



Knowledge as a (Non-factive) Mental State

Adam Michael Bricker¹

Received: 25 October 2022 / Accepted: 7 May 2023
© The Author(s) 2023

Abstract

The thesis that knowledge is a factive mental state plays a central role in knowledge-first epistemology, but accepting this thesis requires also accepting an unusually severe version of externalism about the mind. On this strong attitude externalism, whether *S* is in the mental state of knowledge can and often will rapidly change in virtue of changes in external states of reality with which *S* has no causal contact. It is commonly thought that this externalism requirement originates in the factivity of knowledge. However, despite a number of recent defenses of non-factive accounts of knowledge, epistemology has yet to consider whether a non-factive approach might produce a version of the mental state thesis that can avoid strong externalism. Here I do just that, exploring how three different proposals for weakening factivity might be adapted to theories of knowledge as a non-factive mental state. Contrary to what we might expect, however, none of these proposals are compatible with anything close to attitude internalism about knowledge—or even a substantially weaker externalism. All told, the widespread view that wraps up the severe externalism required for knowledge to be a mental state in factivity is mistaken. Knowledge’s external-world connection runs far deeper than the factivity constraint.

1 Introduction

One of the most widely discussed epistemological research programs of the past few decades, knowledge-first epistemology encompasses an assortment of different positions broadly characterized by eschewing the idea that knowledge is something like a kind of belief (see Carter et al., 2017; Greenough & Pritchard, 2009; McGlynn, 2014; Williamson, 2000). A central feature of knowledge-first epistemology is the thesis that knowledge is a mental state (for defenses, see Bricker, 2022a; Nagel, 2013; Pavese, 2021; Williamson, 2000). On this view, knowledge is not merely a composite of a mental state like belief with some additional non-mental properties like truth and whatever might distinguish true belief from knowledge.

✉ Adam Michael Bricker
adam.michael.bricker@gmail.com

¹ Turku Brain and Mind Center, University of Turku, Turku, Finland

Rather, knowledge is a mental state *in its own right*, analogous to other mental states like belief, desire, perceptual states, and memory. Going beyond the more widely accepted composite view, on which being in some mental state (i.e., belief) is merely necessary for knowledge, this knowledge-first approach maintains that there is some mental state for which being in that state is both necessary and sufficient¹ for knowledge.

Given that knowledge is factive—knowing that *p* entails the truth of *p*—then a kind of externalism about mental attitudes immediately follows from the proposal that knowledge is a mental state. As whether some representational content is true is not generally a fact about *S*'s brain or internal composition, but instead an external fact about *S*'s environment, whether *S* occupies the mental state of knowledge will be at least partially determined by facts that are external to *S*. Williamson defends this consequence by noting that externalism about mental *content*, if not attitudes themselves, is already an established position in the philosophy of mind, seeking to include externalism about factive attitudes under the auspices of more general externalism about the mind (2000, chapter 2). As characterized by Fricker, “Williamson’s suggestion seems to be that, once one has taken one dose of externalism, as it were, one can have no grounds for qualms about accepting another” (2009, p. 53). Objecting to this style of argument, Fricker points out that we can of course still accept more conventional varieties of externalism about mental content while rejecting the kind of attitude externalism engendered by taking knowledge to be a mental state (2009). Taking this a step further, Smith points out that the kind of externalism required for knowledge to be a mental state is “far more radical than any version previously considered” (2017, p. 95). As we’ll discuss below, while a number of accounts have been proposed that involve some variety of externalism for mental content and attitudes, none of them match the degree of externalism required of knowledge as a mental state.

I agree with Smith’s assessment (2017) that, while this isn’t itself a conclusive reason to abandon the mental state thesis, it is a significant cost that defenders of the view must appreciate. There is something highly counterintuitive about the idea that what type of mental state I’m in can and often does instantaneously change in virtue of changes in distant states of reality with which I have no causal contact, and no other variety of externalism about the mind seems to make so strong a claim. Moreover, such a suggestion stands at odds with the widely held view that mental types are causally relevant—Roughly, if the type of state *S* is in can change in virtue of nothing more than aspects of the world that she’s causally isolated from, it’s unclear how such a mental type can be relevant in the causal explanation of *S*'s behavior. All this places a hard limit on the appeal of the mental state thesis, excluding all those uncomfortable with committing to such a severe version of attitude externalism.

Such a hard limit is unfortunate, however, because the mental state thesis is itself highly intuitive, with a large body of evidence suggesting that it is entirely natural to view knowledge as a mental state in its own right: The theory of mind systems we rely on to track what others know represent knowledge like a mental state

¹ Or, at least sufficient (i.e., Bricker 2022a).

(Bricker, 2020; Phillips et al., 2021; Westra & Nagel, 2021); the psychologists who study knowledge attribution generally think of it as a mental state (Nagel, 2013); and even throughout most of the history of philosophy, knowledge was widely regarded as a basic mode of cognition, not a type of belief (Antognazza, 2020). Accordingly, we have a clear interest in understanding whether epistemology might separate out the intuitively appealing mental state thesis from its intuitively troubling brand of attitude externalism.

