

LI Qilin

## Accidentally True Beliefs and the Williamsonian Mental State of Knowing

**Abstract** In this paper, I will explore some philosophical implications of Williamson’s thesis that knowing is a state of mind (KSM). Using the fake barn case, I will introduce a way to evaluate Williamson’s KSM thesis and determine whether the Williamsonian mental state of knowing can be plausibly distinguished from certain other similar but epistemologically distinctive states of mind (i.e., accidentally true beliefs). Then, some tentative externalist accounts of the supposed differences between the Williamsonian mental state of knowing and accidentally true beliefs will be critically assessed, implying that the evaluated traditional versions of externalism in semantics and epistemology do not fit well with Williamson’s KSM thesis. Ultimately, I suggest that the extended-mind or extended-knower approach may be more promising, which indicates that active externalism would be called for by Williamson’s KSM thesis.

**Keywords** knowing, Gettier problem, mental states, externalism, true belief, Williamson, knowing is a state of mind, epistemological

---

### 1 Williamson on Knowing as a State of Mind

In his seminal book *Knowledge and Its Limit*, Timothy Williamson proposes a new approach to the study of knowing or propositional knowledge.<sup>1</sup>

---

<sup>1</sup> Hereafter, without further specification or explanation, “knowing” (or “knowledge”) throughout this paper is restricted to the propositional sense. Some other non-propositional knowledge, such as knowing-how or knowing-by-acquaintance, is set aside.

LI Qilin (✉)

Department of Philosophy, and Institute of Foreign Philosophy, Peking University, Beijing 100871, China

E-mail: liqilin@pku.edu.cn

According to Williamson, “knowing is a state of mind” (Williamson 2000, 21) that “does not factorize as standard analyses require” (Williamson 2000, 33). Bearing this kind of understanding of knowing in mind, Williamson proposes that “we can see epistemology as a branch of the philosophy of mind” (41). Williamson’s epistemological proposal here is to argue against the traditional conceptual analysis of knowledge, of which the justified-true-belief (JTB) account of knowledge is regarded as a paradigm. However, Williamson does concede that “a modest positive account of the concept” of knowledge may be available (33). To fully appreciate his innovative approach to the study of knowing, a few detailed clarifications of Williamson’s account of the thesis that “knowing is a state of mind” would be plausibly called for.

When he suggests that knowing is a mental state, first and foremost, Williamson does not imply that knowing inherits the property of being a mental state from any of its mental components which are constitutive and fundamental to knowledge. In this sense, Williamson no longer commits to the traditional thesis that knowledge is a kind of belief. For Williamson, “knowing is *merely* a state of mind,” which amounts to “the claim that there is a mental state being in which is necessary *and sufficient* for knowing *p*” (21). To put this idea more formally, “For some mental state *S*, being in *S* is necessary and sufficient for knowing *p*” (21). Thus, we can see that Williamson’s claim that knowing is a mental state is actually an abbreviative statement with much philosophical significance, resulting in (at least) two important consequences in Williamsonian epistemology. (i) Being so construed, knowing can be regarded as a paradigmatic mental state which no longer needs to be reduced to any other kinds of mental states (say, beliefs, for example) (cf. 27–33). This would favor Williamson’s knowledge-first epistemology. (ii) Williamson is able to defend an externalist view of knowing, since “a difference in knowledge would *constitute* a difference in mental state” (26). A desirable consequence from this view is that the difference in knowing under discussion “need not be detectable by the subject who lacks knowledge” (26).<sup>2</sup> Williamson’s view of knowing then saves a theoretical space for the possibility of an epistemic subject being in different

---

<sup>2</sup> Williamson himself appeals to this view significantly to argue against epistemic skepticism (see Chapter 8 of his *Knowledge and Its Limits*). Due to the limited length of this paper, I will leave the issue of skepticism aside.

mental states even when he or she is positioned in indiscriminable situations.<sup>3</sup>

Since he rejects the reductionist view of knowing, Williamson only presents qualitative descriptions of the properties of the mental state of knowing. According to Williamson, knowing is the most general factive mental state, and correspondingly, “know” is a paradigmatic factive mental state operator (FMSO). When Williamson claims that knowing is factive, knowing does not take the truth of the target proposition as its subpart or component, even though one may *validly infer* the truth of  $p$  from the mental state of knowing that  $p$ . In this sense, knowing is not reducible to any kind of true belief, for the latter is *not* the factive *mental* state. In the above sense, Williamson’s view of knowing departs significantly from the traditional analysis of knowledge (say, any analysis of knowledge with the form of “JTB + X”), although both views may ostensibly appear to agree on the factivity of knowing.

