



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2022

Persistent burglars and knocks on doors: Causal indispensability of knowing vindicated

Logins, Artūrs

DOI: <https://doi.org/10.1111/ejop.12751>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-215004>

Journal Article

Published Version



The following work is licensed under a Creative Commons: Attribution 4.0 International (CC BY 4.0) License.

Originally published at:

Logins, Artūrs (2022). Persistent burglars and knocks on doors: Causal indispensability of knowing vindicated.

European Journal of Philosophy, 30(4):1335-1357.

DOI: <https://doi.org/10.1111/ejop.12751>

Persistent burglars and knocks on doors: Causal indispensability of knowing vindicated

Artūrs Logins 

Philosophisches Seminar, University of Zurich, UZH, Zurich, Switzerland

Correspondence

Artūrs Logins, Philosophisches Seminar, University of Zurich, UZH, Zürichbergstrasse 43, Zurich CH-8044, Switzerland.
Email: arturs.logins@uzh.ch

Funding information

Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Numbers: 186137, 169293

Abstract

The aim of the present article is to accomplish two things. The first is to show that given some further plausible assumptions, existing challenges to the indispensability of knowledge in causal explanation of action fail. The second is to elaborate an overlooked and distinct argument in favor of the causal efficacy of knowledge. In short, even if knowledge were dispensable in causal explanation of action, it is still indispensable in causal explanation of other mental attitudes and, in particular, some reactive attitudes and factive emotions. Taking into account this sort of causal efficacy in determining which mental states are genuine mental states opens up new perspectives for defending the view that knowledge is the most general factive and genuine mental state.

1 | INTRODUCTION

A well-known aspect of the so-called knowledge-first programme in recent epistemology is that it has two components.¹ The first is the doctrine that knowledge is first (in the relevant sense). This component contains both the claim that knowledge cannot be analyzed (defined) in terms of a combination of some further states and properties, and the claim that appeal to knowledge can define or substantially explain other states or epistemically interesting properties (e.g., epistemic justification, evidence, and so on).² The second component of the knowledge-first doctrine is that despite the fact that knowledge cannot be defined or analyzed, there is a positive theoretically insightful characterization of knowledge. Namely, we can characterize knowledge as the most general factive mental state (cf. Nagel, 2013; Williamson, 2000, p. 39; Williamson, 2011). While the first component of the knowledge-first approach has received sustained interest in recent debates, the second one has been relatively less discussed.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *European Journal of Philosophy* published by John Wiley & Sons Ltd.

In what follows our focus will be exclusively on the second component. In particular, we will focus on some recent challenges to the second component of the knowledge-first doctrine that relies on the assumption that genuine mental states have to be indispensable in causal explanation of action.

The aim of the present article is to accomplish two things. The first is to show that given some further plausible assumptions, existing challenges to the indispensability of knowledge in causal explanation of action fail. The second is to elaborate an overlooked and distinct argument in favor of the causal efficacy of knowledge. In short, even if knowledge were dispensable in causal explanation of action, it is still indispensable in causal explanation of other mental attitudes and, in particular, some reactive attitudes and factive emotions. Taking into account this new sort of causal efficacy in determining which mental states are genuine mental states opens up new perspectives for defending the view that knowledge is the most general factive and genuine mental state.

In what follows, I first specify the second component of the knowledge-first programme; I clarify the challenge from action explanation. I then show how the challenge can be met (e.g., introducing Williamson's well-known burglar and similar cases). Next, I revisit and reject existing objections to Williamson's proposal. In short, it seems that there is no alternative explanation of the relevant cases that would be relevantly superior to knowledge-centred causal explanations. I then show that knowledge can also be causally efficacious with respect to certain attitudes. Finally, I provide a brief discussion about some more fundamental issues that this discussion brings into light, namely, the question of why it is the case that knowledge explains some action better than belief and what is at stake in taking seriously the view that knowledge is a genuine mental state.

2 | KNOWLEDGE AS THE MOST GENERAL FACTIVE MENTAL STATE

An important aspect of the knowledge-first programme is that it proposes a positive characterization of knowledge as the most general factive mental state, or, in Williamson's terms: "knowledge is the most general stative propositional attitude such that, for all propositions p , necessarily if one has it to p then p is true" (Williamson, 2000, p. 39). In other terms, one knows that p , whenever one has a stative propositional attitude that p that can be only held to truths (alternatively: to facts). One knows that p , when one remembers that p , when one is aware that p , and so on. Of course, this characterization is a rough one and stands in need of further specifications to avoid some problematic consequences. For instance, one might think that this proposal entails a problematic overgeneralization of knowledge. Conditions, such as *being right that* or *believing truly that*, satisfy the factivity constraint (e.g., if one believes truly that p , then p is true). Yet, such conditions are not instances of knowing. Being right by mere luck that there are exactly 20,547 entries in the local phone book without even consulting the book is, quite plausibly, not an instance of knowing that there are exactly 20,547 entries in the phone book.

In his discussion of the claim that knowledge is the most general factive mental state, Williamson provides three principles that specify the rough characterization and avoid the problem of over generalization of knowledge. Here are the three principles proposed by Williamson to elucidate the claim:

1. If Φ is an FMSO (i.e., a factive mental state operator), from "S Φ s that A" one may infer "A."
2. "Know" is an FMSO.
3. If Φ is an FMSO, from "S Φ s that A" one may infer "S knows that A." (Williamson, 2000, p. 39)

The focus is on FMSO. The underlying assumption for this focus is that "[w]e can give substance to the category of factive stative attitudes by describing its realization in a natural language." (Williamson, 2000, p. 34). And that "[t]he characteristic expression of a factive stative attitude in language is a *factive mental state operator* (FMSO)" (Williamson 2000, p. 34). The important point for us is that FMSOs are not merely (i) factive (in the relevant sense), (ii) state denoting (used to denote states rather than processes), and (iii) propositional attitude denoting (used to ascribe attitudes to propositions to subjects) but are also (iv) semantically unanalyzable (see Williamson, 2000,

pp. 34–37). That is, their meaning is not composed out of other, simpler elements. To see that the claim that there are semantically unanalyzable terms is not an arbitrary stipulation, compare, for example, “cotton” to “red sweater.” While the former refers to a fiber, is simple, and cannot be analyzed by appeal to some further elements of meaning, the latter is a compound of “red” and “sweater.” Now, while “believe truly that” and similar expressions might satisfy the factivity, state denoting, and attitude denoting constraints, it can be further analyzed into “believes” and “truly” (where the mental state is believing, cf. Williamson, 2000, p. 39). No genuine FMSO can be further analyzed in this manner, for example, seeing that, remembering that, and similar are not compounds of some further elements. And while “is right that” might be factive and perhaps even unanalyzable, it is not attitude ascribing. It is not the case that “is right that” ascribes an attitude typically to a proposition to a subject. It is both natural and standard for “is right that” to be predicated of inanimate objects, for example, “The report is right that CFCs have declined” (source: <https://theconversation.com/are-cfcs-responsible-for-global-warming-14962>, see also Nagel, 2017, p. 530, fn 2 for somewhat similar observations). The conditions (i)–(iv), factivity, state denoting feature, attitude denoting feature, and semantic unanalyzability are all four essential features of FMSOs. Putative problem cases, for example, “believes truly that” and “is right that” lack one or other of these features and hence do not threaten the positive characterization of knowledge as the most general factive mental state.⁴

Knowledge is characterized by FMSOs given that “knows” has a special place among them. “Knows” is the most general among FMSOs. In what follows, I will use the label Superlative Mental Factivity Generalism about Knowledge (K-superlativism for short) for the thesis according to which (a) knowledge is the state picked out by “knows” and (b) “knows” is the most general FMSO. FMSOs are characterized as having features (i)–(iv) from above. I will interpret K-superlativism as a thesis about knowledge, not merely as a thesis about knowledge attributions. In short, I will assume that according to K-superlativism the state of knowing is best characterized by the conjunction of the relevant features of FMSO (i.e., features [i]–[iv]) and by the fact that knowledge is the most general state among states picked out by FMSOs. Before moving on to discuss how K-superlativism deals with the idea that mental states have to explain action, we need still to unpack it further a bit.

First, it is important to keep in mind that K-superlativism is only a thesis about how we can positively *characterize* knowledge. It is not a thesis that proposes to specify individually necessary and jointly sufficient conditions for one to know that *p*. Characterization of a state is clearly distinct from any sort of analysis or substantive explanation of a state. We can, for instance, characterize football as a game where players kick a ball around with their feet and try to score against the opposing team by getting the ball in their goal area. Of course, no one should take this characterization to state individually necessary and jointly sufficient conditions for playing football (maybe there is none). It should not be taken to constitute an analysis, or a substantial explanation of what football is in terms of some other properties and qualities (we are not analyzing football in terms of some combination of being a game and using one’s feet). Yet our rough characterization of football might still be insightful, in particular for someone who needs a rough and ready understanding of how to distinguish football from, say, footvolley. Similarly, it is important to remember that K-superlativism is not proposing to analyze knowledge in some further terms, by giving individually necessary and jointly sufficient conditions for having it, but rather, proposing a characterization that might be helpful for getting a rough and ready understanding of knowledge (in particular if one wants to understand the difference between knowledge and belief).

Second, it is also important to keep in mind that the focus here is on propositional knowledge—knowledge expressed by true statements of the form “*S* knows that *p*.” It is consistent with K-superlativism that there are other sorts of knowledge, that is, sorts of knowledge that are not propositional (cf. perhaps know-how or knowledge of particulars). For the sake of brevity, the specification of “propositional” will not always precede our talk of knowledge, even though, strictly speaking, what follows is only about propositional knowledge.

Third, according to K-superlativism, knowledge is a mental state. Mental states on this view are a subclass of mental attitudes, since some attitudes are nonstative, for example, the process of forgetting (cf. Williamson, 2000, p. 34), while states cannot be processes. It is a mental state, since it has to do with our mind (as opposed to, say, our digestive system).

Fourth, as we saw above, according to K-superlativism, knowledge is the most general factive mental state (again, this is an “is” of characterization, not of analysis or reduction). In short, as Williamson (2000, p. 34) tells us, “The main idea is simple. A propositional attitude is factive if and only if, necessarily, one has it only to truths.” We are abstracting here from the debate on whether facts and true propositions are the same (it is assumed for the sake of brevity that they are, but if they are not, it is possible to restate the view accordingly). Remembering that such and such is the case is factive. One can remember that such and such is the case only if such and such is the case. Similarly, seeing and being aware that such and such is the case are also factive. Now, “[t]he proposal is that knowing is the most general factive stative attitude, that which one has to a proposition if one has any stative attitude to it at all” (Williamson, 2000, p. 34).

