. On the contrary:

> The hypothesis that the concept of knowledge is … mental predicts that it is distinct from the complex concept of the analysans in a traditional analysis; experience [the history of failures of traditional analysis] inductively confirms this

---

[3]The premises involved and Cassam's subsequent discussion suggest that this conclusion should be read in the sense of Unanalysability of the Concept. However, my purpose in this section is not to interpret and examine the Distinct Concepts Argument, but rather to focus on what kind of non-inductive evidence Williamson provides for his negative theses. Sosa (2015: 15–16) reconstructs a Williamsonian argument that is very similar to the Distinct Concepts Argument, but whose conclusion is closer to Unanalysability of the State.

prediction, and thereby supports the hypothesis that the concept of knowledge is … mental by inference to the best explanation. (Williamson 2009*a*: 292)[4]

In fact, it is not just this hypothesis (that the concept of knowledge is mental) that would be abductively supported by such a history of previous failures. Some other parts of WTM are in the same position. So, in general terms, the overall web of relations of evidential support should be understood as follows. The (alleged) history of failures of traditional analysis inductively supports negative hypotheses such as Unanalysability of the State and Unanalysability of the Concept. These hypotheses, in their turn, abductively support WTM, since WTM is a general theory that entails Unanalysability of the Concept and that would make it reasonable to accept Unanalysability of the State. WTM is advanced, then, as the best explanation of such negative hypotheses.

Nevertheless, this scheme also entails that the support for Unanalysability of the State and Unanalysability of the Concept is not merely inductive (Cassam is right about this). WTM results from some generalizations and provides a certain systematicity in a range of phenomena that it would account for, including – presumably – phenomena other than the history of failures of the traditional analysis of knowledge. In that sense, if Williamson is right, WTM would confer additional support on Unanalysability of the State and Unanalysability of the Concept, additional to the merely inductive support coming from the "experience" (the "experience" consisting of the alleged history of failures). Indeed, that is so even if eventually the evidence for WTM comes from that same "experience" and some other related phenomena that it would explain.

Now, my main contribution in this article is to argue for this conditional claim: if WTM supports Unanalysability of the State, it does so at an excessive cost. WTM only supports this negative hypothesis insofar as it also supports some other hypotheses that seem highly implausible. Therefore, the preferable available options are to reject WTM, or to reject the claim that WTM supports Unanalysability of the State. A corollary of this result is that the only adequate evidence for Unanalysability of the State provided by Williamson is the merely inductive evidence (which is controversial, but which – for the sake of argument – I am granting here). When we enter into the details of the discussion, another corollary will emerge: Williamson's definition of knowledge seems to fail.

4. Williamson's theory of mental concepts and states

I have anticipated some theses that belong to WTM, regarding the distinction between the concept of knowledge and the state of knowing (section 2 above).[5] The concept of knowledge is a concept *of* the state of knowing. And, as any other state, the state of knowing can be the state of different concepts (some of them not identical to the concept of knowledge).

Besides this, WTM includes a number of particular claims as well as general theses. In the former group, we find claims such as these:

---

[4]On the proper justification for the prediction relation mentioned by Williamson ("that the concept of knowledge is … mental predicts that it is distinct from the complex concept of the analysans in a traditional analysis") cf. section 4 below.

[5]Let us recall that this way of speaking is a way of referring, for any proposition P, to the distinction between the concept of knowing that P (the concept *knows that P*, the concept expressed by 'knows that P') and the state of knowing that P.

(1) The concept *knows* is mental.
(2) The state of knowing is mental.
(3) The concept *is true* is not mental (at least, not when the true proposition is an ordinary contingent proposition about the external environment; henceforth, I will drop the mention to this restriction, which should be understood).
(4) The state, or the condition, of being true is not mental.
(5) The concept *believes truly* is not mental.
(6) The state of believing truly is not mental.

When we ask ourselves how claims about the mental character of some concepts relate to claims about the mental character of some states, WTM offers us an explicit answer:

(7) A state is mental if and only if there could be a mental concept of that state (cf. Williamson 2000: 28).

Thesis (7) suggests that it is our acceptance that the concept *knows* is mental that allows us to accept that the corresponding state is also mental, not the other way around. It is true that Williamson defends (2) not on the basis of (1) and (7), but by pointing out – among other considerations – that the state of knowing would be a sort of paradigm of a mental state; for instance, he says "Our initial presumption should be that knowing is a mental state" (Williamson 2000: 22). Nevertheless, if an objector who accepts (7) but rejects (2) challenged Williamson to argue for (2), her challenge would amount to a request for the identification of a mental concept that was a concept of (corresponding to) the state of knowing. Williamson would offer – I imagine – the concept *knows* as the obvious example. So, I think that Williamson would extend to the concept *knows* his considerations about our pre-theoretical presumptions and its status as a paradigmatic mental concept. In any case, in the framework of WTM, (1) and (7) entails (2); and, of course, (1) is not entailed by (2) and (7).