As it is common to view the externalist consequence as originating in factivity (e.g., Nagel, 2013, p. 289; Smith, 2017, p. 96; Williamson, 2000, p. 49), an intriguing strategy presents itself: weaken the factivity of knowledge. Much like the mental state thesis, non-factivity for knowledge is particularly intuitive outside of contemporary epistemology, with experimental findings showing a substantial willingness to attribute knowledge of falsehoods among both non-philosopher academics and laypeople, even when controlling for non-literal uses of “knows” (Phillips & Bricker, 2022; Starmans & Friedman, 2020). Moreover, perhaps more importantly, a number of philosophers have recently defended non-factive approaches to knowledge (Bricker, 2018; Bricker, 2022b; Buckwalter & Turri, 2020a; Buckwalter & Turri, 2020b; Kusch, 2009, §5; see also Niiniluoto, 1999; Stjernberg, 2009), and while these aren’t themselves presented as knowledge-first accounts, they can easily be adapted to theories of knowledge as a less-than-factive mental state. If epistemology is correct in assuming that externalism for the mental state thesis originates in factivity, then such accounts of knowledge as a non-factive mental state should allow us to circumvent attitude externalism—or at least weaken it in proportion with our weakening of factivity. However, contrary to our expectations, this does not prove to be the case. No non-factive approach to knowledge appreciably moderates the externalist consequences associated with the mental state thesis; compatibility with attitude internalism is just as out of reach as on the factive approach.

Accordingly, these considerations highlight a critical and easily overlooked feature of knowledge-first epistemology: The associated attitude externalism cannot be solely—or even primarily—attributed to factivity. It is instead a much deeper consequence of taking knowledge to be a mental state. As we’ll find, even a radically non-factive approach, a serious departure from how epistemology normally thinks about the kinds of states that can count as knowledge, broadly retains the same externalist consequences as the factive mental state thesis. While there are options available for severely weakening factivity, the strength of externalism just doesn’t scale with the strength of the knowledge-truth relationship. Because knowledge’s external world connection goes well beyond what is true and false, it is difficult to imagine how we could describe a class of mental states that even remotely resembles knowledge without committing ourselves to a substantial version of externalism about mental attitudes. All told, it does not appear that strong attitude externalism might be avoided.

In surveying the prospects of taking knowledge to be a non-factive mental state, this paper will begin with an overview of Smith’s previous work (2017) on the kind of especially severe externalism required of taking it to be a *factive* mental state (§2). Next, I’ll consider three options for weakening factivity recently suggested in epistemology: stringent truthlikeness (§3), representational adequacy (§4), and

institutional/social sanction (§5). In each case, I'll discuss how to adapt extant theories to a knowledge-first framework, the extent to which the approach reduces externalism-related costs, and what other epistemological costs might be associated with the approach. After this, I'll close with a few general observations, most importantly that the attitude externalism here cannot be thought of as just a consequence of factivity (§6).

Finally, before moving on, note that my purpose with this paper is not to endorse the mental state thesis, any particular account of knowledge as a non-factive mental state, nor even the non-factive approach generally. Instead, I want to provide a preliminary exploration of how these two trends in epistemology might be integrated into a single approach, along with the most important lessons to be gained from this exploration. Nevertheless, I do want to clarify that here I will set aside any general worries that weakening factivity is inherently so prohibitively costly that it precludes serious consideration. Here we'll discuss specific theoretical costs of specific non-factive approaches, but I don't think it's correct to consider non-factivity to itself be one. Non-factivity isn't a cost; the costs of non-factivity are costs. Far from the caricature on which weakening factivity means dropping the truth requirement without offering a replacement (see, e.g., Stjernberg, 2009, p. 40), as we'll discuss, certain versions of non-factivity are much more conservative than might be assumed.

Conversely, the kind of externalism required in taking knowledge to be a factive mental state is far more radical than is often acknowledged, inviting a disconcerting sort of instantaneous switching between mental attitudes even in the absence of causal contact with the relevant parts of external reality. For those otherwise partial to the mental state thesis, but who find attitude externalism (or perhaps externalism about the mind generally) too costly a price, a conservative factivity replacement might then serve as an attractive solution. Particularly given the tendency of both proponents and critics of knowledge-first epistemology to identify this externalist demand as a product of factivity itself, we have a clear interest in understanding whether we might do any differently by taking knowledge to be a non-factive mental state.

2 Knowledge as a Factive Mental State: The Cost of Externalism

Let's begin by discussing the kind of externalism that results from taking knowledge to be a *factive* mental state. Both here and in the following sections, we'll need a theoretically rich system for comparing the mental state thesis with varieties of externalism developed in the philosophy of mind. For this, we'll follow the taxonomy previously introduced by Martin Smith (2017), which contrasts this externalism with four other externalist theories: natural kind externalism, social externalism, demonstrative externalism, and disjunctivism (see Table 1).² In doing so, we observe that the kind of externalism required by the thesis that knowledge is a

² Note that, while most of these were developed as theories of semantic content, I'll follow Smith (2017) in assuming unproblematic translation to mental content.

Table 1 Comparison of theories that entail a version of externalism about the mind (adapted from Smith 2017, Table 5.1; type column added)

Theory	Switching profile			
	Speed	Range	Frequency	Type
Natural kind externalism	Slow	Proximal	Rare	Content
Social externalism	Slow	Proximal	Common	Content
Demonstrative externalism	Fast	Proximal	Rare	Content
Disjunctivism	Fast	Proximal	Rare	Attitude
Knowledge as a factive mental state	Fast	Distal	Common	Attitude

factive mental state is far stronger than any of these other proposals. I concur with Smith that, while this may not itself be reason to abandon the thesis, it is a significant cost that must be acknowledged by knowledge-first epistemologists. Note that here I'll only provide a cursory overview of this material, and I would encourage the reader to seek out Smith (2017) for a more detailed discussion.