Williamson suggests that knowing, regarded as a genuine state of mind, implies that “we can see epistemology as a branch of the philosophy of mind” (41); this, in turn, would enable epistemologists to borrow theoretical resources from the philosophy of mind and then to apply them in epistemological studies. The Williamsonian approach to epistemology should still provide us with sufficient space to engage certain crucial problems that are intrinsically epistemological. I would suggest that one of the legitimate epistemological research projects would be how to tell knowing from other non-knowing states of mind. Some non-knowing states of mind (say, accidentally true beliefs) may appear to be similar to knowing, and epistemologists from Ancient Greece onward have remained enthusiastic about developing various theories that are intended to help people to tell knowing from lucky guessing or purely true beliefs generated by known accidents. We should distinguish the above difference-telling project from the reductionistic project. The latter project may presuppose the former one, but not *vice versa*. For instance, a certain version of the metaphysical reductionism of knowing may reduce knowledge to types of belief with

---

<sup>3</sup> As will be shown in the next section of the paper, this view plays a crucial role in the fake-barn Gettier case.

epistemically desirable properties. If there is no discriminative difference between knowledge and belief, the above reductionism is not possible. The epistemologists who intend to reveal the difference between knowing and non-knowing mental states, on the other hand, are not straightforwardly committed to the reductionism of knowing. Even within the Williamsonian “knowledge-first” framework, the difference-telling project may remain valid. Williamson suggests that knowing is a FMSO that is epistemically primitive and foundational (cf. 39). Belief (and even accidentally true belief), on the other hand, is not an FMSO. According to Williamson, knowledge, and belief are *different types* of states of mind. Thus, even for Williamson, there is some significance to discussing the Gettier cases, because the agents in the Gettier case are in certain states of mind which lack certain properties or elements that are essential or intrinsic to knowing. Of course, the Williamsonian discussion of the Gettier cases is no longer committed to any reductionism of knowing.<sup>4</sup> We may grant the relevant Williamsonian methodological strategy in the paper.<sup>5</sup> In the next section, the famous Gettier case, i.e., the fake barn case, will be introduced so that we can evaluate the Williamsonian thesis of knowing. Once again, a caveat has to be announced in advance: by using the relevant Gettier cases here, I am not committing, implicitly or explicitly, to any metaphysical or conceptual reductionism of knowing. In this paper, I suggest that we should understand the Gettier cases in the following way: the state of mind of the Gettiered subject is just accidentally (or luckily) true belief, which has to be distinguished from knowing, which by no means implies the reductionist thesis that knowing is (conceptually or metaphysically) constituted by some beliefs with specific epistemologically desirable properties.

Bearing the above view in mind, we will continue to assess Williamson’s view of knowing as a state of mind with respect to the classic Gettier case of fake barns.

---

<sup>4</sup> As we will see, Williamson indeed talks about the Gettier cases without any commitments to reductionism. In his discussion of the Gettier cases, Williamson suggests that the agents’ states of mind may be regarded as ones that are closely similar to knowing, but that they are not identical to the states of knowing.

<sup>5</sup> Here I would like to thank one of the anonymous referees for the journal, as his or her comment made me realize that detailed clarification is necessary.

---

## 2 The Gettier Case of Fake Barns and Knowing as a Williamsonian Mental State

Consider a pair of cases as follows:

### THE NORMAL CASE OF KNOWING:

Henry<sub>n</sub> is driving in the countryside with his son. For the boy's edification, Henry<sub>n</sub> identifies various objects on the landscape as they come into view. "That's a cow," says Henry<sub>n</sub>, "That's a tractor," "That's a silo," "That's a barn," etc. Henry<sub>n</sub> has no doubt about the identity of these objects; in particular, he has no doubt that the last-mentioned object is a barn, which indeed it is. Each of the identified objects has features characteristic of its type. Moreover, each object is fully in view, Henry<sub>n</sub> has excellent eyesight, and he has enough time to look at them reasonably carefully, since there is little traffic to distract him. And the countryside is normal. (Goldman 1976, 772, with minor adaptations)

### THE GETTIERED FAKE BARN CASE:

The case-setting is almost the same as the first case, except for the following two variations: (i) it is now Henry<sub>g</sub> who is driving in the countryside; and (ii) unknown to Henry<sub>g</sub>, the district he has just entered is full of papier-mâché facsimiles of barns. These facsimiles look from the road exactly like barns, but are really just façades, without back walls or interiors, and quite incapable of being used as barns. They are so cleverly constructed that travelers invariably mistake them for barns. Having just entered the district, Henry<sub>g</sub> has not encountered any facsimiles; the object he sees is a genuine barn. But if the object on that site were a facsimile, Henry<sub>g</sub> would mistake it for a barn. (Goldman 1976, 773, with minor adaptations)

According to the classical interpretation, we are inclined to agree with the following two knowledge-ascription statements:

(k<sub>n</sub>) Henry<sub>n</sub> knows that it is a barn in front of him.

(k<sub>g</sub>) Henry<sub>g</sub> does not know that it is a barn in front of him.