3 | THE “REAL STATES EXPLAIN” CHALLENGE

It is a widely accepted assumption that for a mental state to be a *genuine* mental state, it has to be the case that all the entities of its type (or kind) can play a crucial role in causal explanation of action (Noonan, 1993, pp. 283–308; Reed, 2005; Williamson, 2000, p. 7). A genuine mental state is a mental state that cannot be reduced to any other mental state in a more fine-grained account of states that one is in. This general thought can be put as a parsimony principle about how to count real ingredients of one's mind: mental entities should not be postulated without necessity, where necessity is understood as the property of being indispensable in an explanation of an action. This principle enjoins us to not count a given mental state as a genuine mental state if it does not have a distinct contribution in a causal explanation of action, and more specifically, when its apparent role in such an explanation can in fact always be played by another (presumably more fundamental) mental state. In other terms, if assuming that X is a genuine mental state is uneconomical in action explanation, then X is not a genuine mental state (cf. Nagel, 2013).⁵

Now, the standard view in the philosophy of action is that we do not really need to appeal to knowledge in causal explanation of action. In other terms, we can dispense with knowledge. On the common view, it is belief that always does the real explanation.⁶ And in all situations where one might be tempted to appeal to knowledge in an explanation of some action, such an appeal to knowledge can really be replaced by an appeal to belief. And this can be done without any loss of explanatory value. Of course, such a view fits perfectly well within the so-called belief–desire or “Humean” model of action/intention explanation (Davidson, 1963; Dennett, 1987; Fodor, 1981; see also Nottelmann, 2011). According to the belief–desire model, the basic element in action explanation is the pair of belief–desire, for example, one buys a house because one believes that buying a house is the best means to achieve happiness and one desires to be happy. Of course, one may hold the common view (e.g., that belief, rather than knowledge, does the real explanation of the cognitive part in causal explanation of action) without endorsing the belief–desire model.

Of course, putting the above claims together, it follows that knowledge cannot be a genuine mental state. In other terms, if it is true that only states that cannot be dispensed with in causal explanation of action are genuine mental states, and knowledge can always be dispensed with in causal explanation of action (e.g., belief can always replace knowledge in these), then it seems that K-superlativism cannot be true. It cannot be true that knowledge is the most general factive mental state.

Not so fast. At this point, there is still room for a proponent of K-superlativism to resist the argument even if the above aspects of the common view are taken on board. Strictly speaking, what follows from the above assumptions, if true, is that knowledge cannot be the most general factive *genuine* mental state. But strictly speaking K-superlativism need not be an account about *genuine* mental states. Recall that it is a characterization of knowledge and a characterization may be compatible with a sort of reduction. Think again about football. It may well be that our characterization of football as a game played with the feet holds and also that there is a reduction of football to, say, outdoor games of some sort. Maybe a proponent of K-superlativism can accept that concerning the purposes of causal action explanation knowledge can be reduced to belief (belief being the really salient cognitive component in causal action explanation) but still also hold that it is best characterized as the most general factive mental state.

In other terms, there is space for a version of K-superlativism according to which knowledge is the most general factive mental state even though it is not the most general factive *genuine* mental state. One might argue that such a view might still be insightful in nontrivial ways. Perhaps, it might provide a rough and ready explanation of what knowledge is, as opposed to, say, forgetting.⁷

However, it appears that the proponents of K-superlativism or, at least some of them, are really concerned with it being a claim about knowledge being characterized as a *genuine* mental state. At any rate, this seems to be a natural interpretation of the fact that they do feel a need to defend K-superlativism in the light of considerations of the above sort (cf. Nagel, 2013; Williamson, 2000, pp. 60–64 and 75–92).

Thus, in what follows, we take it for granted that it is a real challenge for K-superlativism to explain how knowledge explains action (or rather, it is a challenge to the version of K-superlativism that we are exploring in the present context). The following quotation sums up the challenge that K-superlativists need to meet:

For, as already said, if a subset of (so-called) psychological states is demonstrated to be redundant in the psychological explanation of action, this is surely reason to regard them as not, properly speaking, psychological states at all (like knowledge, which is best regarded not as a psychological state, but as a complex consisting of a psychological state (belief) plus certain external factors – not because its status as knowledge is causally irrelevant in action explanation, but because it does not have to be cited, as such, in the psychological explanation of action at all). (Noonan, 1993, pp. 291–292).⁸

4 | THE ROLE OF KNOWLEDGE

The line of response that proponents of K-superlativism have taken to the above challenge is to deny the assumption that knowledge is dispensable in causal explanation of action. Indeed, they take it that knowledge makes a genuine contribution to causal explanation of action, at least in some cases.⁹ This contribution, they claim, cannot be reduced to the purported contribution of belief in action explanation (cf. Williamson, 2000, p. 86).

The idea is that there are situations where one's action cannot be properly explained without appealing to knowledge. Let us focus on two cases that Williamson presents to illustrate the purported indispensability of knowledge in causal explanation of action:

(Burglar—persistence in action)

A burglar spends all night ransacking a house, risking discovery by staying so long. We ask what features of the situation when he entered the house led to that result. A reasonable answer is that he knew that there was a diamond in the house. To say just that he believed truly that there was a diamond in the house would be to give a worse explanation, one whose explanans and explanandum are less closely connected. For one possibility consistent with the new explanans is that the burglar entered the house with a true belief that there was a diamond in it derived from false premises. For example, his only reason for believing that there was a diamond in the house might have been that someone told him that there was a diamond under the bed, when in fact the only diamond was in a drawer. He would then very likely have given up his true belief that there was a diamond in the house on discovering the falsity of his belief that there was a diamond under the bed and abandoned the search. In contrast, if he *knew* that there was a diamond in the house, his knowledge was not essentially based on a false premise. Given suitable background conditions, the probability of his ransacking the house all night, conditional on his having entered believing truly but not knowing that there was a diamond in it, will be lower than the probability of his ransacking it all night, conditional on his having entered it knowing that there was a diamond in it. (Williamson, 2000, p. 62).

And the second case:

(Knocking on the door—taking offence)

You see someone coming to your door; he is about to knock loudly. You are tempted not to reply. How would he react? You ask yourself, “Does he know that I am in?” not, “Does he believe that I am in?” If before knocking he does know that you are in, then he is unlikely to abandon his belief if you fail to reply; he will probably take offence. If before knocking he believes (truly) without knowing that you are in, then he is much more likely to abandon his belief if you fail to reply; he will probably not take offence. If before knocking he fails even to believe that you are in, then he is even less likely to take offence. Whether he would take offence is better predicted by whether he knows than by whether he believes. His taking offence is more highly correlated with knowing that you are in than with believing (truly) that you are in. (Williamson, 2000, p. 86).

Williamson admits that a proof that knowledge is indispensable in causal explanation of action cannot be provided, for to do that one would need to eliminate any possible combination of belief, truth, justification or some further condition. Yet “[t]here are infinitely many potential substitutes which might be proposed” (Williamson, 2000, p. 63). Thus, Williamson limits himself to sketching a recipe for showing that potential substitutes will not suffice for explaining action. The strategy is to focus on the potential substitute that does not provide individually necessary and jointly sufficient conditions for knowing and then to modify the above cases (or indeed, similar ones) in ways that demonstrate that “the failure of the necessity or sufficiency makes a causal difference, making the proposed substitute not even causally equivalent to knowing” (Williamson, 2000, p. 63). Indeed, it seems that cases where the alleged substitutes and knowledge appear to be equal with respect to their efficacy in causal explanation will always be cases where the substitutes implicitly refer to knowledge, in which case, they are committed to a circularity. The potential substitutes that do not appeal to knowledge can be shown to not be causally equivalent by appeal to the burglar or door cases (or, indeed, to appropriate variations thereof). In sum, according to one prominent K-superlativist argument, in cases where action is somewhat persistent and in cases where certain emotions or reactive attitudes are elicited, appeal to knowledge appears to be indispensable in causal explanation of these actions and attitudes. Knowledge is after all causally relevant and can count as a genuine mental state, according to this line of argument.

A further development of a K-superlativist approach has been proposed more recently by Nagel (2013). Elaborating on Williamson's account, Nagel argues that considerations from mental state attribution favor K-superlativism over its rival according to which belief is more fundamental than knowledge. In particular, Nagel has appealed to a vast body of empirical research from developmental, comparative, and social psychology to lend support to the view that knowledge is a genuine mental state and indeed more fundamental than belief (alternatively, Nagel's proposal can be seen as a defence of the fundamentality of the concept of *knowledge*; see footnote 7 above). Nagel suggests that it is a common view in developmental, comparative, and social psychology to treat knowledge as a mental state and indeed to appeal to knowledge (rather than belief) in explaining empirical material (for instance, appeal to knowledge is indispensable in explaining the observed difference between the awareness of ignorance coming before the awareness of error in developmental psychology, and related observations about this distinction in comparative psychology of nonhuman primates).¹⁰

Other philosophers have recently objected to Nagel's proposal (see Butterfill, 2013; McGlynn, 2017; Rose, 2015; Rysiew, 2013). For instance, in a recent and insightful article, McGlynn (2017) proposes a number of considerations aimed at questioning the significance of the empirical data in Nagel's argument. His discussion centers on mindreading and most concretely on data from developmental psychology (as well as from comparative psychology). In a nutshell, McGlynn questions the claim that (overall) data suggest that we acquire the concept of knowledge before the concept of belief.

While McGlynn's suggestions offer a lot to consider, for reasons of space, let me make only two brief points in passing. These are not meant to constitute an exhaustive treatment of McGlynn's criticism, but rather to serve as pointers toward a fuller line of response to his objections that knowledge-first sympathizers might explore. McGlynn (2017, sec. 2.1) observes that proponents of the knowledge-first view have relied importantly on the work of Perner (1993) in particular in making suggestions about the priority of the concept of knowledge. However, McGlynn insists, we should not conclude any priority of knowledge on the basis of the data from Perner, since such a conclusion would be inconsistent with Perner's background assumptions about knowledge. Perner seems to assume a causal theory of knowledge (knowledge as a caused mental representation) and refers to Dretske (1981) and Dretske and Enç (1984).¹¹ And taking this theory seriously seems to be in tension with the idea that we acquire the concept of "knowledge" before that of belief. According to McGlynn (2017, p. 79), "[a]ny argument for the priority of knowledge based on Perner's discussion will have to involve cherry-picking conclusions."

However, it is not clear that rejecting Perner's background Dretskean assumptions about knowledge, while taking on board his empirical data, would need to amount to any cherry-picking or involve any ad-hoc move, for that matter. We have some very good and independent theoretical reasons for thinking that at least early versions of the causal theory of knowledge are mistaken. It does seem to be quite obsolete and is subject to well-known counter-examples. Fake-barn style cases provide one paradigmatic counter-example. Difficulties in dealing with knowledge from sources other than perception, for example, mathematics, seem to be another problematic point for that sort of view. Arguably, our commonsense concept of "knowledge" is not (entirely) captured by the causal theory of knowledge. Thus, it is not entirely clear why we cannot take seriously some data from Perner and from others who might have some otherwise problematic background assumptions. Note also the same point applies even if Perner were to endorse the relevant alternatives theory of knowledge, since it is not clear why we should not be able to take Perner's empirical lessons onboard without endorsing the relevant alternatives theory, which is also well known to have its own problems (see Black, 2003 and Ichikawa & Steup, 2018 for overviews). Finally, note also that the way that Perner characterizes the relevant concept of "knowledge" that interests him is compatible with aspects of Williamsonian knowledge-first K-superlativism. Perner gives three elements that characterize knowledge:

This commonsense theory of knowledge has three critical aspects:

1. Truth (correct representation of facts)
2. Access to relevant information
3. Successful action (Perner, 1993, p. 146).

On one not-fetched interpretation of aspects 1–3, Perner has characterized genuine factive mental states here. Conditions 1 and 2 seem to ensure that knowledge is an attitude that can be held toward truths (facts) only. And condition 3 seems to fit well with the idea that knowledge is a genuine mental state, that is, a mental state that can figure in causal action explanation. If so, contrary to appearances Perner's conception of commonsense concept or view of knowledge is not necessarily in a radical conflict with a knowledge-first approach.