Central to Smith's taxonomy is the concept of "switching," defined as "a change in a person's mental states that is not mediated by any change in the person's internal physical state" (2017, p. 99). The possibility of switching is a hallmark of any externalist theory of mind, as changes in whatever external states of affairs determine the contents or type of a mental state inevitably mean changes in that state. Crucially, however, as Smith notes, not all externalist theories allow for the same kind of switching. To observe this, consider first natural kind externalism, the position that the contents of beliefs about natural kinds are partially determined by external facts about the natural kinds in one's environment. For example, the content of a belief about the natural kind water will depend in part on the physical composition of water— H_2O in our world. This allows for the kind of switching we famously see in Putnam's "Twin Earth" cases (1975): Consider an agent from earth, where water is H_2O , who is unknowingly transported to a twin Earth, on which a substance called "water" and displaying the same macro-level features of water has some other chemical composition XYZ. While the agent's old beliefs about water will initially be about H_2O , as this agent interacts with XYZ, the new beliefs about water will be about XYZ. Eventually, once the agent has spent enough time on twin Earth, the contents of beliefs like, "Water freezes at 0 °C" will no longer be about H_2O , but rather XYZ. As Smith notes (2017, §3), this switching is *slow*, requiring a significant amount of time to occur, *proximal*, requiring causal contact between the agent and the target of her beliefs, and *rare*, being more the stuff of thought experiments than everyday life.

Compare this to the switching profile of *social externalism*, the position that the contents of at least many beliefs—not just those about natural kinds—are partially determined by the agent's linguistic community. For example, the contents of a belief like "I have arthritis in my thigh" might have different contents for members of different linguistic communities, who use the word "arthritis" to mean different classes of ailments (Smith, 2017, p. 100; adapted from Burge, 1979, 1986). As with

natural kind externalism, this allows for slow, proximal switching when an agent moves between linguistic communities. Crucially, however, here switching is much more *common* than on natural kind externalism. It is the sort of thing that happens regularly in the actual world, as agents frequently move between linguistic communities (and the way terms are used evolve within those communities). For more, see Smith (2017, §4).

Finally, let's consider the switching profiles of two other externalist theories, *demonstrative externalism* and *disjunctivism* (see Smith, 2017, §§5–6). On demonstrative externalism (e.g., Evans, 1982; Kaplan, 1989), the contents of beliefs involving demonstratives are partially determined by the target of the demonstrative (e.g., *this* banana or *that* coffee cup). This means that, if the target of a demonstrative is replaced without the agent being aware—perhaps one banana is swapped for a second banana while she isn't looking—then her belief about “that banana” will have different contents, despite no change in her internal state. As with natural kind externalism, this switching is proximal and rare, requiring causal contact with the target of the demonstrative and not featuring in everyday life apart from occasional subterfuge (or, perhaps, remarkably lucky spontaneous object swapping). However, unlike anything discussed so far, this switching is quite *fast*. In as little time as S can lose and re-gain perceptual contact (and the object can be swapped out), the contents of her mental state can switch. We observe a similar switching profile with disjunctivism about perception, the thesis that perceiving and only seeming to perceive are two distinct kinds of mental states (see Soteriou, 2020). As this means that seeing an object and seeing an illusion perceptually indistinguishable from that object are different mental states, even when supported by identical internal states, switching can occur when an object of perception is replaced by an indistinguishable illusion without the awareness of the perceiving agent. Smith offers the example of a cube being vaporized and simultaneously replaced by a perfect holographic counterpart (2017, p. 105). As with demonstrative externalism, this kind of switching is fast, but proximal and rare. It requires causal contact with the represented object to occur, and it would be extremely unusual in ordinary life. But notice now that, unlike the previous examples, this is not merely a change in mental *content*. Rather, this is an example of *attitude* externalism, with switching occurring for the type of mental attitude itself.

Now let's examine the switching profile of the thesis that knowledge is a factive mental state. Quite unlike anything we observed above, this thesis allows for attitude switching that is fast *and* common, a new combination. But more radically, it also allows for *distal* switching, changes in mental states without any causal contact with the objects of those states. Consider Nagel's example of the knowledge state—on April 15th, 1865—that Abraham Lincoln is the president of the United States (2013, p. 289).³ At the point in the morning when Lincoln succumbed to the injuries he sustained the night before (the officially reported time is 7:22 am), all those people who had previously known that Lincoln was president no longer have that mental state. As we have seen with other cases, this switch occurs instantaneously and

³ Nagel adapts the example from Williamson (2000, p. 23).

outside their awareness. However, unlike anything we've seen before, this switching occurs distally, requiring no causal contact with Lincoln to occur. The moment it is no longer true that Lincoln is president, millions of agents worldwide cease to be in the state of knowledge. Observe too that this is not merely the switching of mental content, but of the type of attitude itself. Moreover, note that the switching in this example is not due to any quirk of the Lincoln case, but will instead occur quite frequently, whenever a change in the state of reality voids a knowledge state. In short, the thesis that knowledge is a factive mental state doesn't entail a familiar form of content externalism—like natural kind or social externalism—but an abnormally strong variety of attitude externalism on which fast, distal switching is a common occurrence.