Although some contemporary epistemologists may disagree on (k<sub>g</sub>) and insist

that Henry<sub>g</sub> also has the relevant knowledge,<sup>6</sup> Williamson himself does regard the second case as a genuine Gettier case. Namely, Williamson thinks that Henry<sub>g</sub> in the fake-barn case is not in a mental state of knowing, because “fake barn’ Gettier cases” present certain kinds of cases “in which the agent may lack relevant false beliefs but still the circumstances are not favorable enough for knowledge of the given truth” (Williamson 2013, 12). It is also worth emphasizing that the above claim presented by Williamson is not seriously challenged by the increasing accumulation of data from studies in experimental philosophy,<sup>7</sup> for, as Williamson suggests, the relevant Gettier cases can be formally motivated in epistemic logic. If Williamson is correct, we can conclude from the above pair of cases that Henry<sub>n</sub> and Henry<sub>g</sub> are in different states of mind. Intuitively speaking, Henry<sub>n</sub> is in a mental state of knowing, while Henry<sub>g</sub> is not in a mental state of knowing.<sup>8</sup> But, can such intuitive judgments be backed up or justified (beyond the intuition)? If the answer to the question is “yes,” we may wonder what difference can be identified<sup>9</sup> so that we can tell Henry<sub>n</sub>’s mental state from Henry<sub>g</sub>’s.

It is worth noting that there is an ambiguity (cf. Blackburn 2016, 486) involved in the above question, which might lead to a relatively easy, trivial, and uninformative answer. For instance, if we think the above question aims at the *token* difference between the states of minds of Henry<sub>n</sub> and Henry<sub>g</sub>, the question can be easily answered, for Henry<sub>n</sub> and Henry<sub>g</sub> are different agents who are located in different spatiotemporal positions. But such a reply does not address the real puzzle in the first place and therefore becomes philosophically much less attractive. On the other hand, if we aim at some

---

<sup>6</sup> For instance, some epistemologists who commit to the truth-maker account of knowledge deny that the fake-barn case is a genuine Gettier case; they further claim that the subject in the fake-barn case knows that it is a barn. For the truth-maker account of knowledge, see, for example, Heathcote 2012, 309–14; Jacquette 1996, 115–27. For the argument against the truth-maker theory of knowledge, see Biro, 2013, 57–62; and 2014, 377–81; Vance 2014, 291–305. For a non-traditional but inspiring diagnosis of the fake-barn case without appealing to truth-maker theory, see Hetherington 2012, 217–30; and 2016, 5–9, 89–106, 183–89.

<sup>7</sup> For experimental philosophers’ studies of the Gettier cases and relevant intuition, see, for instance, Alexander and Weinberg 2014, 128–45; Boyd and Nagel 2014, 109–27; Turri 2016, 337–48.

<sup>8</sup> One may suggest, for example, that Henry<sub>g</sub> is only believing (with the belief’s content proposition being accidentally true, of course).

<sup>9</sup> Note that the term “(being) identified” here is to ascribe a performance of ascribers (such as, we who read and evaluate the cases in question), which does not at all imply either Henry<sub>n</sub> or Henry<sub>g</sub> is able to tell whether he is positioned in a *normal* countryside or not.

philosophically significant, informative solution to the problem, it seems that we have to find out a property that is presented in all mental states of knowing (including Henry<sub>n</sub>'s mental state, of course) but absent in the *type* of non-knowing mental states (which is instantiated by Henry<sub>g</sub>). In a word, when we ask what makes the mental states of Henry<sub>n</sub> and Henry<sub>g</sub> different, we are indeed asking, "What difference tells the *type* of Henry<sub>n</sub>'s mental state from the *type* of Henry<sub>g</sub>'s?"<sup>10</sup> Thus, unless further provisos are provided, we would uniformly treat our question as the significant and informative one, which requires us to find the difference between the *types* of mental states of Henry<sub>n</sub> and Henry<sub>g</sub>. Bearing the above consideration in mind, we may suggest that the appropriate answer we are looking for shall satisfy some version of McTaggart's Principle of the Dissimilarity of the Diverse.<sup>11</sup> In other words, since the *types* of mental states of Henry<sub>n</sub> and Henry<sub>g</sub> are distinct, there would be a property that the type of Henry<sub>n</sub>'s state of mind has, and the type of Henry<sub>g</sub>'s does not. Our target then is specifying what the property in question *is*. If there is such a property, thus very property should serve to distinguish the type of Henry<sub>n</sub>'s state of mind from that of Henry<sub>g</sub>'s. We also know that the following properties cannot satisfy the requirement: (1) the common properties shared by the types of mental states of Henry<sub>n</sub> and Henry<sub>g</sub>, and (2) the properties (such as certain relational ones, or the Cambridge ones) that do not impact the identity of the types of an agent's states of mind.<sup>12</sup>

As seen in the next section of the paper, the consideration of a philosophically significant, informative answer to the above question would reveal a challenging difficulty for the Williamsonian view of knowing.

---

### 3 The Difference between Two Types of Mental States

By observing the history of analytic philosophy, we may now easily classify

---

<sup>10</sup> According to Williamson, even granted that, in the given cases, there are no clues for them to tell whether an abundance of fake barns are present in the neighborhood, Henry<sub>n</sub> and Henry<sub>g</sub> are nevertheless in different (types of) states of minds (cf. Williamson 2000, 24–26, 51–64).