Another crucial point in McGlynn's dialectic is the suggestion that there are much better elaborated theoretical options than those considered by knowledge-firsters. According to McGlynn, options considered by knowledge-firsters, and Nagel in particular, are odd and not well-motivated interpretations of the relevant data. In particular, as potential alternatives McGlynn mentions the fully mentalistic approach, for example, Carruthers (2011) among others, and minimalist views, for example, Apperly (2011), Butterfill and Apperly (2013), and Ruffman (2014) among others. According to McGlynn, the oddity of Nagel's proposal (in Nagel, 2013) is that it entails "a relatively full-blooded mentalistic interpretation of children's early success on tasks involving the attribution of knowledge or ignorance but a relatively minimalist explanation of their success on nonverbal false belief tasks" (McGlynn, 2017, p. 88). This apparent oddity, however, appears to be dealt with in a more recent publication by Jennifer Nagel (2017). The focus of this more recent article is on factive versus nonfactive mental states and their attribution in our theory of mind, and Nagel elaborates there in detail the suggestion that actually our ability to attribute factive mental states

might be necessary for developing our ability to attribute nonfactive and, in particular, false mental states, for example, false beliefs. She writes: “Factive mental states are promising entry points of mental state attribution because in these basic nonverbal tasks the attributor needs to make sense of an interaction between the agent and the environment, and some salient fact in that environment needs to figure in setting the content of the attributed mental state” (Nagel, 2017, p. 535). Thus, if she is right, there need not be anything odd about the proposal whereby the capacity of knowledge-attribution is seen in quite full-blooded mentalistic terms, while the capacity to attribute belief is not. Moreover, Nagel seems to suggest that this approach might actually be quite useful in addressing challenges for mindreading in prelinguistic contexts (cf. Nagel, 2017, p. 535; see also the suggestion: “In these basic mindreading tasks, factive mental state attributions launch and guide attributions of nonfactive states,” Nagel, 2017, p. 538).¹²

Unfortunately, a proper full discussion of these issues goes beyond the framework of the present article (in addition to Nagel, 2017, readers might also want to explore another recent knowledge-first-friendly interpretation of the relevant data in Phillips et al., 2021). The focus of the present article is not on the defence of the whole knowledge-first programme. It is in a way more specific. It is not our aim here to discuss whether belief or knowledge is more fundamental (or whether the concept of *knows* is acquired before the concept of *believes*). As we saw in the introduction, the knowledge-first programme contains two components: one about the fundamentality of knowledge, the other about the characterization of knowledge as the most general factive mental state. The focus here is on the latter. However, we will return briefly to the more general and fundamental question of the overall theoretical motivation and consequences of taking knowledge to be a genuine mental state in Section 8.

5 | NEW “REAL STATES” OBJECTIONS

Despite K-superlativist arguments, a number of philosophers have found problematic the claim that knowledge is a genuine mental state that can have an indispensable role in causal explanation of action.

So, for instance, Magnus and Cohen (2003), in what might well be one of the first published reactions to Williamson's version of K-superlativism, argue that knowledge, as a broad mental state, cannot have the expected causal efficacy. To the contrary, they maintain that the explanation of the relevant cases can be provided by the appeal to belief-that-will-survive-scrutiny.¹³ However, one may reasonably doubt whether this objection can be maintained. For one thing, we need to unpack what is understood by “belief-that-will-survive-scrutiny.” On one understanding, it just is a stubborn belief. But we know from Williamson's discussion that stubbornness will not really be helpful in explaining a wide range of the relevant cases. Not all cases of a burglar knowing that there is a diamond in the house are cases of the burglar being stubborn about there being a diamond in the house. Indeed, a knower might in certain circumstances revise her beliefs and lose knowledge. But if by “belief-that-will-survive-scrutiny” one wants to refer to a state that has the status of being safe and sure with respect to truth whatever future inquiry might bring—that is, has some sort of positive modal robustness—then we might wonder whether one is not implicitly referring to knowledge by such a description.¹⁴ Indeed, knowledge does seem to be a state that will survive scrutiny, understood in the sense of positive modal robustness. Thus, it seems that Magnus and Cohen's objection does not undermine the K-superlativist claim that knowledge is indispensable in causal explanations.

Another argument in the literature against K-superlativism questions its claim that knowledge can really explain action, by focusing specifically on an alleged lack of explanatory generality of knowledge-based explanations of action. Bernard Molyneux (2007) assumes that an explanation of action needs to have a certain level of generality. An explanation of why a subject acted in such and such a way also has to work in all the relevantly similar cases where someone relevantly similar to the subject acted in the same (in a sense) way even without knowing the relevant proposition. The suggestion then is that knowledge lacks generality in explanation of action. According to this line of objection, only a narrow, internal state can satisfy the generality constraint in action explanation. Molyneux invites us to consider two situations relevantly similar to the burglar case. The first case is of a perfect internal, that is, physical duplicate of the burglar from Williamson's burglar case (think of the Davidsonian swampman) who

somehow replaces the original burglar before arriving at the house. It is suggested that such a physical duplicate, a duplicate who has exactly the same brain states, will go on to ransack the house even though the duplicate lacks the knowledge that the diamond is in that house. (The suggestion is that the duplicate lacks the knowledge since he does not even have the relevant demonstrative thought referring to “that” house.) The second case involves the original case of the burglar with the following slight modification: the diamond is removed from the house just before the burglar arrives (presumably without the burglar knowing about its removal). Molyneux maintains that the actions would be the same in all three cases: the subjects would go on to ransack the house all night. Yet only in the original case does the burglar know that there is a diamond in that house. Thus, according to this line of objection, knowledge lacks the relevant level of generality in explanation: if an appeal to knowledge cannot explain action in relevantly similar cases, then it is no good explanation of action at all.¹⁵

We should recognize that Molyneux's objection raises a challenge for the K-superlativist. How come that, if appeal to knowledge can explain action in the burglar case, it does not help to explain the same sort of action in physically similar cases (e.g., the duplicate and the replacement)? Indeed, one can even agree with Molyneux that K-superlativism lacks a certain sort of generality. Appeal to knowledge cannot explain *all* instances of similar actions. But we should nevertheless note that this is beyond the point that proponents of K-superlativism are suggesting. The generality to which Molyneux appeals is of the wrong sort. That is, the fact that appeal to knowledge does not explain in some sense similar cases of action is irrelevant with respect to the proposition that *sometimes* appeal to knowledge is what does the real job in explaining action. One needs to demonstrate that contrary to what our pre-theoretical judgments suggest, it is not appeal to knowledge that explains the original burglar's ransacking the house, but rather the appeal to the mysterious and powerful internal, narrow state. Indeed, an opponent of K-superlativism has to show that it is the narrow, internal mental state that does the explanatory job in a wide range of cases where one appears to act on knowledge (think of all the tremendous variety of burglar or door-knocking sorts of cases where one knows and persists in action) and that is *also* because one lacks that uncharacterized internal state in sufficiently similar cases to the burglar where one lacks the relevant persistence in action. On top of that, the opponent has to provide us with an error theory of why we intuitively think of knowledge as doing the explanatory job in these cases and not of that unnamed internal state. Once we see that Molyneux cases do not really help to mount such an argument against K-superlativism, we can admit that Molyneux is right that there are cases in some sense similar to the burglar case where one's action is not explained directly by knowledge (it might perhaps be explained by appearance of knowledge, though), but this is not really a problem. The claim that K-superlativists are putting forward is not that *all* instances of an appeal to belief can be replaced by an appeal to knowledge in action explanation. Rather, it is that in addition to belief we had better also admit into our mental ontology the existence of a genuine mental state of knowledge. That knowledge has the relevant level of generality in explaining a wide range of cases (action on what appears to be knowledge and lack of action in lack of knowledge) is already enough to accept its role as authentic in action explanation. It need not have the absolute level of generality in action explanation.¹⁶

Another more recent line of objection against the K-superlativist proposal that knowledge is a mental state that can do a genuine explanatory job comes from observations about the difference in correlation between action and continuous knowledge and action and continuous belief. This line of objection has been proposed by Jens Kipper (2018) and we now turn to evaluating it in more detail.

6 | CHALLENGE FROM BELIEF AT THE MOMENT OF ACTION

According to a recent new line of argument from Jens Kipper (2018) against K-superlativism, there is an alternative and better explanation of the burglar case (and, presumably, other similar cases) that does not appeal to knowledge in explaining action. Given the availability of such an explanation, we cannot accord an indispensable role to knowledge in action explanation. The alternative explanation of the burglar's ransacking that Kipper proposes appeals to *continuous* belief rather than *initial* knowledge as in Williamson's original example (i.e., the burglar's knowledge when

he entered the house, cf. Williamson, 2000, p. 62). Kipper (2018, p. 2227) characterizes continuous belief as follows: “[the burglar] believes that there is a diamond in the house as he enters it and retains this belief the whole night through (*Continuous Belief*).” Kipper maintains that he follows Williamson in assuming that probabilistic correlation is a good indicator of a (good) explanation of action. The correlation here is captured by the correlation coefficient, which is calculated on the basis of the conditional probability of the explanandum given the explanans and the probability of the explanandum given the negation of the explanans (in short, a higher degree of the former measure and a lower degree of the latter measure contribute to a higher overall coefficient). It is important to note that by continuous belief, Kipper means a belief *at the moment* of action. Unsurprisingly then, Kipper shows that the correlation coefficient between action (persistent ransacking) and continuous belief is higher than the correlation coefficient between action (persistent ransacking) and initial knowledge. Moreover, he also shows that the correlation coefficient between action and continuous belief is higher than the correlation coefficient between action and *continuous knowledge* (e.g., knowledge maintained during the action). This latter conclusion seems to rely ultimately on the idea that it is easier to lose knowledge than to lose belief. Now, these results show, according to Kipper, that (continuous) belief is always explanatorily superior to knowledge in explaining action. Indeed, following this line of reasoning seems to lead us to the conclusion that knowledge can be dispensed with in explaining action and cannot be considered as a genuine mental state.

However, a more attentive look at the details of Kipper's proposal reveals a flaw in this argument. To see this, recall Williamson's initial treatment of the case. In particular, reconsider the following passage: “A burglar spends all night ransacking a house, risking discovery by staying so long. *We ask what features of the situation when he entered the house led to that result.* A reasonable answer is that he knew that there was a diamond in the house” (Williamson, 2000, p. 62, emphasis added). It is important for Williamson's argument that we are focusing on the *prediction* question, that is, a question about the features of the situation that make it the case that a certain event is going to occur. In short, the focus is on a question requiring a *causal* explanation. And indeed, only this sort of question is relevant in this context. To recall that the overall dialectic here is one where we are trying to figure out whether knowledge is a genuine mental state, and we are accepting the constraint according to which only states that play an essential role in causal explanation (of action) are genuine mental states. Thus, the focus has to be exclusively on causal explanations. However, it is clear from the discussion that Kipper's focus is on something else (and actually he seems to be aware of this, since he feels the need to add a section on *causal* efficacy of knowledge, Section 4). For it is difficult to see why a correlation coefficient between a state occurring at the very moment of an action and the action should be seen as an indication of a causal relation between the state and the action in question. Consider my state of digesting and my action of walking around my office in the afternoon—that is, the state of me digesting all afternoon through my walking around my office. The correlation coefficient is extremely high here. But do we really want to say that my digesting has any causal impact on my walking? Or consider the state of being afraid that an aggressor will hurt one during an assault. It may well be the case that the correlation coefficient between being afraid of being hurt by an aggressor and being assaulted has the highest correlation coefficient. But concluding that fear causes assault is not only extremely implausible, it is also not sensible. Thus, it may well be that Kipper is right that the correlation coefficient between an action and a belief occurring during the action has the highest degree, but it is just not clear at all why we should think that this correlation is an indication of a good *causal explanation*. Thus, a proponent of K-superlativism may respond to Kipper's argument by claiming that his results about the correlation coefficients in this context are irrelevant for the question under discussion, namely, the question of whether knowledge can always be dispensed with in *causal explanation* of action (by invoking belief). That a state A is very likely to occur during event B does not make it more likely that B will occur given that A occurred rather than that B will occur given that another state C occurred, that is, that A and B typically tend to happen at the same time does not guarantee that A causally explains B.