In addition to (7), these other general theses also belong to WTM:

(8) If concept C is a conjunction of concepts $C_1,\ldots$, and C_n, and one of these concepts, C_i, is not mental, then C is not mental. This is so even if C_i is a redundant factor (so that at least some other conjunct, C_j, entails C_i).
(9) States S_1 and S_2 are identical if and only if necessarily: (everything is in S_1 if and only if it is in S_2).
(10) The concepts *is C* and *is C and such that T* are different concepts, even if they are logically equivalent (for instance, when T is a logical truth).

Thesis (8) plays a crucial role in the derivation of several important claims entailed or supported by WTM. Thus, thesis (5), the concept *believes truly* is not mental, follows from (3), the concept *is true* is not mental, and (8), under the further assumption that *believes truly that P* is the same concept as the conjunctive concept *(believes that P) and (P is true)*. (I have no objection to this last assumption.)

Now, we arrive at Unanalysability of the Concept, which follows from (1), (3) and (8): since *knows* is a mental concept, by (1), it cannot be identical, by (8), to a conjunctive concept one of whose conjuncts is the concept of truth, which is not mental, by (3).[6]

---

[6]Recall Williamson's words quoted in section 3 above: "The hypothesis that the concept of knowledge is … mental predicts that it is distinct from the complex concept of the analysans in a traditional analysis". In such a prediction, (3) and (8) are being used as implicit premises.

What about Unanalysability of the State? It does not follow strictly from WTM. However, according to Williamson, once we accept Unanalysability of the Concept we should also accept Unanalysability of the State. He expresses the idea in this fragment:

> If a non-mental concept were necessarily coextensive with the mental concept *knows*, they would be concepts of the same mental state. The present account does not strictly entail that no analysis of the traditional kind provides correct necessary and sufficient conditions for knowledge [it does not strictly entail Unanalysability of the State]. But once we accept that the concept *knows* is not a complex concept of the kind traditionally envisaged [once we accept Unanalysability of the Concept], what reason have we to expect any such complex concept even to provide necessary and sufficient conditions for knowledge? (Williamson 2000: 30)

In the same vein, Williamson declares: "If the concept *believes truly* is non-mental, its imagined necessary coextensiveness with a mental concept would be a bizarre metaphysical coincidence" (Williamson 2000: 29). As I take it, when he describes such a possibility as a bizarre metaphysical coincidence, Williamson is highlighting the same general point: the theory, WTM, does not entail the impossibility of that necessary coextensiveness, but it would make much reasonable to reject it. In that sense, WTM would support both claims: given that *knows* is a mental concept, there is no other non-mental conjunctive concept necessarily coextensive with it; and given that *believes truly* is a non-mental concept, there is no other mental concept necessarily coextensive with it. The first of these two claims entails precisely Unanalysability of the State. In the framework of WTM, the second claim entails – in virtue of (7) – Williamson's thesis (6): the state of believing truly is non-mental.

To sum up, WTM entails Unanalysability of the Concept, and makes it reasonable – for Williamson, at least – to accept Unanalysability of the State, as well as to accept (and for the same kind of reason: to discard what would otherwise appear to be a bizarre metaphysical coincidence) that the state of believing truly is non-mental.[7]

5. Concepts and conceptual necessity

Now, it is time for replies. I present in this section a few critical comments. My main criticism comes in sections 6, 7 and 8.

First, theses like (8) and (10) seem to have a certain stipulative flavour, and Williamson does little to dispel this impression. In particular, thesis (8) does not appear as clearly more natural or pre-theoretically intuitive than the following thesis (8*), which one might propose:

> (8*) If it is metaphysically necessary that (if the concept C has an instance, then the non-mental concept D has an instance), then C is not mental.

Of course, Williamson must reject (8*) since it would entail – assuming that 'knows' is factive and the concept *is true* is non-mental – that the concept *knows* is non-mental.

Second, sometimes Williamson attacks opponents who certainly do not represent most contemporary attempts to define knowledge. For instance, occasionally he leads the discussion as if the opponent to Unanalysability of the State should be committed

---

[7]I have included thesis (6), the state of believing truly is non-mental, as part of WTM; although it is not particularly important whether we take it as part of WTM or merely as a thesis supported by WTM.

to rejecting Unreductibility, or to rejecting Unanalysability of the Concept. This is not so (Cassam 2009: 17 makes a similar point).

In fact, even if we look for a definition of knowledge in which the modality of the link between definiens and definiendum is conceptual, that does not entail that definiens and definiendum are the same concept, as demanded by the negation of Unanalysability of the Concept. We can accept that a given analysis 'G if and only if R' is conceptually true, while being agnostic as to whether the concept G is or is not identical to the concept R (cf. also Cassam 2009: 18–19). Therefore, we should not assume that Unanalysability of the Concept is equivalent to this other claim:

> Unanalysability* of the Concept: There will be no conceptually necessary and non-circular definition of knowledge of the form "S knows that P if and only if C_1 and C_2 and … C_n", where one of the conjuncts, C_i, is *P is true*.