Many in epistemology find this kind of externalism to be particularly troubling. As put by McGlynn, “Williamson’s thesis has some deeply counterintuitive consequences—for example, that what happens at other locations can instantaneously change which mental state I am in—and many (myself included) will regard this as reason enough to be wary of accepting the view” (2014, pp. 193–194). Similarly, Fricker notes that “even for one who accepts externalism about content, there remain strong grounds to resist Williamson’s proposed extension of it to allow factive attitudes,” maintaining that while “externalism about mental content is counter-intuitive... externalism about factive attitudes would be even more so” (2009, p. 32 & p. 55). I agree. There is something especially disconcerting about the suggestion that fast, distal attitude switching is a common feature of everyday knowledge states. This is certainly quite unlike anything we observe for most other classes of mental states (credence, desire, pain, hope, fear, etc.), even granting more conventional accounts of externalism. We cannot go from hoping to fearing or experiencing pain to experiencing pleasure in virtue of far-off external factors, with which we have no causal contact. Moreover, there doesn't seem to be any other plausible context where this happens for belief or other cognate representational states. It's not clear that we ever go from believing to not believing in virtue of causally distal changes in the world. And even advocates of disjunctivism about perception, itself a controversial and rather counterintuitive proposal, don't go so far as to think attitude switching can occur in absence of proximal causal contact. All told, this counterintuitive variety of attitude externalism is reasonably understood as a barrier to entry for the view that knowledge is a mental state—perhaps even *the* barrier.

At this point, a proponent of the mental state thesis might interject that the concerns here are a bit overblown. Sure, it might be counterintuitive and unusual, but do we have any more substantive reason to resist distal attitude switching for at least some types of mental states? While I can't answer this question definitively, I do think that there is a looming problem here that goes well beyond simply being counterintuitive and different. It's widely assumed in philosophy that mental states are causally efficacious. Roughly, mental states play meaningful, non-redundant roles in the causal explanation of behavior, and the type of mental state is *causally relevant*—Ss in mental state type T will exhibit different patterns of behavior than Ss

not in T, with mental type playing a causal role in these behavioral differences.⁴ For example, whether S believes that it's sunny outside or simply desires, without believing, that it's sunny outside will *ceteris paribus* result in very different patterns of behavior in S. Crucially, this causal relevance of mental types isn't merely acknowledged by Williamson, but serves a central role in his case for the factive mental state thesis. Knowledge states play distinct causal roles from (especially) mere belief states, particularly over extended time courses, motivating the claim that knowledge is a *sui generis* mental state (2000, pp. 61–4 & pp. 75–92): A burglar who knows that a diamond is in a house will spend more time searching the house than a burglar who merely believes it; someone knocking on your door is much more likely to take offense at a non-response if they know that you're at home, as opposed to merely believing it, and so on.

The problem, however, is that the prospect of distal attitude switching, for knowledge or any other putative mental type, raises doubts about how causally relevant that type might have been in the first place. If S has no causal contact with the change in reality that precipitated the switch, then presumably S's behaviors can and will go on in quite the same way pre- and post-switch, for however long we play out the scene in either direction. It's only when causal contact with the relevant aspect of reality is reestablished—and/or there is some change in what S believes, desires, etc.—that we might expect S's behaviors to be meaningfully different. Crucially, going from being in a knowledge state to not being in a knowledge state seems to be causally irrelevant in all this. Say, for example, the burglar initially knows that the diamond is in the house. But at a point at which he lacks causal contact with the diamond, it's removed from the premises. Now the burglar no longer knows, but it's difficult to imagine how this alleged shift in mental type can be causally relevant to the burglar's behavior, since it only occurs in virtue of an external change that he has no causal contact with. The burglar will continue on just as if he knew, and we can only expect his behavior to meaningfully change once he reestablishes causal contact with the diamond, perhaps gaining some new evidence that it's no longer in the house,⁵ or simply ceases to *believe* that it's there. In short, we find that the prospect of distal attitude switching for knowledge cuts against the suggestion that knowledge, as a mental type, is causally relevant. Other mental types may be causally relevant. And attitude-individuating aspects of external reality that S is in causal contact with might even be causally relevant. But to the extent that it depends on causally distal aspects of reality, being in a knowledge state itself seems to have no causal relevance for S's behaviors.

Notice that this isn't an issue for all externalist attitude switching. On the proximally constrained attitude switching permitted on disjunctivism, the behavior of S₁ who sees object O versus the behavior of S₂ who only experiences the visually indistinguishable illusion of O will likely diverge over time, largely in virtue of the fact that S₁ is in causal contact with O while S₂ is in causal contact with

⁴ The causal relevance of mental types featured especially heavily in discussions of mental causation (see, e.g., McLaughlin 1989; Burge 1993; Ludwig 1994).

⁵ E.g., he learns that a member of his own crew pocketed the diamond and slipped away.

an illusion (Williamson discusses a similar example of seeing water vs. seeing a mirage; 2000, p. 75). And, moreover, notice too that none of this is a problem for the majority of epistemologists who reject the mental state thesis outright, as they don't think that knowledge is causally efficacious in the first place. Further still, if this were merely a rare occurrence, perhaps it could be written off as insufficiently informative. Mental states, especially when viewed in isolation, don't generally entail any particular behaviors. The causal relevance of mental types is best expressed in terms of patterns of and correlations with behavior that usually occur given some type of mental state, so exceptional cases may not be enough to undermine causal relevance. However, because the factive mental state thesis opens the door to distal attitude switching as an ordinary feature of mental states, it sits uneasy with the widespread view of mental types as causally relevant. To be clear, I don't want to give the impression that I'm the first to notice this sort of tension within factive mental states. While not quite the familiar "action at a distance" charge (e.g., Williamson, 2000, p. 61), which also calls into question the causal relevance of proximally constrained attitude externalism, there is certainly a shared worry that, as put by Nagel, "the distal vulnerability of factive states may make them seem unfit to serve in causal explanations of behavior" (2013, p. 289). Moreover, I don't mean to suggest that this is an insurmountable problem, only that there are deep tensions in this particular version of strong attitude externalism, which go well beyond simple intuitive discomfort. Again, this is the sort of thing that can serve as a barrier to entry for the mental state thesis.