That a belief is more likely to occur during an action is beyond the K-superlativist point that sometimes knowledge is causally efficacious. Yet one might want to hear an explanation from a proponent of K-superlativism of why even in situations where knowledge plays a genuine role in a causal explanation of an action, the corresponding belief is still more likely to co-occur with the action in question than the corresponding state of knowledge is.

However, it is not clear that there really is a genuine need for a deeper explanation of this than what we already have in Kipper's own initial observation about the difference in the relevant correlation coefficients. That is, it seems that a proponent of K-superlativism is warranted to reply to the request for an explanation of why even in cases where knowledge provides causal explanation of an action, continuous belief seems to be more strongly (probabilistically) correlated with the relevant action than continuous knowledge with the action, by simply pointing to Kipper's claim that there are many more ways of losing one's continuous knowledge than there are ways of losing one's continuous belief. Recall that in making his case for the claim that continuous belief is more strongly (probabilistically) correlated with action than continuous knowledge is, Kipper suggested that:

There are many ways in which the burglar could fail to have Continuous Knowledge even though he has Continuous Belief, and it is clear that the probability of his continuing his search is much higher if he continues believing that there is a diamond in the house without knowing it, than if he does not continue believing it. (Kipper, 2018, p. 2228).

In other terms, the burglar would lose continuous knowledge if any of necessary conditions for knowledge ceased to obtain at any given time during the course of the action. It would happen, for instance, if the relevant belief at any given time failed to be safe while the burglar still had the relevant continuous belief. This fact about there being more ways to lose continuous knowledge than to lose continuous belief, then, is all that is needed to explain why action is more strongly correlated with continuous belief than with continuous knowledge, even given that knowledge is still indispensable in some causal explanations of action. Given that mere probabilistic correlation does not exhaust causal explanation and as long as Kipper's own initial observations about the relevant correlation coefficients are accepted, it seems that we need not seek any further explanation of the relevant contrast in correlation coefficients in continuous belief versus continuous knowledge. K-superlativism is concerned with causal explanation of action (behavior prediction), not mere probabilistic correlation between continuous knowledge and continuous action.

Finally, note that Kipper's argument implausibly predicts that not only is initial knowledge always dispensable in causal explanation of action, but so is initial belief, and indeed, any initial state whatsoever. For the correlation coefficient of continuous belief and action will always be higher than the correlation coefficient of any other initial state and action. However, this is problematic, since we do want to have legitimate causal explanations that appeal not only to initial knowledge states but also to initial belief states (initial desires, initial intentions and so on). I suggest that this result is another indication of the flaw in Kipper's argument discussed above. Namely, it illuminates its failure to focus specifically on causal explanations rather than probabilistic correlations of states and actions in general. Thus, I conclude that the objection that relies on probabilistic comparisons of knowledge and continuous belief fails.

7 | REINFORCING THE CASE FOR K-SUPERLATIVISM: CAUSAL EXPLANATION OF ATTITUDES

Our discussion until now has mostly focused on the claim that knowledge is a genuine mental state. We considered an initial challenge to this claim, according to which only genuine mental states can play an indispensable role in causal explanation of action, and knowledge is always dispensable in causal explanation of action. We then revisited Williamson's case-based argument in favor of the contrary claim, according to which knowledge does in some situations play an indispensable role in causal explanation of action. Finally, we explored and replied to existing objections to Williamson's positive argument. The bulk of the discussion until now has been centered in particular on the burglar case from Williamson (2000) and aimed to show that the existing objections to Williamson's argument from the existence of the burglar and similar cases missed their targets in one way or another. In other words, the dialectic until now has been mostly defensive. I take it that what precedes has shown that existing attempts at undermining the argument from burglar-style cases toward the claim that knowledge is indispensable in causal explanation of

(some) actions are all unsuccessful. I would now like to go on the offensive and elaborate an overlooked independent positive case for K-superlativism. The gist of it is that even if there were a good undermining argument against burglar-style cases, there are other options available for K-superlativists to defend the claim that knowledge can be causally indispensable and thus has to be considered as a genuine mental state. The trick is to admit that the causal efficacy in action is not all the causal efficacy that genuine mental states can have.

Let us start with an unduly forgotten aspect of one of Williamson's initial cases. Reconsider the case of knocking on doors. In describing the case, Williamson wrote: "You see someone coming to your door; he is about to knock loudly. You are tempted not to reply. How would he react? [...] If before knocking he does know that you are in, then he is unlikely to abandon his belief if you fail to reply; he will probably *take offence*. [...] Whether he would *take offence* is better predicted by whether he knows than by whether he believes" (Williamson, 2000, p. 86, emphasis added). The unduly overlooked part of this example is that, strictly speaking, the example is not about action. It is about a reactive attitude. It is about a visitor taking offence, an attitude that involves the visitor participating in a human relation with you and the visitor becoming upset or angry that you do not open the door. Thus, elaborating on Williamson's initial insight, what this example shows is that someone's reactive attitudes, for example, an attitude of someone being upset/angry that you do not open the door is better causally explained by appeal to that person knowing that you are at home than by that person's true belief that you are at home. Someone being angry that you do not open the door entails/implies that you are at home (*being angry that p* is factive; see below); thus someone being angry that you do not open the door at least requires that that person has a true belief that you are at home, but such a true belief is more likely to be there after the person has knocked for a while, given that it is an instance of knowing that you are at home rather than just mere true belief. An appeal to someone knowing that you are at home, then, is a better causal explanation of that person being angry that you do not open the door, than that person merely having a true belief that you are at home. Thus, the door-knocking case is a case of causal explanation of an emotion; it is a case of causal explanation of attitudes. It is not a case about one's knowledge better explaining some action. All of the existing objections to K-superlativism have focused on the indispensability of knowledge for action explanation. However, we should not forget about the causal explanation of attitudes. And it seems that this and similar cases suggest that knowledge has a particular role to play in the causal explanation of some attitudes (e.g., factive emotional attitudes). Even if it were the case that behavior is not best explained by an appeal to knowledge, one might still maintain that knowledge has an irreducible role to play in the mental economy.

Two more remarks are relevant here before concluding. First, the current argumentative move presents some parallels with Eric Schwitzgebel's defence of *phenomenal dispositionalism* about belief (cf. Schwitzgebel, 2002, 2010, 2013). In short, Schwitzgebel proposes to avoid standard objections to classic dispositionalism about belief (e.g., roughly, the view that a belief just is reduced to/analysed in terms of one's dispositions to behave in certain ways) by including in his account of belief not only the appeal to one's dispositions to behave, but also an appeal to one's other mental states, and, in particular, ways one might feel about things. This is not the place to discuss at any length the merits and pitfalls of phenomenal dispositionalism. The parallel is only intended to show that there might be independently plausible grounds for focusing on mental states' causal efficacy beyond its role in the causal explanation of action.

Second, a clarification about factive emotions might be relevant here. It is a rather common view in the philosophy of emotions that some emotions seem to presuppose the truth of their content. A *locus classicus* defence of this view was presented by Gordon (1987); see Dietz (2018) for a more recent defence). Gordon has suggested (cf. Gordon, 1969, 1987) that all emotions with propositional content are either "*factive emotions*" or "*epistemic emotions*." Examples of the former include being amazed (that p),¹⁷ being annoyed, being disappointed, being excited, being grateful, being horrified, being proud, being surprised, being upset, being amused, being ashamed, being disgusted, being furious, being indignant, being resentful, being thankful, regretting, being angry, being delighted, being embarrassed, being glad, being pleased, being sad, being sorry, and being unhappy. Epistemic emotions include being afraid, hoping, being worried, fearing, being hopeful, being frightened, and being terrified (Gordon, 1987, p. 27). On Gordon's view, factive emotions *require knowledge*, while epistemic emotions *preclude knowledge*. Now, some philosophers of emotions have objected to Gordon's view, on various grounds. In particular, some doubt that

factive emotions really require knowledge (see de Sousa, 1987, pp. 109, 138, Wollheim, 1999, pp. 103–110, Ben-Ze'ev, 2001, p. 543, n. 20). But it still appears that a common view accepts at least apparent factivity of emotions of the first group (“factive emotions”). That is, it seems that according to a common view, emotions from the first group seem, *prima facie*, to require the truth of their content (but may contain exceptions in atypical cases). Of course, Gordon's view is a stronger one. Consider, for instance, one of Gordon's examples of factive emotions: according to Gordon, it is bizarre to utter: “Yes, and some people are quite *upset* that there are Martian spaceships circling the Earth”; or “I have a crazy neighbor who is glad that there are Martian spaceships up there” (Gordon, 1987, p. 36). According to Gordon, these assertions are not felicitous because “no one, crazy or not, is upset or glad that *p*, unless *p*” (Gordon, 1987, p. 36). Fortunately for us, we do not need to settle that debate here. All that is needed for the present purposes is that what we have called factive emotions, at least, typically presuppose the truth of their content. Thus, it would require exceptional circumstances (or be impossible, if Gordon is right) for someone to be upset that you are at home and do not open the door, while you are not at home.

Now, one might worry that an appeal by a proponent of K-superlativism might backfire. For, if these states are really factive mental states, then by its own lights, K-superlativism has to classify these as instances of knowledge. But, clearly, factive emotions are not states of knowledge (see Hyman, 2014 for this line of worry about K-superlativism). However, one might, I think, avoid this worry. The move would be to admit that factive emotions, like being sad that *p*, are partly constituted by one knowing that *p*; they do not amount to states of knowing that *p* properly understood. Another move would be to insist on differences between other standard factive mental states, for example, remembering that *p*, seeing that *p*, and factive emotions. One might, for instance, explore the suggestion that while ascriptions of full-blown factive mental states *semantically* entail true propositions (e.g., “that *S* sees that *p*” semantically entails “*p*”), ascriptions of factive emotions only pragmatically entail/presuppose the relevant true propositions (e.g., “*S* is sad that *p*” pragmatically presupposes “*p*”). This difference might signal a mere constant correlation between truth and state in the case of factive emotions but a constitution relation in the case of factive seeing and factive memory. Finally, a more radical move would explore the idea that factive emotions are indeed states of knowing, perhaps knowing of values. A full-blown elaboration of these possible strategies, however, goes beyond the scope of the present project.