<sup>11</sup> The principle can be roughly formulated as follows: if *x* and *y* are distinct then there is at least one property that *x* has and *y* does not, or *vice versa*. See Chapter 10 of McTaggart 1921/1968.

<sup>12</sup> The elaboration concerning the nature of the property we are looking for is to meet one of the anonymous referees' demand for the clarification here.

certain theoretical strategies as unviable for accounting for the difference between the types of the mental states of Henry<sub>n</sub> and Henry<sub>g</sub>.

For one, the classical version of semantic externalism (cf. Burge 2007, 100–50; Davidson 2001, 15–38; Putnam 1975, 215–71) cannot help much, for semantic externalism normally holds that semantic content fails to supervene upon the agent's or speaker's internal features (Kallestrup 2012, 61–63). The comparison between the cases of Henry<sub>n</sub> and Henry<sub>g</sub> does not accurately parallel the comparison of the water on the Earth and the water on the twin Earth, for instance. Because the water on the Earth is H<sub>2</sub>O and the water-like liquid on the twin Earth is XYZ, the Earthian's belief that water is colorless and the twin Earthian's belief that water is colorless are actually different, no matter how qualitatively similar the relevant beliefs are, even granted that both the Earthian and the twin Earthian can sincerely report their beliefs by uttering the same sentence.<sup>13</sup> In this sense, the Earthian's and the twin Earthian's relevant beliefs are different, for their beliefs are about different watery stuff (H<sub>2</sub>O and XYZ, respectively). Semantic externalism can also readily explain why an Earthian's mental state fails to be knowing when he or she mistakes XYZ for H<sub>2</sub>O and falsely forms the relevant mental state in question. But the above strategy cannot be directly extended to generate a desirable solution to the cases of Henry<sub>n</sub> and Henry<sub>g</sub>, for they both form causally appropriate mental attitudes towards the genuine barns in question. In sum, the comparison between Table 1 and Table 2 nicely illustrates the significant dissimilarity between the Gettier-related cases and the twin-Earth-related cases.

Conceptually speaking, it would be rather unsurprising that semantic externalism does not help much here because semantic externalism mainly provides theoretical insights about the content of a proposition that may be embedded in a mental attitude. Our central concern, however, is the difference in the *states* of the minds in question (rather than the *mere* content of the mental states in question). This may also remind us of Williamson's discussion of the broadness of the mental state of knowing. Williamson explicitly criticizes Burge's account of factive mental states (which, of course, include knowing) (Williamson 2000, 50–51), for Williamson thinks Burge fails to appreciate the broadness of knowing fully. Williamson suggests that

---

<sup>13</sup> For a detailed discussion of Putnam's twin Earth argument, see Kallestrup 2012, 58–74.

knowing is a broad mental state, and its property of being factive does not imply that knowing is a hybrid state of mind that is composed of a purely mental part and a non-mental part (say, truth-part). In other words, even if one, in a certain case, rationally and truly believes that  $p$ , there is no case for this subject to be positioned in the same state of mind and to know that  $p$ . Thus, Williamson concludes that knowing, as a distinctive type of state of mind, is sharply distinguished from other types of mental states with epistemic merits (say, rationally and truly believing).

**Table 1** The Semantic Values of the Term “Water” in the Twin Earth Case

The term “water”	The watery stuff	The sentences uttered	The content of the mental states/thought
The Earth	H <sub>2</sub> O	Water is colorless	H <sub>2</sub> O is colorless
The twin Earth	XYZ		XYZ is colorless

**Table 2** The Semantic Values of the Term “It ” in the Gettiered Fake Barn Case

The term “it”	The observed object	The sentences uttered	The content of the mental states/thought
Henry <sub>n</sub>	The genuine barn	It is a barn	The genuine barn observed is a barn <sup>14</sup>
Henry <sub>g</sub>	The genuine barn		

The above consideration would naturally broaden the scope of our investigation from the *mere* content of the mental states to epistemological studies that center upon the states of minds in question. The very mental or psychological states, the formation processes, the causal mechanism, and the interaction between the cognitive subjects and their epistemic environments, etc., then become relevant. We would, in turn, shift our attention from semantic externalism to epistemic externalism.

Evidently, not every externalist theory of knowledge would work for us. For instance, the causal theory of knowing (Goldman 1967) cannot handle the fake barn case properly, since Carl Ginet and Alvin Goldman designed the fake barn case to deliberately reveal the insufficiency of this very theory. We

<sup>14</sup> It is confessed that, strictly speaking, the content of Henry<sub>n</sub>'s and Henry<sub>g</sub>'s mental states are different in a *token* way, since different genuine barns are observed by Henry<sub>n</sub> and Henry<sub>g</sub>. But again, this approach remains insufficient to provide a desirable answer to the given question, because it only provides us with a philosophically trivial description that fails to informatively tell us what distinguishes the type of Henry<sub>n</sub>'s mental state of knowing from the type of Henry<sub>g</sub>'s mental state.

have to appeal to some other versions of externalist theories of knowledge to explain the differences of Henry<sub>n</sub>'s and Henry<sub>g</sub>'s mental states.