Moving on, I want to observe that this strong attitude externalism is commonly identified as a specific product of factivity. Unsurprisingly, this idea seems to have originated with Williamson, proliferating subsequent discussions without much additional consideration. For example, Williamson states, "That knowing is a mental state is inconsistent with internalism... for the truth of what one knows may involve the external environment" (2000, p. 49; see also Williamson, 1995, §IV). Smith embraces a similar framing, contending that this externalism arises because "knowledge, unlike belief, desire, fear, hope, intention, etc., is *factive*" (2017, p. 96; original emphasis). Fricker likewise identifies the "factivity" of knowledge as the key reason why it's surprising to identify it as a mental state (2009, p. 31), and Nagel, too, views the above-mentioned "distal vulnerability" as a feature specific to "factive states" (2013, p. 289). All this, of course, seems perfectly reasonable. Go-to examples of switching—like the Lincoln case—rely on true-to-false changes in the state of reality, and it's not immediately obvious how comparable switching might occur otherwise. This then suggests an alternative approach for those disquieted by the strong externalism of fast, common, distal attitude switching: weaken the factivity constraint so that knowledge is a kind of non-factive mental state. While certainly a minority position in epistemology, a number of notable philosophers have expressed support for non-factive accounts of knowledge, including Ilkka Niiniluoto (1999), Martin Kusch (2009), and John Turri (Buckwalter & Turri, 2020a/b). In the following sections, I'll consider how these different approaches might be adapted into accounts of knowledge as a less-than-factive mental state, laying the groundwork for a non-factive knowledge-first epistemology. However, contrary to what might be expected, we'll find that no

extant technique for weakening factivity even comes close to underwriting an internalism-friendly version of knowledge as a mental state. The roots of externalism go far deeper than factivity.

3 Knowledge as a Factiveoid Mental State

3.1 Overview

Likely the most straightforward approach to weakening factivity is just to say that knowledge entails *approximate truth*. Rather than taking S to know that p only if p is true, S knows that p only if p is at least very close to being true. This approach is naturally formulated as a *truthlikeness condition* on knowledge. Taking truthlikeness to be some scalar measure of a proposition's closeness to truth, with truthlikeness 1 equal to strict truth, we might say that knowledge that p requires some very high degree of truthlikeness less than 1. To use terminology from Bricker (2018, §4.0), knowledge is *factiveoid*, a kind of state that is very close to truth but not necessarily strictly true.

A number of reasons have been offered for preferring a truthlikeness condition over a stricter truth condition, with perhaps the most compelling coming from the philosophy of science.⁶ In *Critical Scientific Realism*, Niiniluoto observes that “even the best claims of science are normally inexact, approximate, or idealized, i.e. they are not true in a strict sense” (1999, p. 84). This leads him to endorse an explicit truthlikeness condition in lieu of a truth condition for knowledge: S knows that p only if p is truthlike (1999, p. 84).⁷ Niiniluoto illustrates his rationale using the familiar example of Newtonian mechanics (1999, p. 84): While strictly false, the descriptions of the world provided by Newtonian mechanics are highly truthlike, and so should count as knowledge by those who believe them to be true (or at least believe them to be highly truthlike). In a reoccurring theme for arguments against factivity, there is a potential skeptical threat if we set the requirement for knowledge at strict truth. A good deal of putative scientific knowledge—perhaps even scientific knowledge as a category—appears threatened.

As already indicated, neither the truthlikeness proposal, nor any of the proposals that we'll discuss below, were introduced as knowledge-first theories. On the contrary, every account we'll consider here employs at least some degree of belief-first formulation. Nevertheless, we might easily adapt truthlikeness to the requisite type of knowledge-first account: Knowledge is a *factiveoid* mental state, a kind of mental state that entails truthlikeness, but not necessarily strict truth. Put another way, knowledge is a kind of mental state that can only be held to (at least) approximate

⁶ For alternative arguments supporting a similar conclusion, see Bricker 2018 and Bricker 2022b.

⁷ Things are complicated by Niiniluoto's conception of truthlikeness, which he doesn't take to always be entailed by truth. He does, however, favorably consider replacing this with a more straightforward notion of “approximate truth” (1999, p. 84). In either case, the resulting account is explicitly one on which strictly false approximate truths can be known.

truths. Other than not being strictly factive, this kind of mental state should be easily recognizable to the knowledge-first epistemologist. Advocates for this approach might even say that this is precisely the same mental state that knowledge-first epistemology has already identified. It's just gone unnoticed that many of the attitudes that constitute these states aren't actually held to strict truths.

Let's now consider the extent to which this approach might reduce externalism-associated costs, as well as what new costs might emerge in doing so.