The plausibility of Unanalysability of the Concept will depend – among other things – on specificities and details of our preferred theory of concepts; Williamson commits himself to some of those specificities about identity conditions for concepts when he includes in his WTM theses like (8) and (10). The issue is obviously related to, but independent of, our preferred view of conceptual necessity, which is the subject matter of Unanalysability* of the Concept. In general, I think that a significant portion of contemporary epistemologists searching for an interesting definition of knowledge typically commit themselves to the denial of Unanalysability of the State.[8] Some of them would also reject Unanalysability* of the Concept. However, Unanalysability of the Concept is a different matter; it is of comparatively limited interest to epistemologists (philosophers of language and philosophers of mind may be more interested in it). Finally, I think that Williamson's criticism to Unreductibility has a modest scope, since the search for a reductive definition of knowledge (in the sense specified in section 2) is a relatively old-fashioned project (and it was already so in 2000, when Williamson published his book). Putting it in other words: how many influential epistemologists make (now, or in 2000) explicit claims committing themselves with definitions of knowing where all the concepts in the definiens would supposedly be more basic than the concept *knows that P*?

6. A metaphysical anthropological hypothesis

Let us recall: for Williamson, *knows* is a mental concept and *believes truly* is a non-mental concept. Why is this so? The basic idea, underlying this discriminatory treatment, seems to be this: the concept *believes truly* is a kind of hybrid composite concept, which stipulatively combines a mental concept (*believes*) and a non-mental concept (*true*). The state corresponding to the concept (the state of believing truly) is "reached" or referred to by the concept in just such a composite way.[9] Thesis (8) in WTM states the (alleged) intuition that composite concepts of this sort cannot be mental. Unlike *believes truly*, the concept *knows* would be – for Williamson – simple or primitive; it cannot be factored into different independent components, one of which is non-mental

---

[8]Very few contemporary characterizations of knowledge have the superficial appearance of the definitions forbidden by Unanalysability of the State. However, whether or not they really instantiate that form may depend on subtle issues concerning logical form and semantic analysability. Cf., for instance, Sosa (2009: 206; 2015: 15–16) and Williamson (2000: Ch. 1; 2009c: 364). I will not enter here into this issue; cf. a few comments on semantic analysability in sections 7 and 8 below.

[9]Williamson (2009a: note 1) also uses a notion of concept according to which concepts are types of linguistic or mental representations. In that sense, we may say that a concept refers to its corresponding state.

(in section 1 above, I already advanced this idea). The concept *knows* refers to, or "reaches", its corresponding state in a simple, non-hybrid way, we may say.

Note that it would be complicated to argue for the non-mental character of the concept *believes truly* – thesis (5) in WTM – on the basis of the non-mental character of the state of believing truly, i.e. on the basis of Williamson's thesis (6) (pace Reed 2005: 138). Granting (7), it would be an open possibility that even if the concept *believes truly* was non-mental, some other concept of the same state could be mental. Williamson rejects this possibility as a bizarre metaphysical coincidence (cf. the previous section). However, his rejection assumes that the concept *believes truly* is non-mental. Its being non-mental, therefore, would be accounted for – in Williamson's view – by the fact that I have emphasized: it is a hybrid, non-simple combination of a mental concept and a non-mental concept.

I am going to present a specific metaphysical-anthropological hypothesis. We have no evidence for its truth. However, its mere possibility poses a serious problem for Williamson's view. The whole pack of Williamson's theses that I have laid out lead us to think that if the hypothesis is possible, then the state of believing truly is mental.

This is the hypothesis:

> (MAH) There are communities of rational thinkers who speak a language that have a linguistic expression, X, with the following characteristic: X expresses a concept that is a concept of (it refers to) the state of believing truly. But, this concept refers to the state in a simple, non-hybrid way; it refers to the state in the same primitive way as the concept *knows* refers to its corresponding state.[10]

Williamson's thesis (5) indicates that the concept *believes truly* is non-mental. MAH says nothing about such a concept. It concerns a different concept, which is a concept of (that corresponds to or refers to) the same state. Unlike *believes truly*, the concept *X* is "simple-primitive" in the same sense in which – for Williamson – *knows* is "simple-primitive". The concept *X* refers to, or "reaches", its corresponding state in a simple, primitive, non-hybrid way.

Now, MAH seems clearly to be a possible hypothesis (more on this below and in the next section). But, Williamson's theorizing about concepts and states entails that if MAH is possible, then the concepts postulated by it really exist – not just as a mere possibility. In particular, according to Williamson the state of believing truly exists; and, if MAH is possible, then there could be a concept *X*, with the traits indicated in MAH, of that state. Meanwhile, if we grant that *knows* is a mental concept, we have no good reasons to deny that *X* is also a mental concept. An important consequence is the following: by (7), the state of believing truly is a mental state, because there could be a mental concept of it, e.g., the concept *X*.[11]

---

[10]Strictly speaking, 'X' is a variable that ranges over linguistic expression. Therefore, there is no implication that MAH concerns expressions composed of a single letter, or even that the language in question (spoken by the communities mentioned by MAH) uses Latin script. However, I think that the discussion may be more fluid if we ignore these details and, allowing some laxity, use 'X' as if it were the expression itself (that is, one of the expressions MAH speaks of). So, not in the formulation of MAH itself, but henceforth, I will speak of "the concept expressed by 'X'" or "the concept *X*" (analogously to speaking of "the concept expressed by 'knows'" or "the concept *knows*").