This is one seemingly promising approach. According to the mainstream diagnosis of the cases, Henry<sub>n</sub>'s and Henry<sub>g</sub>'s mental states have different statuses of anti-epistemic-luck; in other words, Henry<sub>n</sub>'s mental state of knowing is epistemically robust, but Henry<sub>g</sub>'s mental state is only accidentally true, which can be easily defeated. In this sense, Henry<sub>n</sub>'s mental state of knowing is epistemically reliable and safe, properties which are entirely absent in Henry<sub>g</sub>'s mental state. Let me briefly summarize the above idea as follows: it is the property of being epistemically safe<sup>15</sup> that distinguishes Henry<sub>n</sub>'s mental state of knowing from Henry<sub>g</sub>'s. Namely, Henry<sub>n</sub>'s mental state is epistemically safe while Henry<sub>g</sub>'s is not. Thus, according to Leibniz's law (or, the indiscernibility of identicals), it is epistemic safety that helps us in telling Henry<sub>n</sub>'s mental state apart from Henry<sub>g</sub>'s.<sup>16</sup> At first glance, the above "solution" seems to be promising, and appears to satisfy our criterion for a plausible account of the difference between the types of Henry<sub>n</sub>'s and Henry<sub>g</sub>'s mental states. As will be shown, however, the above so-called "solution" is not tenable under scrutiny. The crux of the above "solution" is, I think, whether the property of being epistemically safe is a proper discriminative one that is applicable within the schema of Leibniz's law, so that we can reasonably tell the difference in the *types* of mental states. For instance, currently (say, it is time  $t_1$ ), when I am composing this paper, I am mentally conscious, and let us call this mental state  $C_1$ . Namely, I am in  $C_1$  at  $t_1$ . After a good sleep through the night, when I wake up the next day and continue to compose the paper, I am also conscious. Let us mark this situation by the statement that I am in  $C_2$  at  $t_2$ . Intuitively, my mental states of  $C_1$  and  $C_2$  do not seem to be different in type,

---

<sup>15</sup> It should be acknowledged that my way of using the concept of epistemic safety is somehow different from the paradigmatic one, since it is normally used to describe the epistemically favorable property of *belief*, which is in turn normally regarded as a component of knowing. In this sense, the mainstream talk of epistemic safety is in tension with Williamson's anti-reductionist view of knowing. Since Williamson (*Knowledge and Its Limits*, 41, for example) suggests that his account of knowing is (at least) conceptually neutral to the concepts of being justified, being caused, being reliable, etc. I deliberately adapt the use of the term "epistemic safety" to make the relevant discussion compatible with Williamson's general view of knowledge.

<sup>16</sup> Thanks to Edouard Machery for composing this remedy.

although  $C_1$  has the property of being marked by the temporal instance  $t_1$ , which is absent in  $C_2$ . In this case, we cannot validly apply Leibniz's law to derive a conclusion about the difference in types of  $C_1$  and  $C_2$ . The remaining task for us is to find a reasonable way to evaluate whether we can validly apply Leibniz's law, together with epistemic safety, to derive the "favorable" conclusion.

To avoid the accusation of begging the question or being *ad hoc*, I would like to present the following bacteria case with Dretske's Gettier-style adaptation:<sup>17</sup>

There are certain marine bacteria with internal magnets, and they originally reside in the northern hemisphere. Since oxygen is toxic to them, the bacteria's internal magnets align themselves with the geomagnetic north; they can then swim downwards from the surface of the ocean and reach the oxygen-free zone. In this sense, when their internal magnets function well in the northern hemisphere, we may name the relevant physiological state of the bacteria  $P_n$ . Thus, evidently,  $P_n$  is a *safe* state for guiding the bacteria in the northern hemisphere to the oxygen-free zone.<sup>18</sup> Now suppose a bacteriologist migrates some of the bacteria from the northern hemisphere to the southern hemisphere. Because their internal magnets remain the same during the transplantation, the type of the transplanted bacteria's physiological state would be identical with that of  $P_n$ . Thus, the transplanted bacteria, without any further change of their environmental setting, would be guided towards the ocean's surface and highly probably be killed in the oxygen-rich zone of the southern hemisphere's ocean. Accidentally, from the bacteriologist's pocket, a small bar magnet is unintentionally dropped into the water area where the bacteria have just been transplanted, which happens to change the local magnetic field. The sinking bar magnet would guide all transplanted bacteria towards the oxygen-free zone in the deep ocean. In this latter situation, although the

---

<sup>17</sup> Dretske's original bacteria case is presented in Dretske 1994, 164, 166. Hereafter, I would call my adaption of the case "the Gettiered bacteria case."

<sup>18</sup> There is a hot debate concerning what the bacteria's internal magnets functionally represent in the northern hemisphere. For instance, Dretske and Millikan significantly disagree with each other on that issue (see Dretske 1994, and Millikan 1994, 243–58). I set aside the relevant discussion about (mis-)representation throughout my paper, for nothing crucial in my Gettiered bacteria case hinges on the verdicts with respect to the content of the (mis-)representation.

type of the transplanted bacteria's physiological state is identical with that of  $P_n$ , the very state of the transplanted bacteria under discussion is no longer a *safe* (but rather a *lucky*) one for guiding the bacteria to the oxygen-free zone in the southern hemisphere.