3.2 Switching Profile

Unsurprisingly for such a structurally similar account, taking knowledge to be a factivoid mental state generates a switching profile very similar to its factive counterpart. For example, the above Lincoln case will still constitute the same kind of attitude switch. Once Lincoln dies, the proposition that he is president is no longer highly truthlike, and the knowledge state ceases to be a knowledge state. I would submit that most ordinary, human-scale changes in the state of reality are like this, large enough in magnitude that the shift to falsity will also be a shift to insufficient truthlikeness.⁸ In all such cases, the factivoid thesis has nothing new to say about switching. It will still be fast, common, and distal. Notice, however, that the factivoid account does require a greater magnitude change in the state of reality for a switch to occur. Small, true-to-highly-truthlike changes will not generally switch a factivoid mental state. Accordingly, switching on the factivoid account will be *less common*, however marginally, than switching on the factive account.

Before moving on, I want to offer something of a refinement on Smith's notion of frequency. Here, as well as in the following sections, we'll specifically discuss frequency of switching *relative to the total body of knowledge on a given account*. This is necessary because there are two distinct effects that can change the absolute frequency of switching: (i) changes in the conditions on which switching can occur and (ii) changes in the total number of cases that are classified as knowledge (and thereby potential switches). In the case of the factivoid mental state account, for instance, there could be an increase in the number of switches that occur in virtue of more mental states now being classified as knowledge, but that's not what we're interested in here. For our purposes, frequency itself isn't an interesting metric. It is only valuable to the extent that it tells us what switching is like. For example, while it might be true that switching is highly uncommon on a radically skeptical version of knowledge as a mental state, this itself isn't especially informative. We want a frequency metric that can capture differences in the nature of switching between different approaches, so it's important to exclude frequency changes driven by other sources. Accordingly, here we'll specifically focus on relative frequency, restricting the frequency metric to changes in the commonality of switching driven by changes in the conditions on which switching can occur (e.g., truth-to-false vs.

⁸ This is not a coincidence, but instead a feature of these truthlikeness accounts, which are built to capture the micro-level misrepresentations present in domains like scientific theorizing that might initially appear to produce representations that are perfectly veridical.

truthlike-to-untruthlike) and excluding those driven by changes in the overall volume of potential switches (i.e., more total knowledge vs. less total knowledge).

3.3 Epistemological Costs

By design, truthlikeness accounts are extensionally quite close to their factive counterparts. They maintain not that we consider conspicuously false mental content to constitute knowledge, but rather that much of the content we already consider to be known is in fact, however slightly, false. Nevertheless, depending on one's intuitions about knowledge and how one chooses to fill in the details, adopting the factivoid framework may come at the cost of some minor degree of epistemological revisionism. For many, it's likely counterintuitive that when approximately-true-yet-false content is fully accessible, as in quantitative measurement, this content can still constitute knowledge. For example, when S measures the radius of a coffee table as 84 cm, when it is in fact 84.61 cm, it may seem counterintuitive to say that S knows that the radius is 84 cm (although see Buckwalter & Turri, 2020a, 2020b). Or, in the kinds of cases that motivate Niiniluoto, many might be hesitant to grant that—given S's full access to how a scientific theory's descriptions misrepresent the world—these descriptions could still constitute knowledge for S. The advocate of the factivoid approach might simply accept this intuitive disagreement as a cost, or instead restrict the kind of false content that can be known to cases where the approximation isn't fully accessible (see Bricker, 2022b, §3 for more).

Additionally, observe that, in requiring a very high degree of truthlikeness for knowledge, the factivoid approach retains the more general intuition that knowledge and truth are closely related. Even among advocates of certain non-factive approaches to knowledge, it is recognized that there is a “conceptual connection between knowledge and truth” (Buckwalter & Turri, 2020a, p. 97; see also Bricker, 2022b, §4), and while the explicit requirement that knowledge be very close to truth trivially satisfies this, we'll shortly find that things aren't always this easy.

4 Knowledge as a Representationally Adequate Mental State

4.1 Overview

The representational adequacy approach to weakening factivity, notably advocated by Buckwalter and Turri (2020a, 2020b), shares many similarities with the truthlikeness approach discussed above. Not only is it motivated by similar skeptical concerns, but its central organizing principle is also that approximately true representations can constitute knowledge, even when they are strictly false. Nevertheless, there are substantial differences between the two approaches, perhaps most notably in how the level of approximate truth sufficient for knowledge is determined. On truthlikeness, this is a matter of meeting some truthlikeness threshold, similar to how some epistemologists might think about having a sufficient degree of justification for knowledge or sufficient degree of credence for belief. However, on the

Buckwalter-Turri account, distance to truth is secondary to the representational demands of an individual belief. Whatever degree of truth a belief needs to have to do the things we need it to, that's the level of truth required for that belief to constitute knowledge. While regrettably a bit vague on the general characterization of this representational adequacy, Buckwalter and Turri do offer one specific way in which a representation could be true enough for knowledge: It could "serve our purposes well enough to facilitate action and help us to achieve our goals in a particular circumstance" (2020a, p. 97). They illustrate this with the example of a belief that pi equals 3.14 (2020a, p. 97), contending that while this could be representationally adequate for the purposes of a grade school classroom—and therefore known—it would not be known in more demanding scientific applications, like the design of a global positioning system. In short, the degree of truth required for knowledge is on something of a sliding scale, varying, at a minimum, with the accuracy demands for the goal-oriented action guided by the representation in question.