[11]The conjunction of two theses of WTM, theses (6) and (7), entails that the concept *X* is not mental. However, this fact cannot be used against the thesis that the concept *X* is mental; because the fact depends on thesis (6), which in the context of this discussion – as I have just said – would be questioned. I am indebted to an anonymous referee of *Episteme* for bringing up this point.

Note that, in the present discussion, it would be question-begging to say something along the lines of the following reasoning:

(**) X is a non-mental concept; it cannot have the simplicity-primitiveness of *knows*, because it necessarily admits an analysis that factors it into different independent components, of which one is non-mental; this is the obvious analysis: S Xs that P if and only if (S believes that P) and (P is true).

Why would it be question-begging? At present, we are looking for (not merely inductive) evidence supporting Unanalysability of the State. More specifically, we are asking ourselves how WTM and in particular the thesis that *knows* is mental – thesis (1) in WTM – would support Unanalysability of the State. Therefore, we cannot assume Unanalysability of the State. In other words, it is dialectically illegitimate to assume that the concept *knows* does not admit of an analysis with the form: S knows that P if and only if (S believes that P) and (P is true) and Q. But (**) only works as a justification for the alleged non-mental character of the concept X if it is assumed that *knows*, unlike X, does not admit of such an analysis.

We can expand this point. My main claim in this work is that WTM only supports Unanalysability of the State if it also supports ruling out as impossible certain metaphysical hypotheses – MAH is one such example – that clearly seem to be possible. There are two ways for this claim to be true: (a) WTM does not support Unanalysability of the State; or (b) WTM supports ruling out as impossible hypotheses such as MAH which seem to be possible. Now, MAH involves a certain notion of simplicity-primitiveness and establishes that there could be a concept X (which refers to the state of believing truly) that is simple-primitive, as *knows* is. I have not said much about what exactly this notion of simplicity-primitiveness amounts to, precisely because the burden of clarifying it would rest with Williamson, if doubts about its intelligibility were raised. I am using 'simple-primitive' to refer to what Williamson must say to justify the claim that *knows* is mental and *believing truly* is non-mental. This simplicity-primitiveness is somehow related to the idea that the concept "does not admit of an analysis that factors it into different independent components, one of which is non-mental". Whatever the exact connection between them, I am invoking the notion of simplicity-primitiveness under this assumption: saying that the concept *knows* refers to its state in a simple-primitive way does not entail by definition that *knows* does not admit of a traditional-conjunctive analysis. In other words, I am assuming that the simplicity-primitiveness of *knows* does not entail by definition Unanalysability of the State (of Knowing). Analogously, I am assuming – against (**) – that the simplicity-primitiveness of X does not entail by definition Unanalysability of the State of X, which is a false thesis, as (**) correctly says.[12] If the assumption is right, then (**) is wrong. X can have the simplicity-primitiveness of *knows*, and consequently it should be classified as mental; (b) holds. What happens if I am wrong regarding that assumption? If the alleged simplicity-primitiveness of X entails by definition Unanalysability of the State of X, then MAH is impossible; (b) does not hold. But then, (a) holds. The fact that *knows* is a mental concept – thesis (1) in WTM – entails by definition the very thesis Unanalysability of the State (of Knowing). There is no interesting support for Unanalysability of the State coming from WTM, and particularly none coming from thesis (1).

---

[12]By "Unanalysability of the State of X" we should understand this thesis: there will be no necessarily true and non-circular definition of the state of X having the form "S Xs that P if and only if C_1 and C_2 and … C_n", where one of the conjuncts, C_i, is *P is true*.

Let me make some further consequences of accepting that X is a mental concept explicit (besides the fact that believing truly would be a mental state). Granting that *believes truly* is a non-mental concept, its necessary coextensiveness with a mental concept is not – pace Williamson – a bizarre metaphysical coincidence; in other words, it is not unreasonable to think that there is a mental concept that is necessarily coextensive with *believes truly*. On the contrary: it would be a bizarre metaphysical coincidence if there were no such mental concept; because that non-existence would require that MAH were impossible, and that would indeed be a metaphysical impossibility that was extremely hard to explain.

Analogously, the necessary coextensiveness of *knows* with a non-mental concept would not be a bizarre metaphysical coincidence; i.e., it would not be unreasonable to think that there was a non-mental concept that was necessarily coextensive with *knows*. Therefore, pace Williamson, it is not unreasonable to reject Unanalysability of the State. If WTM supports Unanalysability of the State, it does so at an excessive cost: it also supports MAH being impossible.