In sum, focusing on factive emotions allows a proponent of K-superlativism to provide a different line of argument in favor of the idea that knowledge is a genuine mental state. It seems that examples like Williamson's initial door-knocking example, and indeed many other cases of factive emotions, suggest that sometimes one's initial knowledge is causally relevant for some of one's emotional attitudes.

8 | BACK TO THE FUNDAMENTAL QUESTIONS

In the light of our preceding discussion, one might well accept our replies to the existing arguments against the indispensability of knowledge in the causal explanation of action but still be perplexed about the general knowledge-first project. In particular, one might be still uneasy with the idea that knowledge is a genuine mental state. If knowledge is a genuine mental state, then one has not merely to give up the internalist (rough) view according to which one's mental states are determined (in a sense) by one's internal physical state; one has to embrace quite a radical externalist approach to the mental on which the mental states one has may depend commonly on physically and temporally quite distant factors. How could K-superlativism account for the apparent oddness of such a consequence? Furthermore, one might wonder about the exact connection between action, knowledge, and belief according to K-superlativism. In particular, one might think that unless one gives up the idea that belief is entailed by knowledge, belief is still somewhat relevant in causal explanation of action. That is, belief is present in all cases of action explanation, even in cases that seem to speak in favor of the idea that knowledge is indispensable. If so, how exactly then are we supposed to understand the claim that knowledge is a better explanation of (some) action? One might think that if belief is always present, perhaps it can also constitute an explanation of action in

cases where knowledge appears to explain better, an explanation that would not be as good as the explanation by knowledge, but an explanation nonetheless. But what exactly makes a better explanation of action in the relevant cases? The aim of this section is to tackle these more fundamental questions in greater detail.¹⁸

Let us start with the question about why knowledge is a better explanation of (certain) action than belief is. That is, let us first consider whether proponents of K-superlativism are in a position to offer something of a more fundamental explanation of why exactly knowledge sometimes (causally) explains action better than belief. One might suspect that if it is really the case that sometimes knowledge provides a better explanation, contrary to what some might think, then there must be some more fundamental reason for this. It is not by pure chance that knowledge sometimes explains better. And after all, if knowing that *p* entails believing that *p* (without being constituted by belief), why is it the case that it is not a belief that is always a better explanation? An opponent of K-superlativism is then in a position to require such a more fundamental explanation. Without it, one might think, the case for K-superlativism is somewhat incomplete.

I would like to propose two tentative and complementary lines of thought in reply to this question/worry. The first appeals to some recent empirical work on mental state attribution, that is, theory of mind, and one possible interpretation of this. It is a speculative suggestion, since it relies crucially on some recent empirical work that is yet to be further tested (I rely here on the literature overview and arguments from Phillips et al., 2021, Turri, 2017, and Nagel, 2017; see these for further references; see also Bricker 2020, 2021 for a different line of argument in favor of the thought that knowledge is a genuine mental and noncomposite state, an argument partly based on evidence from cognitive neuroscience). According to a common view in the theory of mind, the attribution of belief states (to other subjects) is more fundamental than the attribution of knowledge. And correspondingly, it has been thought that attribution of belief and desire is correlated with prediction of behavior. This common view, however, has been challenged recently. It has been observed that data from developmental, comparative, and social psychology are either nonconclusive or seem to speak in favor of knowledge attributions being more fundamental in the relevant sense (Nagel, 2017; Phillips et al., 2021). The sense of fundamentality here may be first introduced in terms of one mental state attribution depending on another. Say, if belief is more fundamental or basic than knowledge, then the capacity to attribute knowledge to others depends somehow on the capacity to attribute belief. Phillips et al. (2021, p. 8) also provide a more specific way of understanding the basicness in this context by providing four criteria that scientists tend to appeal to in theorizing about which aspects of human cognition are more basic. Among these are:

1. being more ancient with respect to evolutionary history. The idea here is that aspects we share with other related primate species should be taken to be more basic, since they are “older” in the evolutionary sense.
2. Emerging earlier in human ontogeny. The idea here is that the earlier the capacity emerges, the less it requires experiences and is in this sense more basic.
3. Being automatic. The idea here is that it would be unexpected for a more basic capacity not to operate automatically, that is, without “conscious initiation” (Phillips et al., 2021, p. 8).
4. And being present in special populations. The idea is that capacities “that are more basic tend to be conserved across populations despite radical differences in experiences or deficits in other cognitive processes” (Phillips et al., 2021, p. 8).

According to Phillips et al., some of the data from the relevant sciences are inconclusive, but some data seem to speak in favor of knowledge attributions being more basic in the sense specified by these four criteria. The conclusion about knowledge being more basic in the relevant sense of basicness in mental state attribution seems to fit particularly well with the results of another recent study of social cognition that focuses on the role of belief attribution and knowledge attribution in behavior predictions (Turri, 2017). The conclusion in that study is that knowledge attributions seem to be more strongly correlated with behavior predictions than belief attributions. In short, knowledge leads to action predictions more reliably than belief does. To give an idea of one of the experiments, consider the following (rough) presentation of a setting that was used. Participants were asked whether in a given situation

with relevant background information a subject, Laurie, will say that Josh is in the apartment, whether Laurie believes that Josh is in the apartment, and whether Laurie knows that Josh is in the apartment. The study found that attribution of knowledge to Laurie is more strongly correlated to the attribution of the relevant behavior (i.e., Laurie will say that Josh is in the apartment), or in Turri's (2017, p. 2256) terms, "at least in some ordinary contexts, knowledge attributions guide behavioral predictions in human social cognition" (see Turri, 2017 for details). Another experiment tested whether belief or knowledge attributions contributed to faster behavior predictions and found that knowledge attributions did better on this score as well (Turri, 2017, p. 2258). If this is on the right track, then knowledge attributions seem to be at least in some contexts better correlated with predictions about action than belief. One tentative lesson that can be drawn from this, according to Turri (2017, p. 2258), is that "[i]f belief attributions are more complex, slower, and less reliable for predicting behavior, then that would naturally lead people to rely, by default, on knowledge attributions instead." Thus, my first line of thought in reply to the above question is the tentative suggestion, inspired by the abovementioned studies, that knowledge explains some action better than belief because knowledge is itself a more basic or fundamental element of cognition than belief is. Again, this is a very speculative suggestion, and it relies heavily on empirical theses that need to be further confirmed, but it would not be surprising if K-superlativism and the idea that knowledge is more fundamental than belief is in theory of mind go together. Knowledge explains better, on this tentative suggestion, because it is psychologically more fundamental in our folk theory of mind. Of course, important further fundamental questions might be raised (e.g., What is a genuine action explanation at all? How can we adjudicate between folk action explanations and theorist action explanations if they come into conflict?). However, if the above suggestions are on the right track, we can see here one potentially deeper explanation of why knowledge explains in a sense (some) action better, or at least why it may appear to us on pretheoretical grounds that knowledge provides a better explanation of some actions.¹⁹

My second thought in reply to the request for a deeper explanation of why knowledge explains better goes back to Williamson's original discussion on the ineliminable role of knowledge in causal explanations. Recall in particular Williamson's claim that knowledge is in a sense more robust, than, say, a true belief that does not rely on false lemmas but is not knowledge because of some (somewhat easily) discoverable misleading evidence: "Although knowing is not invulnerable to destruction by later evidence, its nature is to be robust in that respect" (Williamson, 2000, p. 63). It seems that knowledge enjoys an important sort of robustness. I would like to expand on this thought by suggesting that knowledge is more resilient to counter-evidence than mere true belief is. To see this, consider a series of knowledgeable burglar cases, C₁, C₂, ... C_n. All these cases are initially just like Williamson's burglar case, that is, when the burglar, in all these cases, enters the house, he knows that there is a diamond in the house. But they differ somewhat in how the subsequent events unfold. Let us say, in all of these cases, burglars receive some counter-evidence for the proposition that the diamond is not in the house. The difference lies in what exactly the evidence is and how/when they receive it. In these cases, the counter-evidence is not massive and burglars maintain the knowledge (and the belief) that there is a diamond in the house. For example, discovering that the safe is empty is one possible piece of counter-evidence in a burglar-style case. If a knowledgeable burglar discovers that the safe is empty, however, it is unlikely that he will lose knowledge that there is a diamond in the house (assuming that knowledge cannot be based on a false lemma). Of course, a burglar could lose knowledge (in a different case), if, for instance, he discovers massive counter-evidence, such as, perhaps, discovering signs of a diamond-owner's hasty departure and a highly reliable informant communicating to the burglar during the ransacking that the diamond has been taken out of the house just before the ransacking. But apart from cases of massive counter-evidence, it is most plausible that the burglar maintains the relevant piece of knowledge. And this is so, in the series of C₁, C₂, ... C_n. Things are different in the parallel cases C*₁, C*₂, ... C*_n of burglars who only have mere true belief that there is a diamond in the house, but otherwise are exactly the same as the knowledgeable burglar cases. It is natural to think that even in nonmassive counter-evidence cases, many of the burglars in C*₁, C*₂, ... C*_n will lose their beliefs (we can imagine that these cases are not cases of knowledge, but cases of mere true belief due to aspects that are not connected to the amount or nature of the counter-evidence that we have there, but, say, to the failure of the safety condition on the belief-forming method). That is, it seems that the ratio of burglars who maintain

their true belief that there is a diamond in the house in C^*1, C^*2, \dots, C^*n in the face of nonmassive counter-evidence is lower than the ratio of burglars who maintain their knowledge (and true belief) that there is a diamond in the house in $C1, C2, \dots, Cn$. The tentative suggestion then is that this is because of this ratio, that is, the resilience ratio in the face of counter-evidence factor, that knowledge seems to explain better the action of persistent ransacking in Williamson's burglar case. For it is because of knowledge that burglars in $C1, C2, \dots, Cn$ maintain their true beliefs.

Now, as Williamson himself observes (see Williamson, 2000, p. 63), one might be tempted to reply to this resilience-based explanation by noting that stubborn belief can also be highly resilient and perhaps even more resilient than knowledge in the face of counter-evidence. In reply to this worry, however, we can follow Williamson and point to the fact that appeal to true stubborn beliefs will not be satisfactory in this context. For not all cases of knowledge are cases of stubborn belief (cf. Williamson, 2000, p.63). But in order to demonstrate that knowledge is not a better explanation of action in cases like that of the burglar, we need to show that in *all* burglar-style cases where one knows one also has a stubborn belief. Yet, of course, such a condition need not be accepted. Not all our cases $C1, C2, \dots, Cn$ are cases of stubborn belief. Thus, this line of objection seems to be missing its target.

Finally, one might also object that even if appeal to knowledge seems to be appropriate in action explanation in burglar-style cases, there is still a sense in which it is belief that really explains action in such cases, given that knowledge entails belief. One could argue that it is nonetheless belief that is the common element in all cases of action explanation. There are of course cases of action where agents know the propositions upon which they act; in a sense these are "good" cases, since no error is involved here. But there are also cases where agents act without knowing the relevant propositions upon which they seem to act. We can call these the "bad" cases, since an error or at least the failure to know is involved. One might argue that perhaps knowledge explains best action in good cases, but not in bad cases, since *ex hypothesis*, the agents do not know the relevant propositions there. But our theory of action explanation, according to this line, should be simple, and we should always appeal only to one cognitive element in action explanation. Given that belief is the common element in good and in bad cases, it seems that an appeal only to belief offers us a simpler theory of action explanation. According to this line of thought, then, simplicity considerations favor the view that it is belief that is indispensable in causal explanations of action, given that knowledge entails belief.