Buckwalter and Turri offer a familiar anti-skeptical argument for their representational adequacy account (2020a). As many of our ordinary beliefs are only approximately true, but strictly false, considering these beliefs to be known requires that we replace factivity with a weaker constraint on knowledge. In this particular case, rather than appealing to scientific knowledge, Buckwalter and Turri utilize what might be called "naïve approximations," which rely on taking apparent approximations in ordinary language at face value (for a critical discussion, see Bricker, 2022b, §2). These include beliefs like "a mile equals 1.6 km," "the current temperature is 75° F," and "the current time is 9:03" (Buckwalter & Turri, 2020a, p. 94). While unfortunately Buckwalter and Turri don't spend much time discussing their choice to relativize truth requirements to the demands of individual contexts, they have subsequently provided some empirical evidence that indicates—although certainly not conclusively—that ordinary knowledge attributions are sometimes sensitive to representational adequacy (2020b).

As with truthlikeness, the Buckwalter-Turri account isn't presented as a knowledge-first theory. Nevertheless, we again can easily specify a knowledge-first version of this approach to weakening factivity: Knowledge is a representationally adequate mental state, a kind of mental state that sufficiently matches reality to, among other things, guide successful goal-directed action in some particular context. Fitting nicely with knowledge-first epistemology's emphasis on the causal efficacy of knowledge (discussed above), we might say that knowledge states are mental states that are characterized by causing the right kinds of behaviors and actions. Other mental states like belief may or may not represent the world in a way that can cause successful action, but knowledge states always will. Notice the pragmatist flavor, too—knowledge is just whatever allows us to do the cognitive work that we associate with knowledge.

4.2 Switching Profile

Representational adequacy follows a pattern not entirely dissimilar to that of the factive account. Much of the attitude switching entailed by the factive mental state

thesis will be retained here too, whenever a change in the state of reality changes whether some given mental content is representationally adequate for some given task. Notice, however, that now these changes won't always propagate instantly, nor will they propagate at the same speed for all knowers. For a newspaper editor in New York, switching on the morning of Lincoln's death occurs very quickly, as soon as it would be disadvantageous for the editor to carry on as if Lincoln were still president (which is to say, likely almost immediately after his death). However, for a settler living in an isolated frontier community, the representation that Lincoln is president will be adequate for the purpose of political discussion with peers, right up until the moment the news of Lincoln's death reaches them—days or perhaps even weeks later. While still allowing for very fast switching, these variable latencies are perhaps the most interesting feature of representational adequacy's switching profile.

Next, notice that while the changes in the state of reality that precipitate a switch might take time, causal contact with these changes is still not necessary for the switch to occur. Although we can still observe this in the Lincoln case, perhaps a cleaner example can better illustrate how distal switching is supported by the representational adequacy account. Consider Lukko's knowledge of the access code to enter the front door of the office building she works in. Assuming that her practical interests with the keycode are closely tied to being able to access the building, if that code is changed without Lukko's awareness, she instantly loses representational adequacy—and therefore knowledge—of what the access code is. This occurs regardless of whether she has any causal contact with the keypad in that moment.

Finally, as representational adequacy is generally a lower bar to clear than either truth or truthlikeness, switching will be marginally less common on this approach. Not only will low magnitude truth-to-truthlike changes in the state of reality not produce switches in many contexts, neither will some higher magnitude truthlike-to-untruthlike changes. While there might be some exceptions in contexts where an extremely high degree of accuracy is required for representational adequacy, overall we might expect switching to occur with a somewhat lower frequency than on either the factive or factivoid approach.

4.3 Epistemological Costs

Likely the most conspicuous consequence of replacing factivity with representational adequacy is that, in some cases, the close connection with truth widely assumed to characterize knowledge (see §3.3) is no longer tightly preserved. While successful goal-oriented action will frequently require at least moderately truthlike representations of the world, this is not always the case. To borrow a previous example (Bricker 2022b, §4), under the right conditions, misrepresenting one's hiking destination as 60 miles up the trail will be more than sufficient for successfully reaching a destination only 15 miles up the trail. In cases where wildly inaccurate representations can still guide successful action, truth and knowledge will come apart in a way not possible on the factivoid account.

Notice that there may be downstream effects of this loosened truth-knowledge connection, particularly for the usefulness of “knows” in communicative contexts. In

ordinary communication, we frequently use “knows” to indicate awareness of some truth in the world. For example, “You can stop pretending; the whole office knows by now that you’ve been the one eating everyone’s lunches.” Or, “Trust me, I know all about his long history of stealing lunches at work.” Moreover, “knows,” and the attribution of knowledge more generally, is frequently used to flag good sources of information—If *S* is represented as knowing whether *p*, then *S* is a reliable informant whether *p* (see especially Craig, 1990; Phillips et al., 2021, §6). If knowledge no longer entails truth, nor preserves a sufficiently tight connection with truth to at least serve as a close proxy, then this communicative role of signaling awareness of truth may be threatened, at least assuming that “knows” and knowledge attribution must display the same relation with truth as knowledge itself. Perhaps one way to avoid this consequence would be to downgrade the connection between knowledge (non-factive) and how we think/talk about knowledge in everyday contexts (factive)—something of the structural converse of Hazlett’s suggestion (2010) that knowledge, the thing epistemologists are interested in, is factive, but the verb “knows” is not. While I won’t speculate further here, such a move would be naturally understood as an additional cost for the view.