7. Discussion of possible responses and counter-responses

In this section, I will briefly comment on some options Williamson has to avoid the problem I have detailed. If MAH overtly established that X was a mental concept, then Williamson would have no choice but to reject the claim that MAH is possible. The rejection could have two different bases; I have preferred to articulate MAH in its present form, so that we can discern those two bases as two different options open to Williamson. He may opt for one of these claims: (I) MAH is impossible; or alternatively (II) the concept X is non-mental. Let us focus on (I) and then (II) separately.

What reasons might someone have to claim that MAH is impossible? Think of the state of knowing. The history of human evolution has given rise to different natural languages. Those natural languages typically have expressions, such as 'knows' and its derivatives, which refer to the state of knowing. We confidently assume that different expressions, belonging to different languages, express the same concept of knowledge (i.e., the concept *knows*, the concept expressed by 'knows', 'sabe', 'sai', etc.). For Williamson, some complex expressions do not express the concept *knows*, but refer to the same state (cf. the next section). Of course, knowing entails believing and knowing entails truth. Nevertheless, the alleged fact that – at least in Williamson's view – *knows* is a mental concept is accounted for (as far as it is contrasted with the alleged fact that some other concepts, such as *believes truly*, are non-mental) by its simple-primitive character, by its referring to its corresponding state in a non-hybrid way; a way that does not involve a combination of the concepts *believes* and *true*.

Now, think of the state of believing truly, and consider the possibility of a different history of human evolution. Suppose that there were at least one natural language that had a given expression, 'X', with the following traits. 'X' refers to the state of believing truly. Of course, the state of X entails believing and it also entails truth. Nevertheless, the concept X – the concept expressed by 'X' – has a simple-primitive character; it refers to its corresponding state in a non-hybrid way; a way that does not involve a combination of the concepts *believes* and *true*. What exactly is unacceptable in this story?[13] What kind of physical, biological or psychological trait in our animal nature poses a

---

[13]Recall the discussion in the previous section. To say that "any concept X that refers to the state of believing truly will refer to it in a way that *does* involve a combination of the concepts *believes* and *true*" would be a reformulation of (**). I repeat the core of my response: if we read the notion of *simple-primitive* so that it entails (**), then MAH is impossible but Unanalysability of the State (of Knowing)

metaphysical obstacle to its mere possibility? Or is it rather some other logical or meta-physical trait, related to the state of being true, that causes the problem? At first sight, there is no such obstacle. If, unknown to all of us, there really is some empirical, logic-conceptual or metaphysical impediment to the possible different evolution of our lan-guages and cognitive system, then the philosopher who denies the possibility of MAH should describe it. She should explain to us why it is impossible for the relation that holds between the state of knowing and the concept *knows* to hold between the state of believing truly and a certain concept, X.

In fact, her task is even harder. The tag 'anthropological' in the denomination of MAH is not entirely accurate. MAH concerns possible communities of rational thin-kers, so that it is not restricted to our species. MAH may be true in virtue of possible evolutionary histories concerning other biological species, not even necessarily from our planet. Taking this reflection into account, it is even more unreasonable to think that MAH is metaphysically impossible.

Consider this empirical restrictive hypothesis: if a natural language has a simple-primitive expression that expresses the concept *believes* and has a simple-primitive expression that expresses the concept *true*, then it is highly improbable that it also has a simple-primitive expression that expresses the concept *believes truly*. It might be true. However, it is of little help to Williamson. First, he would need the impossibility of MAH, not merely its low probability in the logical space. Second, MAH does not entail that users of 'X' also use linguistic simple-primitive terms expressing the concept *believes* and the concept *true*.[14]

The second option, (II), is to argue for the non-mental character of the concept X. However, when we scrutinize the – direct or indirect – evidence provided for the thesis that *knows* is a mental concept, it is hard to find a crucial trait that would not also apply to the concept X, once we grant that the concept X has the characteristics described by MAH (i.e., once we grant that MAH is possible). Some of the relevant issues here have already been discussed in the previous section: (7) makes it difficult to argue for the non-mental character of the concept X on the basis of the alleged non-mental character of the state of believing truly; and if doubts about the intelligibility of the notion of simplicity-primitiveness were raised, it must be chiefly Williamson's concern to dispel them.

An objector has suggested to me that the considerations in section 2.4 of *Knowledge and its Limits* might be used in support of option (II) (J. Comesaña made this point, in the *VIIIth Meeting of the SEFA*; M. García-Carpintero joined the objection). This sup-port would come from thesis (11), that Williamson defends in that section:

---

cannot receive any interesting support from WTM. In this section, I am working under the alternative assumption that *simple-primitive* does not entail (**).

[14]The considerations in the three last paragraphs show that my strategy has not been just to say something like "let us introduce a semantically unanalysable term, 'X', to refer to the state of believing truly". Such a strategy would dubiously be legitimate. In contrast, I have made it reasonable to accept this conditional: if 'knows' is simple-primitive, then it is not impossible the existence of rational thinkers who use a term 'X', referring to the state of believing truly, which is also simple-primitive. And – more importantly – MAH does not say that the linguistic community using 'X' has stipulated that it refers to the state of believing truly. So, the line I take differs from that followed by Reed (2005). He makes reference to a highly specific kind of state: looking at a barn in good visual conditions, in an area with many fake barns, without being aware that the fake barns are in fact fake (the original example – provided by C. Ginet – appeared in Goldman 1976). Then he writes: "Let us introduce a semantically unanalyzable term for cases of this sort: …, in the above case, [the subject] *gnows* […] that there is a barn in front of him" (Reed 2005: 138). Williamson would consider 'gnow' to be a semantically analysable term; cf. Williamson (2000: 36; 2009*b*: 364).