To this, I would like to reply that it is not certain that such a proposal is explanatorily fruitful. If one takes seriously the K-superlativist proposal that knowledge is an indivisible, simple state, and cannot be explained or reduced to further more basic elements, then appealing to belief in cases where knowledge explains better just because belief is also entailed by knowledge does not seem particularly insightful. Such an explanation, by an appeal to belief, would still piggyback on the presence of knowledge. It is the knowledgeable status of the subject that does the relevant explanatory job, as we have seen above. Of course, knowledge still entails belief on the standard K-superlativist account, but it is not the presence of mere belief or true belief that is relevant in our target cases. What seems to matter in our cases is the fact that subject knows the relevant propositions. I would like to suggest then that the urge to simplify action explanation by looking for one common element in all action explanation should be resisted. Of course, simplicity *ceteris paribus* is a theoretical virtue. But it does not seem that we are here in a *ceteris paribus* situation. Action in burglar-style cases and in "good" cases in general does not seem to be comparable *ceteris paribus* to cases where one acts on a proposition without knowing it. The above discussion about resilience to counter-evidence is one aspect in which the two are importantly different.

Let us turn now to the other one of the fundamental questions that we mentioned above, namely, the question about further consequences for philosophy of mind of taking knowledge to be a genuine mental state.

More specifically, the worry that one might have here is that the view that knowledge is a full-blown, genuine mental state leads to counter-intuitive consequences with respect to causal and temporal interconnections of our mind and world. If knowledge is a genuine mental state, then our mental states can be altered not only unbeknownst to us, but even without the altering events having any causal connection to us. Even worse, as Smith (2017) has recently argued, the events that change our mental state may then be physically as well as temporally distant. Consider the following example from Smith that illustrates this point vividly:

I currently believe, and know, that my local fishmonger has a blue shopfront. Suppose that, while I am on a trip overseas, the shopfront is restyled and painted red. As soon as I return home and set eyes upon it, I will lose my belief that my local fishmonger has a blue shopfront. This, of course, will be a perfectly standard case of mental state change, in which my beliefs are revised in light of new information. But if knowledge is a mental state, then this is not the first time that the new paint job will have had an effect on my state of mind. Rather, this will have happened days or weeks before, while my body was thousands of miles away and in no causal contact with the fishmonger – for this is the point at which I will have ceased to *know* that my local fishmonger has a blue shopfront. And this change, whenever it happens, won't be due to any change in my internal state, which will continue exactly as it would have if the shopfront had not been touched. (Smith, 2017, pp. 106–107).

This example appears to put into light a rather counter-intuitive consequence of K-superlativism. If knowledge is a genuine mental state, then the mental states that someone is in can be affected (instantly) by very distant events of which the subject is not and cannot easily be aware *and* remains during this change of the mental state in the same internal physical state. Such a possibility appears to many to be too odd.

Proponents of K-superlativism recognize that one might find such consequences of their view counter-intuitive and be tempted to give up the idea that knowledge is a genuine mental state. However, they also suggest that the oddness here can be mitigated if one already accepts a form of externalism about the mental. Williamson (2000, p. 50) writes: “[A]n externalist conception frees us to affirm that knowledge is a mental state.” The main point of Smith's criticism, however, is that taking knowledge to be a mental state is a much more radical departure from the spirit of internalism, that is, from the above intuition that can be rendered explicit as the intuition that an event cannot alter one's mental states without interfering in one way or another (even if only indirectly) with one's body and sense organs. Smith objects that taking knowledge to be a genuine mental state is a much more extreme form of externalism about the mental than any presently known version of externalism about the mental in this respect. For versions of externalism about the mental predict that one's mental states can switch in ways that do not affect one's internal physical state but are less radical than the switches/changes in mental states that are predicted by the view that knowledge is a mental state. So, for instance, while natural kind externalism (e.g., think of the Twin Earth thought experiment, cf., Burge, 1979; McGinn, 1977, Putnam, 1973) predicts that switches in mental state are slow, proximal, and rare, knowledge-as-mental-state externalism predicts that such switches can be fast, distant, and rather common (see Smith, 2017, p. 108). Smith stops short of claiming that such consequences give us sufficient reasons for rejecting the view that knowledge is a mental state, but he does contend that they constitute a serious theoretical cost for the view nonetheless.

In reply to this line of criticism, I would like to propose two complementary thoughts. These are in no sense intended to provide an exhaustive treatment of the worry, let alone to solve the internalist–externalist debate. They are supposed to mitigate the force of this worry/objection from the apparently radical mental-state changes that K-superlativism seems to imply.

First, I would like to point to the fact that if the mental-state switch/change is a problem for K-superlativists who think that knowledge is a mental state, then it is also a problem for anyone who thinks that regret, happiness, joy, awareness, memory, perception, and so on (e.g., factive mental states) are genuine mental states. This is so, since the fishmonger case (as well as others) above can be easily adapted to other factive states. Consider, for instance, regret. Imagine that before the trip overseas, the subject of the above example tried to convince the owner of the fishmonger's shop to paint the shopfront yellow, but did not manage to do so. When he left, the shopfront remained blue. Now, a thousand miles away, he regrets that the shopfront is still blue despite all of his efforts. At this very moment, however, and unbeknownst to the subject, the owner changes his mind dramatically and paints the shopfront red. It would seem that we now have to admit that the subject no longer regrets *that the shopfront is still blue*. This is so because it is no longer true that the shopfront is still blue. Thus, if one takes the apparently counter-intuitive consequences in the original example to constitute an argument against the view that knowledge is

a genuine mental state, then one also has to take the modified example to constitute an argument against the view that regret (and any other factive mental state for that matter) is a genuine mental state. However, denying that all these states—for example, being annoyed that *p*, being disappointed that *p*, being excited that *p*, being proud that *p*, being surprised that *p*, being upset that *p*, being amused that *p*, being ashamed that *p*, being disgusted that *p*, being furious that *p*, being indignant that *p*, being angry that *p*, being delighted that *p*, being embarrassed that *p*, being glad that *p*, being sad that *p*, and so on (see Section 7 above)—are mental states seems to involve its own theoretical cost. If anything, such a denial would arguably amount to quite an impoverished vision of our mental lives.²⁰ Thus, given that all so-called factive mental states have similar consequences mental-state switch-wise, and given that a strong *prima facie* case can be made for counting these states as genuinely mental, I would like to suggest that we have good reasons to think twice before giving up the idea that knowledge is a genuine mental state merely on the basis that there appear to be counter-intuitive consequences in cases like that of the fishmonger.

The second thought that I would like to offer in reply to Smith's criticism is a proposal of how K-superlativists might reply more constructively to intuitions from Smith's cases. If Smith is right and we do have intuitions for internalism that are elicited in cases like that of the fishmonger, then we find ourselves in an uncomfortable situation. We seem to have both pretheoretical judgments that back up internalism and pretheoretical judgments that speak in favor of radical knowledge-as-a-genuine-mental-state externalism. What should we do in such a case of apparently clashing intuitions? If this is our predicament, we should not rely on intuitions alone. Our theoretical choice here should also be informed by considerations about the theoretical fruitfulness of all options. K-superlativism and the view that knowledge is a genuine mental state does seem to provide some fresh theoretical fruitfulness with respect to action explanation and factive-emotion explanation, as we have seen above. Moreover, the full K-superlativist picture brings in an elegant and simple overall view, where knowledge is a central element. Notably, it allows us to have a systematic approach to epistemology and beyond. Again, I am not saying we have anything close to an ultimate solution to the internalist–externalist debate here, or even a comprehensive treatment of the challenge raised by Smith. However, I would like to suggest that in addressing the suggestion that knowledge cannot be a genuine mental state because of internalist intuitions, we should also take into account these further two last points and should not rush to throw out K-superlativism too quickly. Considerations of theoretical fruitfulness and simplicity might guide us in the light of the internalist intuition challenge.

9 | CONCLUSION

The knowledge-first programme has two tenets. The first is that knowledge cannot be defined/analyzed in terms of some more fundamental conditions, but rather can be appealed to in theorizing about other states and properties. The second is that knowledge can be positively characterized as the most general factive mental state. This article has focused on the latter—for some reason less discussed—aspect of the knowledge-first approach. A crucial part of our discussion concerned the vexed question of whether knowledge is a mental state at all. A widely accepted criterion for distinguishing whether a state is a genuine mental state is said to be that the state in question is indispensable in causal explanation of action. We explored Williamson's positive argument based on possible cases towards the conclusion that in some situations appeal to knowledge is the best possible causal explanation of action. We examined the existing objections to Williamson's argument and concluded that none of them is successful in undermining it. We also offered a somewhat new (or unduly overlooked) argumentative line in favor of the idea that knowledge is a genuine mental state, from its role in causal explanation of reactive attitudes, and factive emotions. Finally, we considered two fundamental questions, namely, the question of why exactly knowledge provides a better explanation of some action and the question of what further consequences follow from the view that knowledge is a genuine mental state. Overall, our verdict is that knowledge does seem to be a genuine mental state and that there appear to be non-ad-hoc ways to deal with the fundamental questions just mentioned, which is not yet to vindicate

the full-blown knowledge-first programme. Yet, of course, it fits best with the general ideology of putting knowledge first.

ACKNOWLEDGMENTS

Many thanks to Edgar Phillips, Fabrice Teroni, and Lamba Boyz reading group at the University of Geneva (Benjamin Nesser, Constant Bonard, and Kyrill Khromov) for discussion and comments on earlier versions of this article. The research work that led to this article was supported by the Swiss National Science Foundation Grants Number 169293 and 186137.