The above Gettiered bacteria case structurally parallels the fake barn case. Intuitively, the transplanted bacteria are merely lucky to be guided by the bar magnet towards the bottom of the ocean and therefore avoid the fatal disaster of being killed by oxygen. To elaborate it modally: although the transplanted bacteria in the actual world are guided towards the oxygen-free zone due to the accident of the dropped bar magnet, in a nearby possible world where no bar magnet is dropped, the transplanted bacteria, by their internal magnets, are led into the oxygen-rich zone and are all killed there. In this sense, the transplanted bacteria with  $P_n$  state only *accidentally* get the right information about the new environment to which they have been transplanted. On the other hand, the bacteria in their original habitat always get the information about the oxygen-free zone correctly—but their internal magnets are in exactly the same state  $P_n$ . In sum, the state of bacteria's internal magnets,  $P_n$ , provides safe guidance in their original habitat, but becomes unsafe in the new one. I think the bacteria case vividly illustrates that the same physiological (or metaphysical) state can be safe in one situation and unsafe in others. Thus, the contrast of being safe and being unsafe is not a plausible way for us to distinguish the states in question.

If the same type of  $P_n$  can be safe in one situation and unsafe in another, we cannot reasonably distinguish the type of  $Henry_n$ 's mental state from that of  $Henry_g$ 's by suggesting the former is (epistemically) safe and the latter is unsafe. Someone may object that there is a significant dissimilarity between the Gettiered bacteria case and the fake barn case—it is  $Henry_n$ 's and  $Henry_g$ 's *mental* states that are under discussion, while  $P_n$  is rather a *physiological* state of internal magnets in the bacteria. Actually, it is not so difficult to meet this challenge by adapting the original Gettiered bacteria case and introducing the relevant mental state  $M_n$  that supervenes upon  $P_n$ . Suppose the bacteria in the case now evolve and develop a certain mental state  $M_n$  that supervenes upon  $P_n$ . According to the classic supervenience thesis in philosophy of mind, the bacteria remain in the same mental state  $M_n$  as long as its supervenient physiological basis is still  $P_n$ . By a similar pattern,

we can see that the mental state  $M_n$  of the bacteria before the transplantation is (epistemically) safe, but becomes unsafe after the transplantation.

If the above analysis is plausible, I think it is fair to suggest that standard externalism in epistemology cannot help Williamson with his thesis of knowing as a state of mind.

---

#### 4 Some Further Issues: Knowing as the Extended State of Mind in Prospect

In the previous discussion, I showed that some traditional externalist strategies cannot distinguish knowing as the Williamsonian state of mind from certain accidentally or luckily true beliefs. I confess, however, that this paper on its own does not suggest a conclusive rejection of Williamson's thesis of knowing, for the theoretical alternatives have not yet been exhaustively surveyed. We may still hold a certain hope that some to-be-developed version(s) of externalism, when introduced into Williamson's account, could lead to a desirable outcome. Some hints can probably be revealed from a reflection on our previous discussion.

Some philosophers who endorse Williamson's Knowledge-First approach may, for instance, suggest that Williamson is able to account for the difference of  $Henry_n$ 's and  $Henry_g$ 's mental states by appealing to his primeness externalism. Namely, according to Williamson, knowing, as a prime state of mind, is used primitively to explain certain other epistemological terms as well as epistemic phenomena rather than the other way around.<sup>19</sup> I think some clarification of my approach in this paper is in order here. Williamson's argument for the primeness of knowing is based upon the *reductio-ad-absurdum* strategy (cf. Williamson 2000). Suppose knowing is a composite mental state constituted by the mental, internal component, and the non-mental, external component. There would then be a case where the recombination of one's mental state is constructible out of the mental component from one case and the non-mental component from another, and one is still in the state of knowing. Williamson, however, strongly argues that the subject in the newly constructed case is not in the state of knowing. Here, again, I would like to re-emphasize my use of the

---

<sup>19</sup> I am grateful to one of the anonymous referees for this point.

Henry<sub>n</sub>-Henry<sub>g</sub> case (as well as the bacteria case) does not imply that Henry<sub>g</sub> is in the same type of state of mind as Henry<sub>n</sub> (let alone the so-called reductionism of knowing). I grant in the cases that Henry<sub>g</sub> is in a different type of mental state from Henry<sub>n</sub>. What is called for in this paper is an account of the intended differences (or different properties) that distinguish Henry<sub>g</sub>'s mental state from Henry<sub>n</sub>'s. In this sense, to embrace Williamson's primeness externalism does not further the understanding of the genuine difference between Henry<sub>g</sub>'s and Henry<sub>n</sub>'s mental states, since the primeness externalism used in this situation amounts to a restatement of the granted supposition I have already accepted and therefore fails to solve the relevant puzzle.