(11) Some rational actions are better explained by invoking the concept *knows* than by invoking the concepts *believes* or *believes truly*.

Williamson focuses – among others – on an imaginary case, then he identifies a determinate explanandum concerning the actions of an agent and he convincingly argues that the substitution of 'believe' or 'believe truly' for 'know' in our explanation of it involves explanatory loss.

However, the attempt to support option (II) on the basis of (11) has at least three drawbacks. First, the objector would need something stronger than (11), on the lines of (12):

(12) In general terms, rational actions are better explained by invoking the concept *knows* than by invoking the concepts *believes* or *believes truly*.

The idea – as far as I can see – is, then, that the mental character of concepts is closely connected with their role in intentional explanations of actions so that (12) would warrant a further claim:

(13) The mental character of the concept *knows* is more firmly established than the mental character of the concepts *believes* and *believes truly*.

The second drawback is obvious. Claims (12) and (13) do not concern the concept *X*, but a concept which is coreferential to it: the concept *believes truly*. Williamson compares the explanatory role played by *knows* to the explanatory role played by *believes truly*. But, of course, he does not compare it to the explanatory role played by *X*. I do not think that we have clear intuitions about this other contrast. Let us suppose that some of the thinkers mentioned in the hypothesis (MAH) also have a linguistic term, 'K', that expresses the concept *knows*. Sometimes they explain rational action by using 'X'. Can we assume that their explanations are better if they use 'K' instead of 'X'? Even if we accept (12) and (13), there is no sufficient evidence for (12*) and (13*):

(12*) In general terms, rational actions are better explained by invoking the concept *knows* than by invoking the concept *X*.
(13*) The mental character of the concept *knows* is more firmly established than the mental character of the concept *X*.

The final – and most important – problem with this objection is clear too. Option (II) requires something significantly stronger than (13*). It requires that the concept *X* is not mental at all. We can think of an idea that would be useful to the objector: with regards to the mental character of concepts, there is a line of demarcation determined (or partially determined) by how fruitful they are when used in explaining rational action, and it turns out that this line separates concepts like *knows* and *believes* as mental from concepts like *believes truly* and *X* as non-mental. However, the hypothesis that such a line of demarcation is placed precisely there seems a little arbitrary.

8. Some corollaries about general factive mental states

In Chapter 2 of his book, Williamson presents and defends an alternative definition of knowledge. Knowing, according to that definition, would be the most general factive mental state; i.e., the factive mental state metaphysically implied by all factive mental states.

A propositional attitude is factive if and only if, necessarily, one has it only to truths. … Not all factive attitudes constitute states; forgetting is a process. Call

those attitudes which do constitute states *stative*. The proposal is that knowing is the most general factive stative attitude, that which one has to a proposition if one has any factive stative attitude at all. (Williamson 2000: 34)

Believing truly is a factive state, and it is more general than knowing. Williamson does not see it as a counterexample to his definition, since he thinks that believing truly is non-mental. However, in the light of our previous results, the concept X would be a mental concept. So, by (7), believing truly would be a mental state. In that sense, apparently Williamson's proposal would fail.

Things are not so easy, however, since Williamson is somewhat unclear as to exactly what the status of his characterization of knowledge is. Sosa (2009) claims that the state of safely believing is a counter-example to Williamson's definition. In response, Williamson writes:

The full account in the book is formulated in terms of factive mental state operators (FMSOs) ['know' and 'remember' would be typical FMSOs, in Williamson's view] … by definition FMSOs are semantically unanalysable. … Sosa in effect stipulates that 'safely believes p' is to be synonymous with 'believes p and would believe p only if p were true'. But then 'safely believes p' is semantically analysable in just that way, and is therefore not an FMSO. Correspondingly, safe belief is not a factive mental state, and the counter-example fails. (Williamson 2009c: 364)[15]

There are at least two problems with this response to Sosa (besides the problem anticipated by Sosa himself). First, the counterexample concerns the state, not a concept. Even if 'safely believes p', like 'believes truly', is semantically analysable, there is the possibility, according to (7), that some other concept that refers to the same state is semantically unanalysable. Second, suppose we accept that by (stipulative) definition, FMSOs are semantically unanalysable.[16] Then, we cannot unproblematically assume that 'knows' is an FMSO. Otherwise, we would obtain the apparently substantive thesis Unanalysability* of the Concept, or something very close to it, from a stipulative definition and an unproblematic assumption.