ORCID

Artürs Logins  <https://orcid.org/0000-0003-4760-9027>

ENDNOTES

- ¹ See Williamson (2000) for a canonical statement of the programme in contemporary epistemology.
- ² See, for instance, Williamson (2000, pp. 184–208), Sutton (2007), Bird (2007), Hawthorne and Stanley (2008), Ichikawa (2014), Kelp (2016, 2018), Littlejohn (2017), Logins (2014, 2017), Silva (2017), Fratantonio (2018); see Benton (2014) and Silva Jr (2020) for recent overviews.
- ³ The principles (1), (2), and (3) are numbered as (18), (19), and (20) respectively in the original publication.
- ⁴ Thanks to an anonymous referee for this journal for urging me to clarify the formulation of the view in the light of the potential objection from mere veridical stative propositional attitudes.
- ⁵ One might wonder at this point, whether the idea (that we will explore in length in what follows) that knowledge is a genuine mental state and the fact that there are species of knowledge sit well together. One might worry that given the above assumptions about genuine/non-reducible mental states playing an indispensable role in the causal explanation of action, the following three claims seem to stand in tension: knowing that *p* is a genuine mental state, remembering that *p* (and other species of knowing) is a genuine mental state, yet remembering that *p* (seeing that *p*, being aware that *p*, etc.) reduces to knowing that *p*. More concretely, the worry is that it is unclear how knowing that *p* and remembering that *p* could be genuine mental states while remembering that *p* reduces to knowing that *p*. In reply to this worry, I would like to observe that the remembering-knowing case differs importantly from cases of standard non-genuine mental states, that is, states composed of more fundamental elements (i.e., surface mental states). Schadenfreude is most likely a composed state that involves a belief that someone unlikely is having some misfortune and pleasure at realizing that this person has that misfortune. Arguably, all cases of action explanation involving schadenfreude can be explained by appeal to the relevant beliefs and pleasure of seeing someone's misfortune. Remembering is not a composed state in this sense. The above tension can be solved by specifying that remembering does not reduce to knowing in the relevant sense, the sense of "being composed of." Knowledge does not compose remembering, while belief and pleasure do compose schadenfreude. We do not get a more fine-grained explanation of action by focusing on knowing (genus) instead of focusing on remembering (species), when we explain a given action. A more fine-grained appeal to more fundamental, more basic elements that compose the relevant surface state, for example, an explanation in terms of belief and pleasure, is more fine-grained than an explanation of the relevant behavior in terms of the surface state of schadenfreude. Second, I would like to note that the worry here might be an instance of a more fundamental puzzle of counting items belonging to a genus and items belonging to a species. Imagine that you have to count how many animals are there in a town. Should you count in all the town's huskies? It seems so. Should you count in all the town's dogs? Certainly, but if so, your animal inventory risks being inaccurate since you would double count some individuals (e.g., counting them both as dogs and as huskies). Note that whatever the solution to your double-counting problem is, it cannot be to rule out dogs or huskies as not being genuine animals. Something similar seems to be going on also in the case of knowing and remembering. Both of these seem to be genuine mental states. Which of these will figure in a given instance of an action explanation might depend on a situation's dialectical or other contextual factors. Moreover, note that this is different from what is happening with knowledge and belief according to a view discussed below. On the view we will discuss shortly, belief is always the relevant, genuine, fundamental mental state when knowledge seems to be involved in a causal explanation of action. That view is reductive in our sense. On that view knowing is a surface state, much like schadenfreude, and can be ultimately disposed of in causal action explanation. It is inconsistent with maintaining that both knowledge and belief are genuine mental states and that in all cases of action explanation where knowledge is involved in the causal explanation of action needs only appeal to belief. Many thanks to a reviewer for drawing my attention to the need to specify this point.

- ⁶ See Pettit (1986) for a dissenting view.
- ⁷ See also Ichikawa and Jenkins (2017) and Ichikawa (2017) on a somewhat related line of thought according to which there are two general possible interpretations of the claim that knowledge is basic: a representational and a metaphysical reading; and the related idea that metaphysical fundamentality might come in degrees. Perhaps a proponent of K-Superlativism might hold that her thesis is to be understood along a representational reading (e.g., that it is merely the claim that “the concept KNOWS does not have BELIEVES as a constituent”; Ichikawa, 2017, p. 343, and see further references therein). Alternatively, a proponent of K-superlativism might also hold that the degree of fundamentality at play in the common view, according to which belief is the only cognitive element that is indispensable in causal explanation of action, is more fundamental than the claim that knowledge is the most general factive mental state. I leave a proper investigation of these options to a proponent of K-superlativism on another occasion.
- ⁸ See Steglich-Petersen (2005) and Reed (2005) for objections along similar lines.
- ⁹ See Williamson (2000, p. 64): “When one works through enough examples of this kind [presented below], it becomes increasingly plausible that knowing can figure ineliminably in causal explanations. It is causally efficacious in its own right if any mental state is [...]”
- ¹⁰ See, however, Rose (2015) and McGlynn (2017), for a challenge to the evidence from the empirical literature. See Dudley (2018) for discussion.
- ¹¹ Note also that it is somewhat surprising to refer only or even mainly to Dretskean approach as the causal theory of knowledge, since Dretske is better known for contributing to the relevant alternatives theory of knowledge. More standard examples of the causal theory of knowledge would be Goldman (1967). Yet, Perner writes: “The causal theory of knowledge (Dretske, 1981; Dretske & Eng, 1984) stipulates that a true representation has to be caused by the fact to be known” (Perner, 1993, p. 304, note 7.1) and seems to endorse the idea that this theory captures our commonsense concept of “knowledge” or commonsense theory of knowledge cf. Perner (1993, p. 146). Thanks to a reviewer for drawing my attention to this point.
- ¹² Note also Nagel’s observation with respect to minimalist programmes: “It is noteworthy that if even minimalist programmes end up with something akin to belief contents (“belief-like states”), they derive these contents from some type of factive condition, a condition such as encountering or witnessing, of a type that can only link an agent to a fact or real state of affairs” (Nagel, 2017, p. 540).
- ¹³ See also Kaplan (2003) for a similar line of objection. The difference between Magnus and Cohen’s objection and Kaplan’s proposal boils down to Kaplan focusing on the state of confidence that is resilient to some further set of propositions rather than the state of belief-that-will-survive-scrutiny. However, as far as I can see, it has the same pitfalls as Magnus and Cohen’s line of objection.
- ¹⁴ See also Nagel (2013) for a suggestion along similar lines and McGlynn (2014, pp. 190–193) for a more detailed similar response to Magnus and Cohen.
- ¹⁵ See Molyneux (2007, p. 270): “If the burglar’s knowing that *p* is what explains his actions, then it ought to explain similar actions of similar burglars in similar situations; its failure to do so is a disturbing indicator.”
- ¹⁶ Let me add one last consideration here. It might be possible to convince some opponents by enlarging slightly the kind of cases that we are focusing on here. In particular, it might be helpful to focus on accomplishments (e.g., finding the diamond) rather than mere processes (e.g., ransacking the house). It might be claimed that the pre-theoretical judgement that nothing less than knowledge will suffice to explain the relevant action in accomplishments is even stronger than the judgement in persistence cases. For the truth and a modally robust character of the relevant state of mind seem to be even more essential for accomplishments. See also McGlynn (2014, pp. 190–193) for a similar suggestion. In this context, Williamson’s case of someone digging up treasure seems to be more relevant (cf. Williamson, 2000, pp. 61–62). Thanks to Edgar Phillips for discussion on this point.
- ¹⁷ In what follows I omit the tedious reference to the “that *p*” clause in describing the relevant emotions, but it is important to remember that the focus here is on emotions having promotional contents.
- ¹⁸ I would like to thank an anonymous referee for this journal for urging me to consider these questions in a more sustained way.
- ¹⁹ The above suggestion (i.e., the suggestion that knowledge itself explains some action better than belief because it is a more fundamental element in our psychology) involves a non-trivial gap, though. In suggesting that knowledge explains (some) action better than belief since knowledge attributions are more fundamental elements of our theory of mind (compared to belief attributions), I rely on the assumption that we can conclude (the metaphysics of) mental states (e.g., knowledge itself being psychologically more fundamental and being non-dependent on belief) from observations about knowledge attributions and more generally from how our theory of mind works (how our brains attribute mental states to others). I acknowledge that this gap between the observations about mental state attribution and mental states

themselves cannot be closed in a theoretically cost less and straightforward way (see McGlynn, 2014, pp. 185–186 for a worry along these lines for the account in Nagel, 2013 in particular). Moreover, the studies that I referred to above do not seem to concern knowledge itself but are rather focused on knowledge attributions. Yet, I also think that it is not unreasonable to endorse some bridge principle connecting mental state attributions to claims about the states themselves. Here are two possible and potentially compatible ways to motivate such bridge principles. First, one might think that our theory of mind is commonly reflected in our natural language semantics of mental state ascriptions. It would be surprising if it were not. It is a common assumption that our ordinary language use and semantics of mental state ascriptions, in particular, give us at least *prima facie* ingredients for our philosophical theories of mental states themselves. It would seem that it is a virtue of a philosophical theory of some mental state if it can accommodate our folk understanding of this state, which is arguably expressed in our ordinary language. If so, and if ordinary language mental state ascriptions commonly express our mental state attributions (theory of mind), then we have here a fallible yet a plausible bridge principle connecting theory of mind to metaphysics of states (assuming that philosophical theories of mental states are in the business of specifying metaphysics of mental states). Second, one might endorse a principle recently defended by Adam Bricker (Bricker, 2021) that connects neurocognitive mechanisms involved in the production of (perceptual in particular) knowledge attributions to the content of judgments that attribute knowledge and thus to the metaphysics of the knowledge states. According to this very interesting proposal, called “Neurocognitive Parity,” “[t]he contents of judgement J about knowledge reflect the structure of knowledge only if the mechanics of the neural and cognitive processes responsible for J also reflect the structure of knowledge.” (Bricker, 2021, p. 11). In short, the idea here is that if our intuitive judgements about cases of knowledge (e.g., Gettier cases or what have you) are evidence for certain claims about metaphysics of knowledge (e.g., factivity, safety, but also being a simple non-composed state), then the processes that produce knowledge attributions (e.g., self-perspective inhibition, and others) are also evidence with respect to claims about metaphysics of knowledge. In short, according to Bricker: “[t]he point of neurocognitive parity is simply that when the content of judgements about knowledge can evidence the structure of knowledge, so too can the neurocognitive mechanics responsible for that content.” (Bricker, 2021, p. 12). Bricker also thinks that empirical evidence from cognitive neuroscience suggests that processes that produce knowledge attribution are such that they treat knowledge as a mental state and also thinks that “[i]f the mechanics of some class of judgements (a) reflect the structure of knowledge and (b) treat knowledge as if it is a mental state, then knowledge is a mental state for the cases that are the subject of those judgements.” (Bricker, 2021, p. 13). Naturally, Bricker then concludes that knowledge is a genuine, non-composite mental state. Regardless of the merits of Bricker's own very intriguing and in my view promising argument, I would like to suggest that endorsing the Neurocognitive Parity principle provides us with one plausible way of responding to the gap challenge, the challenge of explaining how and why observations about knowledge attributions can tell us anything about metaphysics of knowledge itself. I do not pretend here to have dealt with the gap problem exhaustively. This task will be left for another occasion. Yet, the above two suggestions might constitute the first steps towards such treatment. Many thanks to an anonymous referee for this journal for drawing my attention to the need to consider this issue more attentively.