Some philosophers may suggest that even though primeness by itself cannot account for the relevant difference between Henry<sub>g</sub>'s and Henry<sub>n</sub>'s mental states, it is quite evident that Henry<sub>n</sub> is positioned in a normal epistemic situation and therefore has knowledge, but Henry<sub>g</sub> is in an epistemically unfriendly situation and thus fails to know. In other words, it is the situational difference that distinguishes Henry<sub>n</sub>'s knowing from Henry<sub>g</sub>'s belief. Again, I agree with this strategy in a general way—but this strategy is too coarse-grained to enhance our understanding of the difference between a Williamsonian mental state of knowing and merely true belief. We all realize that the fake barns in Henry<sub>g</sub>'s case prevent him from knowing. The genuine problem that interests us most is how to explain this prevention by appealing to a property of knowing that is missing in Henry<sub>g</sub>'s case. The relevant worry presented in this paper also sheds some light on how to accurately interpret the bacteria case: the bacteria case should be regarded as an illustration of a situation where a traditional epistemic safety account cannot successfully distinguish a Williamsonian mental state of knowing from the externalist merely truly believing.

Someone may suggest that the worry presented in this paper is committed to a strawman fallacy, for the term “a state of mind” is a technical term in knowledge-first epistemology and should not be understood in an ordinary sense.<sup>20</sup> According to this line of thought, Williamson could bite the bullet

---

<sup>20</sup> For instance, Brueckner suggests that the primeness of knowing makes the Williamsonian account of “a state of mind” distinct from the ordinary concept of “mental state.” See Brueckner 2002, 197–202. Recently, Smith has also argued that the Williamsonian thesis of knowing as a state of mind cannot be plausibly appreciated by the standard understanding of mental states in philosophy of mind. See Smith 2017, 95–112.

by saying that Henry<sub>n</sub>'s mental state is different from Henry<sub>g</sub>'s only in the Williamsonian sense of "a state of mind." But, in the ordinary sense of mental states, both Henry<sub>n</sub> and Henry<sub>g</sub> are in the same mental state. In this case, the comparison of the cases of Henry<sub>n</sub> and Henry<sub>g</sub> loses its initial power to generate the relevant puzzle for us.<sup>21</sup> I sincerely doubt Williamson would embrace such a rescuing strategy, which amounts to winning a battle but losing the war. This strategy severely undercuts the significance of the relevant disputes concerning Williamson's knowledge-first epistemology, for the disagreements between Williamson and his opponents may become just terminological. I think the strategy is not charitable enough for Williamson's project if we cannot prove that there is no further viable account for his thesis of knowing as a mental state.

I would lastly consider a criticism of the bacteria case, which may, in turn, shed some light on the prospect of Williamson's thesis of knowing as a mental state. One may correctly indicate that there are two necessary presuppositions in the analysis of the bacteria case here. One *explicit* presupposition is the commitment to the supervenience thesis, which remains contentious in contemporary philosophy of mind; the other presupposition is *implicitly* committed, for the physical, physiological, or mental states of the bacteria in question are all located strictly within the spatial boundaries of the biological organisms (i.e. the bacteria). To translate this implicit presupposition back to the analysis cases of Henry<sub>n</sub> and Henry<sub>g</sub>, we may discover that it amounts to presupposing that the mental states of Henry<sub>n</sub> and Henry<sub>g</sub> are spatially restricted to their respective bodies. Both presuppositions may be abandoned when we evaluate the mental states of Henry<sub>n</sub> and Henry<sub>g</sub>. Without these two presuppositions, the similarities between the Gettiered bacteria case and the fake barn case no longer hold.

I am quite sympathetic to the above fair criticism. The setup and the information in the cases of Henry<sub>n</sub> and Henry<sub>g</sub> are not explicitly committed to either presupposition listed above. The mental states of Henry<sub>n</sub> and Henry<sub>g</sub> in question may be functionally multi-realizable and epistemically extended beyond the skulls or bodies of the agents. In other words, what has been shown so far is that "passive externalism" (Clark and Chalmers 2002, 643–51) is unable to provide sufficient support to Williamson's thesis of knowing as a

---

<sup>21</sup> I owe this idea to one of the anonymous referees.

genuine state of mind. “Active externalism” (Clark and Chalmers 2002) has not entered the scene yet.

I also think the extended-mind or extended-knower approach may be promising, but this theoretical account still calls for substantial development. Elsewhere I have argued that Clark and Chalmers’ original functionalist version of active externalism is not straightforwardly applicable here because the case of Henry<sub>g</sub> and Henry<sub>n</sub> is substantively different from Clark and Chalmers’ case of Inga and Otto. Since extended cognition, the extended knower, and the extended mind are hotly debated issues, and the relevant literature multiplies fast, I do not have space in this paper to address active externalism, though it may set an agenda for further research. I would be satisfied with the limited conclusion so far that traditional, passive externalism is not sufficient to supply Williamson’s thesis of knowing as a genuine state of mind as a solution to the problem in question.