Regardless of these problems, I take it that Williamson's exact definition of knowing involves the notion of factive mental state operators, and it would take approximately the following form:

(W) Knowing is the most general mental state referred to by an FMSO.[17, 18]

---

[15]Sosa (2009: note 3) anticipates this kind of reply, and suggests that – although syntactically complex – 'safely believes that p' is no more semantically analysable than 'knows by the sense of touch that p', which in Williamson's eyes is semantically unanalysable (cf. Williamson 2000: 36–7; 2009c: note 1). For Williamson "semantically unanalysable" means "not synonymous with any complex expression whose meaning is composed of the meaning of its parts" (Williamson 2000: 34).

[16]Williamson (2000: 36, 39, 40) comes very close to saying it; but he is not as explicit about this as in Williamson (2009a: 285; 2009c: 364). Reed (2005: 137) also assumes that the concept of FMSO does not definitionally entail semantic unanalysability.

[17](W) defines the state of knowing. Obviously, the FMSO Williamson has in mind is 'knows'. This also allows him to define the concept *knows* (i.e., the concept expressed by 'know'). His first informal definition of the state (still without mentioning FMSOs; quoted at the beginning of this section) appears on p. 34. When he goes into the details, what he presents on p. 39 ("the full account in the book" mentioned in his reply to Sosa) is a definition of the concept *knows* (more exactly, a characterization of it up to logical equivalence). Granting some of his relatively uncontroversial claims about the relation between concepts and states, this second definition entails (W).

[18](W) is not a conjunctive definition (cf. section 2 above). Furthermore, for Williamson, definiens and definiendum in (W) are not the same concept. Now, we have the following results. Suppose that concept G is necessarily coextensive with the concept *knows* and concept R is necessarily coextensive with the concept

(W) implies at least the following two substantive hypotheses: the state of knowing is mental; and it is the most general factive mental state. If by definition FMSOs are semantically unanalysable, then (W) also implies this other substantive hypothesis: 'knows' is semantically unanalysable. Of course, Williamson is perfectly clear in defending this last hypothesis. The lack of clarity I refer to concerns whether or not he wanted that hypothesis to be implied by his characterization of knowledge. Taking into account his later comments – Williamson (2000*a*: 285; 2000*c*: 364) – the answer is positive.

Let me now examine more carefully whether or not the state of believing truly is a counterexample to (W). If semantic unanalysability is not built into the definition of 'FMSO' as a necessary condition, then there is no reason to reject the claim that 'X' is an FMSO. From the discussion in previous sections, I take it that 'X' expresses a mental concept (if 'knows' expresses a mental concept). Therefore, (W) would fail, because believing truly is a counterexample.

However, let us accept that the notion of FMSO definitionally entails semantic unanalysability. Then, there are two issues. First, it is controversial that 'knows' is indeed an FMSO; if it is not, then (W) fails. Even if 'knows' is an FMSO, it is an open question whether or not 'X' is semantically unanalysable. Williamson says: "An artificial verb stipulated to mean the same as 'believe truly' would not be an FMSO" (Williamson 2000: 36). However, 'X' is not an artificial verb stipulated to mean the same as 'believe truly' (cf. our section 7, in special note 14). Certainly, 'X' admits of a traditional-conjunctive analysis; Unanalysability of the State of X is a false hypothesis. But we cannot infer from this that the analysis in question is purely semantic; we cannot assume that 'X' and 'believes truly' are synonyms. If 'X' is semantically unanalysable, again there is no reason to reject the claim that 'X' is an FMSO. So, (W) would fail. If 'X' is semantically analysable, then we have not found a counterexample to (W).[19, 20]

---

believes truly. Then, according to Williamson: if *G* is a conjunctive concept, its coextensiveness with the concept *knows* is a bizarre metaphysical coincidence; if *G* is not a conjunctive concept, its coextensiveness with the concept *knows* is not a bizarre metaphysical coincidence; and if *R* is a mental concept, its coextensiveness with the concept *believes truly* is a bizarre metaphysical coincidence (cf. our section 4). Again, all this sounds somewhat ad hoc. These differential treatments would seemingly require some theoretical justification.

[19]An anonymous referee of *Episteme* points out that Williamson might raise this objection: it would be a bizarre coincidence that the community using 'X' (having devoted many resources to investigate about X) did not discover that concept X is coextensive with the concept of true belief. The idea – if I understand it correctly – would be that this discovery (and the conditions that led to it) would be evidence for the thesis that 'X' is semantically analyzable. I grant that some possible communities using 'X' would discover this relation of coextensiveness. And I also grant this other statement: if Williamson is willing to accept that although 'X' is not stipulatively equivalent to 'believes truly' it is semantically analyzable, then he has a way of holding on to (W). However, this fact does not substantially change our previous considerations against the existence of non-inductive evidence for Unanalysability of the State. Supposedly, the semantic analysability of 'X' would be evidence for the identity between the concepts X and *true belief*. This identity, combined with thesis (8), would entail that concept X is non-mental. But this chain of reasoning depends (among other things) on Williamson's thesis (8), which I have already said has a stipulative flavour (cf. our section 5).