²⁰ One might reply to the above line of thought by noticing that not everyone accepts the view that the above-mentioned states, the states of being annoyed that p, being upset that p, being happy that p, regretting that p, and so on are factive states. If none of these is factive (in the relevant sense), then the conclusion that none of these are genuine mental states does not follow from the above example, and thus the above problem with Smith's line of criticism against K-Superlativism might be avoided. And as it happens, some have indeed rejected the view that (all of) these states are factive, for example, Fantl (2015) rejects the view that being happy that p is factive. I would like to make two quick points in reply to this response. First, while I acknowledge that this move could help to avoid my objection, it would also involve some non-trivial theoretical costs. In particular, one would need to appeal to a massive error-theory that could explain why contrary our linguistic and common sense judgements *none* of the relevant states is factive. Such a theory would need to explain why we are so radically wrong about these commonly accepted judgements that seem to constitute evidence for factivity of these states. While there exist attempts of rejecting factivity of some of these states, for example, Fant on happiness, I am not aware of any such overarching attempt of explaining away all of the data concerning all of these *prima facie* factive states. I am not suggesting here that it cannot be done. I am merely registering that we need to acknowledge the theoretical costs that such a possible error theory would involve. Second, it is not clear to me that it is possible to provide such an overarching error theory without entailing that knowledge itself and other canonical factive states are not factive. The worry here would be that there is a serious risk that by arguing against factivity of factive emotions, one also ends up arguing against factivity of knowledge and other factive states. However, denying factivity of knowledge seems to involve a whole new level of theoretical costs that I am not sure we are ready to take on board. Again, I am not saying it is impossible to have good arguments against factivity of knowledge, but the cost of denying factivity of knowledge seems even more massive. Note also that the suggestion here is not only about knowledge. Think about “forgetting that p,” “explaining that p,” “showing that p” and so on. Consider the fishmonger's case again. Our subject travels overseas, and while on the trip, begins to forget what color the shopfront was. It seems perfectly fine to

describe his situation as “Subject is forgetting that the shopfront is blue.” But now suppose that the owner paints the shopfront red at that moment. If forgetting were not factive we should be able to still say the same thing: “Subject is forgetting that the shopfront is blue.” But it does seem odd to say such a thing when we know that the shopfront is not blue now. Note also that if forgetting is factive then the initial fishmonger case would also constitute a parallel argument for the claim that forgetting cannot be a mental attitude. Yet neither the rejection of factivity of forgetting nor the conclusion that it is not a mental attitude appears to be acceptable. Similar considerations apply to “S explains that p” and “S shows that p.” Thus, arguably, the oddness of the fishmonger’s shopfront case is better dealt with without rejecting factivity of neither factive emotions nor other attitudes or states. I suggest then that considerations about the fishmonger’s case (and similar cases) do not lead us to reject the idea that knowledge is a mental state. If we accept they do, then given some arguably plausible assumptions about factivity of some mental attitudes, we should also conclude that neither factive emotions nor other factive attitudes, for example, forgetting, are mental. My positive suggestion is that a more general solution to the problem of the fishmonger’s case would be optimal since the problem is not only a problem for the view that knowledge is a mental state. It is a problem for all the views according to which factive emotions and other factive attitudes are mental. Thanks to a reviewer for this journal for bringing this potential response to my attention.

REFERENCES

- Apperly, I. (2011). *Mindreaders: The cognitive basis of ‘theory of mind’*. New York, NY: Psychology Press.
- Benton, M. A. (2014). Knowledge norms. *Internet Encyclopedia of Philosophy*.
- Ben-Ze’ev, A. (2001). *The subtlety of emotions. A Bradford book*. Cambridge, MA: The MIT Press.
- Bird, A. (2007). Justified judging. *Philosophy and Phenomenological Research*, 74(1), 81–110.
- Black, T. (2003). Contextualism in epistemology. *Internet Encyclopedia of Philosophy*.
- Bricker, A. M. (2020). The neural and cognitive mechanisms of knowledge attribution: An EEG study. *Cognition*, 203, 104412. <https://doi.org/10.1016/j.cognition.2020.104412>
- Bricker, A. M. (2021). Knowledge is a mental state (at least sometimes). *Philosophical Studies*, 1–21. <https://doi.org/10.1007/s11098-021-01714-0>
- Burge, T. (1979). Individualism and the mental. *Midwest Studies in Philosophy*, 4(1), 73–122.
- Butterfill, S. A. (2013). What does knowledge explain? Commentary on Jennifer Nagel, ‘Knowledge as a mental state’. *Oxford Studies in Epistemology*, 4, 309.
- Butterfill, S., & Apperly, I. (2013). How to construct a minimal theory of mind. *Mind & Language*, 28, 606–637.
- Carruthers, P. (2011). *The opacity of mind: An integrative theory of self-knowledge*. Oxford: Oxford University Press.
- Davidson, D. (1963). Actions, reasons, and causes. *Journal of Philosophy*, 60(23), 685–700.
- de Sousa, R. B. (1987). *The rationality of emotion*. Cambridge, MA: MIT Press.
- Dennett, D. C. (1987). *The intentional stance*. Cambridge, MA: MIT Press.
- Dietz, C. (2018). Reasons and Factive emotions. *Philosophical Studies*, 175(7), 1681–1691.
- Dretske, F. (1981). *Knowledge and the flow of information*. Oxford: Wiley-Blackwell.
- Dretske, F., & Enç, B. (1984). Causal theories of knowledge. *Midwest Studies in Philosophy*, 9, 517–528. <https://doi.org/10.1111/j.1475-4975.1984.tb00075.x>
- Dudley, R. (2018). Young children’s conceptions of knowledge. *Philosophy Compass*, 13(6), e12494.
- Fantl, J. (2015). What Is It to Be Happy That P?. *Ergo: An Open Access Journal of Philosophy*, 2(12), 267–297.
- Fodor, J. A. (1981). Three cheers for propositional attitudes. In *Representations. Philosophical essays on the foundations of cognitive science* (pp. 100–123). Cambridge, MA: MIT Press.
- Fratantonio, G. (2018). Armchair access and imagination. *Dialectica*, 72(4), 525–547.
- Goldman, A. I. (1967). A causal theory of knowing. *The Journal of Philosophy*, 64(12), 357–372. <https://doi.org/10.2307/2024268>
- Gordon, R. M. (1969). Emotions and knowledge. *Journal of Philosophy*, 66, 408–413.
- Gordon, R. M. (1987). *The structure of emotions: Investigations in cognitive philosophy*. Cambridge, England: Cambridge University Press.
- Hawthorne, J., & Stanley, J. (2008). Knowledge and action. *Journal of Philosophy*, 105(10), 571–590.
- Hyman, J. (2014). The most general factive stative attitude. *Analysis*, 74(4), 561–565.
- Ichikawa, J. J. (2014). Justification is potential knowledge. *Canadian Journal of Philosophy*, 44, 184–206.
- Ichikawa, J. J., & Steup, M. (2018). The analysis of knowledge. *The Stanford encyclopedia of philosophy* (Summer 2018 Edition), E. N. Zalta. Retrieved from <https://plato.stanford.edu/archives/sum2018/entries/knowledge-analysis/>
- Ichikawa, J. J. (2017). Basic knowledge first. *Episteme*, 14(3), 343–361.
- Ichikawa, J., & Jenkins, C. S. I. (2017). On putting knowledge ‘first’. In J. A. Carter, E. C. Gordon, & B. Jarvis (Eds.), *Knowledge first: Approaches in epistemology and mind*. Oxford: Oxford University Press.
- Kaplan, M. (2003). Who cares what you know? *The Philosophical Quarterly*, 53(210), 105–116.

- Kelp, C. (2016). Justified belief: Knowledge first-style. *Philosophy and Phenomenological Research*, 93, 79–100.
- Kelp, C. (2018). *Good thinking: A knowledge-first virtue epistemology*. New York, NY: Routledge.
- Kipper, J. (2018). Acting on true belief. *Philosophical Studies*, 175(9), 2221–2237.
- Littlejohn, C. (2017). How and why knowledge is first. In A. Carter, E. Gordon, & B. Jarvis (Eds.), *Knowledge first*. Oxford: Oxford University Press.
- Logins, A. (2014). The problem of massive deception for justification norms of action. *Acta Analytica*, 29(4), 457–468.
- Logins, A. (2017). Common sense and evidence: Some neglected arguments in favour of E=K. *Theoria*, 83(2), 120–137.
- Magnus, P. D., & Cohen, J. (2003). Williamson on knowledge and psychological explanation. *Philosophical Studies*, 116(1), 37–52.
- McGinn, C. (1977). Charity, interpretation, and belief. *Journal of Philosophy*, 74(9), 521–535.
- McGlynn, A. (2014). *Knowledge first?*. Basingstoke, England: Palgrave Macmillan.
- McGlynn, A. (2017). Mindreading knowledge. In J. A. Carter, E. C. Gordon, & B. W. Jarvis (Eds.), *Knowledge first: Approaches in epistemology and mind* (pp. 72–94). Oxford: Oxford University Press.
- Molyneux, B. (2007). Primeness, internalism and explanatory generality. *Philosophical Studies*, 135(2), 255–277.
- Nagel, J. (2013). Knowledge as a mental state. *Oxford Studies in Epistemology*, 4, 275–310.
- Nagel, J. (2017). Factive and Nonfactive mental state attribution. *Mind & Language*, 32(5), 525–544.
- Noonan, H. W. (1993). Object-dependent thoughts: A case of superficial necessity but deep contingency? In J. Heil & A. R. Mele (Eds.), *Mental Causation*. Oxford: Oxford University Press.
- Nottelmann, N. (2011). Belief-desire explanation. *Philosophy Compass*, 6(12), 912–921.
- Perner, J. (1993). *Understanding the representational mind*. Cambridge, MA: MIT Press.
- Pettit, P. (1986). Broad-minded explanation and psychology. In P. Pettit & J. McDowell (Eds.), *Subject, thought and context*. Oxford: Oxford University Press.
- Phillips, J., Buckwalter, W., Cushman, F., Friedman, O., Martin, A., Turri, J., ... Knobe, J. (2021). Knowledge before belief. *Behavioral and Brain Sciences*, 44(e140), 1–37.
- Putnam, H. (1973). Meaning and reference. *Journal of Philosophy*, 70(19), 699–711.
- Reed, B. (2005). Accidentally Factive mental states. *Philosophy and Phenomenological Research*, 71(1), 134–142.
- Rose, D. (2015). Belief is prior to knowledge. *Episteme*, 12(3), 385–399.
- Ruffman, T. (2014). To believe or not believe: Children's theory of mind. *Developmental Review*, 34, 265–293.
- Rysiew, P. (2013). Is knowledge a non-composite mental state? *Oxford Studies in Epistemology*, 4, 333–343.
- Schwitzgebel, E. (2002). A phenomenal, dispositional account of belief. *Noûs*, 36(2), 249–275.
- Schwitzgebel, E. (2010). Acting contrary to our professed beliefs or the Gulf between occurrent judgment and dispositional belief. *Pacific Philosophical Quarterly*, 91(4), 531–553.
- Schwitzgebel, E. (2013). A dispositional approach to the attitudes: Thinking outside of the belief box. In N. Nottelmann (Ed.), *New essays on belief* (pp. 75–99). Basingstoke, England: Palgrave.
- Silva Jr, P. (2020). Knowledge-first theories of justification. Internet Encyclopedia of Philosophy.
- Silva, P. (2017). Knowing how to put knowledge first in the theory of justification. *Episteme*, 14(4), 393–412.
- Smith, M. (2017). The cost of treating knowledge as a mental state. In A. Carter, E. Gordon, & B. Jarvis (Eds.), *Knowledge first, approaches to epistemology and mind* (pp. 95–112). Oxford: Oxford University Press.
- Steglich-Petersen, A. (2005). Williamson on knowledge, action, and causation. *SATS: Northern European Journal of Philosophy*, 6(1), 15–28.
- Sutton, J. (2007). *Beyond justification*. Cambridge, MA: MIT Press.
- Turri, J. (2017). Knowledge attributions and behavioral predictions. *Cognitive Science*, 41, 2253–2261.
- Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford University Press.
- Williamson, T. (2011). Knowledge first epistemology. In S. Bernecker & D. Pritchard (Eds.), *The Routledge companion to epistemology* (pp. 208–218). London, England: Routledge.
- Wollheim, R. (1999). *On the emotions*. New Haven, CT: Yale University Press.

How to cite this article: Logins, A. (2021). Persistent burglars and knocks on doors: Causal indispensability of knowing vindicated. *European Journal of Philosophy*, 1–23. <https://doi.org/10.1111/ejop.12751>