**Acknowledgments** Some early versions of the paper were presented on various occasions, and I would like to thank all its audiences for their comments, critiques, and objections. Thanks also to Peter Baumann, Michael Blome-Tillmann, Bo Chen, Haiqiang Dai, Linhe Han, Xingmin Hu, Yunpeng Jiang, Chuang Liu, Shuai Liu, Jianhua Mei, Ju Wang, Timothy Williamson, Zhu Xu, Chuang Ye, Feng Yu, and Zijian Zhu—they provided me with valuable suggestions and feedback during the development of the paper. I got detailed and insightful comments from two anonymous referees for the journal, to whom I am incredibly grateful. The research was supported by the National Social Sciences Fund of China (project no. 14CZX038) and the Project of Key Research Institute of Humanities and Social Sciences in Universities by the Ministry of Education of China (project no. 16JJD720003).

---

## References

- Alexander, Joshua and Jonathan M. Weinberg. 2014. “The ‘Unreliability’ of Epistemic Intuitions,” in *Current Controversies in Experimental Philosophy*, edited by Edouard Machery and Elizabeth O’Neill, 128–45. New York: Routledge Publishing
- Biro, John. 2013. “Showing the Time.” *Analysis* 73.1: 57–62.
- Biro, John. 2014. “Clocks, Evidence, and the ‘Truth-Maker Solution.’” *Acta Analytica* 29.3: 377–81.
- Blackburn, Simon. 2016. “Type-Token Ambiguity,” in *The Oxford Dictionary of Philosophy*, 3rd ed. Oxford: Oxford University Press.

- Boyd, Kenneth, and Jennifer Nagel. 2014. "The Reliability of Epistemic Intuitions," in *Current Controversies in Experimental Philosophy*, edited by Edouard Machery and Elizabeth O'Neill, 109–27. New York: Routledge Publishing.
- Bruেকner, Anthony. 2002. "Williamson on the Primeness of Knowing." *Analysis* 62.3: 197–202.
- Burge, Tyler. 2007. "Individualism and the Mental," reprinted in *Foundations of Mind: Philosophical Essays*, Vol. 2, 100–50. Oxford: Oxford University Press.
- Clark, Andy, and David J. Chalmers. 2002. "The Extended Mind," reprinted in *Philosophy of Mind: Classical and Contemporary Readings*, edited by David J. Chalmers, 643–51. Oxford: Oxford University Press.
- Davidson, Donald. 2001. "Knowing One's Own Mind," reprinted in *Subjective, Intersubjective, Objective*, 15–38. Oxford: Oxford University Press.
- Dretske, Fred. 1994. "Misrepresentation," reprinted in *Mental Representation: A Reader*, edited by Stephen P. Stich and Ted A. Warfield, 157–73. Oxford: Basil Blackwell Ltd.
- Goldman, Alvin I. 1967. "A Causal Theory of Knowing." *The Journal of Philosophy* 64.12: 357–72.
- Goldman, Alvin I. 1976. "Discrimination and Perceptual Knowledge." *The Journal of Philosophy* 73.20: 771–91.
- Heathcote, Adrian. 2012. "Gettier and the Stopped Clock." *Analysis* 72.2: 309–14.
- Hetherington, Stephen. 2012. "The Gettier-Illusion: Gettier-Partialism and Infallibilism." *Synthese* 188.2: 217–30.
- Hetherington, Stephe. 2016. *Knowledge and the Gettier Problem*. Cambridge: Cambridge University Press.
- Jacquette, Dale. 1996. "Is Nondefectively Justified True Belief Knowledge?" *Ratio* 9.3: 115–27.
- Kallestrup, Jesper. 2012. *Semantic Externalism*. New York: Routledge.
- McTaggart, J. Ellis. 1921/1968. *The Nature of Existence*, Vol. 1. Cambridge: Cambridge University Press.
- Millikan, Ruth. 1994. "Biosemantics," reprinted in *Mental Representation: A Reader*, edited by Stephen P. Stich and Ted A. Warfield, 243–58. Oxford: Basil Blackwell.
- Putnam, Hilary. 1975. "The Meaning of 'Meaning,'" reprinted in *Mind, Language and Reality: Philosophical Papers*, Vol. 2, 215–71. Cambridge: Cambridge University Press.
- Smith, Martin. 2017. "The Cost of Treating Knowledge as a Mental State," in *Knowledge First: Approaches in Epistemology and Mind*, edited by J. Adam Carter, Emma C. Gordon and Benjamin W. Jarvis, 95–112. Oxford: Oxford University Press.
- Turri, John. 2016. "Knowledge Judgments in 'Gettier' Cases," in *A Companion to Experimental Philosophy*, edited by Justin Sytsma and Wesley Buckwalter, 337–48. Malden: Blackwell Publishing.
- Vance, Chad. 2014. "Truthmaker Theory Does Not Solve the Gettier Problem." *Ratio* 27.3: 291–305.
- Williamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press.
- Williamson, Timothy. 2013. "Gettier Cases in Epistemic Logic." *Inquiry* 56.1: 1–14.