[20]Under the assumption that my argumentation was on the right track, it has been suggested to me that some non-orthodox Williamsonians might hold on to Unanalysability of the Concept (of *knowledge*) if they cling to this burning nail: they could say that concept X is the concept of *knowledge*. After all, 'X' would be the most general factive mental state operator, and so it would fit Williamson's definition of knowledge. I do not know if that position would preserve Unanalysability of the Concept. In any case, I think that keeping Unanalysability of the Concept at that cost would be straying too far from Williamson.

9. Conclusions

My main subject matter in this article has been the prospect of defining knowledge; in particular, what kind of evidence do we have for Williamson's negative thesis Unanalysability of the State. I have granted, for the sake of the argument, a controversial thesis: all traditional-conjunctive definitions proposed so far have counter-examples. That provides inductive evidence in favour of Unanalysability of the State.

For Williamson, the history of previous failures to define knowledge would abductively support his general theory WTM; in that sense, WTM would provide additional, not merely inductive evidence, for Unanalysability of the State. I have defended the following crucial claim: within the general framework of WTM, this theory only supports Unanalysability of the State if it also supports ruling out as impossible certain metaphysical hypotheses – such as MAH – that clearly seem to be possible. MAH concerns a concept, X, which refers to the state of believing truly in the same simple-primitive way as the concept *knows* refers to the state of knowing. If *knows* is mental, X is mental. WTM entails that if X is mental, then the state of believing truly is mental.

As a consequence, (W), Williamson's own definition of knowing – as the most general factive mental state – is in danger: if 'X' is semantically unanalysable, then it is a FMSO and the state of believing truly is a counter-example to (W).

I will finally now repeat the main lesson of this discussion from a more practical point of view. Suppose you are an epistemologist, or at least you are interested in epistemology. You focus on a specific definition of knowing, (D). (D) may be a definition of your own, or it may have come from someone else. You have carefully reflected on (D) and think that all the alleged counterexamples to it can be properly answered. As far as you know, (D) is a correct non-circular analysis, which provides necessary and sufficient conditions for knowledge. It might be a controversial question whether or not (D) is a traditional-conjunctive definition (its apparent form does not immediately decide the point). But suppose also that you think or you grant that (D) is traditional conjunctive. Now, what kind of worries should you have in the light of Williamson's criticisms in *Knowledge and its Limits*? None at all. Williamson proposes a general theory, WTM, which would explain the previous failures of other traditional-conjunctive definitions and would predict that (D) also fails. But WTM would explain too much; it makes highly implausible predictions about metaphysical possibilities. Meanwhile, the merely inductive evidence against (D) should not worry you too much. There is a better explanation for it than WTM: other epistemologists failed just because philosophy – and epistemology, in particular – is hard.[21]

References

Cassam Q. (2009). 'Can the Concept of Knowledge be Analysed?' In P. Greenough and D. Pritchard (eds), *Williamson on Knowledge*, pp. 12–30. Oxford: Oxford University Press.

Goldman A.I. (1976). 'Discrimination and Perceptual Knowledge.' *Journal of Philosophy* **73**, 771–91.

Pérez Otero M. (2016). 'La teoría de Williamson sobre los conceptos y los estados mentales.' *Proceedings of the Eighth Meeting of the SEFA*, pp. 148–51. Oviedo: KRK Ediciones.

Pérez Otero M. (2018). 'El dominio de lo mental en la filosofía de Williamson.' *Theoria* **33**, 23–40.

Pritchard D. (2012). 'Anti-luck Virtue Epistemology.' *Journal of Philosophy* **109**, 247–79.

Reed B. (2005). 'Accidentally Factive Mental States.' *Philosophy and Phenomenological Research* **71**, 134–42.

Sosa E. (2009). 'Timothy Williamson's *Knowledge and its Limits*.' In P. Greenough and D. Pritchard (eds), *Williamson on Knowledge*, pp. 203–16. Oxford: Oxford University Press.

Sosa E. (2015). *Judgment and Agency*. Oxford: Oxford University Press.

Williamson T. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.

Williamson T. (2009*a*). 'Reply to Quassim Cassam.' In P. Greenough and D. Pritchard (eds), *Williamson on Knowledge*, pp. 285–92. Oxford: Oxford University Press.

Williamson T. (2009*b*). 'Reply to Frank Jackson.' In P. Greenough and D. Pritchard (eds), *Williamson on Knowledge*, pp. 330–2. Oxford: Oxford University Press.

Williamson T. (2009*c*). 'Reply to Ernest Sosa.' In P. Greenough and D. Pritchard (eds), *Williamson on Knowledge*, pp. 364–9. Oxford: Oxford University Press.

Manuel Pérez Otero is Professor at the Department of Philosophy of the University of Barcelona, and member of LOGOS (*Research Group in Analytic Philosophy*) since it was founded. His main research areas are Epistemology, Philosophy of Language, Philosophical Logic and Metaphysics. On these topics, he has published several books as well as articles in journals such as *Dialectica*, *Pacific Philosophical Quarterly*, *Philosophical Studies*, *Philosophy and Phenomenological Research* and *Synthese*. http://www.ub.edu/grc_logos/manuel-perez-otero