Moving on, taking knowledge to be a representationally adequate mental state almost certainly demands more revision to our intuitive judgements about knowledge than is required of the factivoid approach. This is most obvious when the truth-knowledge gap is greatest—Surely, we don’t intuitively judge that *S* knows that her destination is 60 miles up the trail when it is in fact only 15. However, even in more moderate cases, it is likely that our intuitive judgements don’t match the classifications of representational adequacy. Returning to one of the main examples offered by Buckwalter and Turri, I would submit that the standard intuitive judgement is that no *S* actually knows that π equals 3.14, even if only a grade-schooler. Again, while I certainly don’t want to say that intuitive revision is itself a reason to disregard a theory, it is a cost to be considered.

5 Knowledge as a Widely Shared Mental State

5.1 Overview

Looking beyond epistemology’s analytic project, we find a third option for theorizing about knowledge as a non-factive mental state: using “knowledge” like sociologists of knowledge⁹ to pick out a class of institutionally endorsed or otherwise widely held representational states. Consider the following characterization from Bloor (1976, pp. 2–3; emphasis added):

The sociologist is concerned with knowledge, including scientific knowledge, purely as a natural phenomenon. His definition of knowledge will therefore be rather different from that of either the layman or the philosopher. Instead of

⁹ Note that historians of knowledge do something similar. See, e.g., Burke 2015 (especially ch. 1).

defining it as true belief, **knowledge for the sociologist is whatever men take to be knowledge.** It consists of those beliefs which men confidently hold to and live by. In particular the sociologist will be concerned with beliefs which are taken for granted or institutionalised, or invested with authority by groups of men. Of course knowledge must be distinguished from mere belief. This can be done by reserving the word ‘knowledge’ for what is collectively endorsed, leaving the individual and idiosyncratic to count as mere belief.

Unlike the previous two accounts, this approach does not derive from skeptical concerns or conceptual arguments about how knowledge is attributed. Instead, it emerges from theoretical concerns about how best to characterize “regularities and general principles” observed by the sociologist in culture and society (Bloor, 1976, p. 3). In this case, the argument goes, there is a regularity to what people believe that warrants a distinction between knowledge—in some way driven or underwritten by social processes—and purely individual belief.

It is on the basis of this theoretically robust motivation that Kusch defends the legitimacy of the departure from factivity entailed by the sociology of knowledge program. Responding to Kvanvig’s claim that non-factive uses of knowledge like “it used to be known that the earth was flat” (2003, p. 190) are illegitimate and not theoretically motivated, Kusch writes (2009, p. 73):

To accuse generations of social scientists of misspeaking on the grounds that their use is not that favoured by mainstream epistemology strikes me as a misplaced attempt at linguistic legislating . . . Clearly the non-factive use of ‘knowledge’ favoured by social scientists is very much ‘related to theoretical concerns’. Social scientists aim to develop theories that explain why certain types of belief—or indeed, certain types of knowledge—are found credible in certain types of communities.

As with the previous two non-factive accounts, the explicit belief-centric formulation might easily be eschewed in favor of a knowledge-first approach to the sociology of knowledge, resulting in a theory of knowledge as a *widely shared mental state*, which is then subject to at least a few minimal constraints: First, crucially, this must be some kind of propositional attitude, which represents the world as being a certain way. Widely shared phenomenal states (hopes, fears, joys, pains, etc.), for instance, don’t qualify. In this way, this kind of mental state is reminiscent of group belief (see, e.g., Gilbert, 1987; Wray, 2001). Importantly, however, knowledge may go beyond mere group belief in that it is sanctioned by the group—taken to be knowledge, explicitly or implicitly. Notice that this would classify knowledge amongst strong social kinds, like money or marriage, the existence of which depends on our attitudes about the kind (see, e.g., Khalidi, 2015).¹⁰ And in this particular case of knowledge, there’s little mystery for where such attitudes might come

¹⁰ To be clear, I don’t mean to suggest that non-factivity is *necessary* for taking the view that knowledge is a social kind. Social kinds are commonly subject to mind-independent constraints (again, see Khalidi 2015 for discussion), so an alternative view of knowledge as a social kind might easily retain a factivity stipulation.

from. Humans have a deep, innate capacity to represent others as having knowledge, which displays a long evolutionary history, hallmark developmental stages in early childhood, and does not depend on the capacity to represent belief (see Bricker, 2020; Krupenye, 2021; Phillips et al., 2021). Knowledge, then, might be viewed as a widely shared propositional attitude that is at least as basic as group belief. And as these innate knowledge representations are computed like representations of a *sui generis* mental type (see especially Bricker, 2020; Phillips et al., 2021), all this adapts naturally to a mental state thesis. Now clearly more detail is required to make this a proper theory of knowledge, and not merely a sketch of one. Nonetheless, there's a natural knowledge-first view in here—Knowledge is, all told, a widely held representational state produced and maintained through significant human cooperation, and in some way culturally or institutionally sanctioned through the implicit or explicit application of the innate human concept of knowledge.

5.2 Switching Profile

The widely shared mental state thesis entails a very different kind of knowledge-reality connection than that of the other two approaches previously considered. Rather than knowledge requiring some connection with the represented state of reality, it now involves some connection with the community or social group responsible for the knowledge state obtaining. This means that changes in the represented state of reality will not necessarily precipitate switching. Lincoln's death won't cause members of a community to lose their knowledge that Lincoln is president, only their widespread abandonment of that attitude will. Crucially, however, this doesn't mean that fast and even distal switching cannot still occur.¹¹ Imagine that an isolated frontier community doesn't learn of Lincoln's death until a week after it occurs. In this time, on the widely shared mental state view, members of this community might still all know that Lincoln is president. But say that when news of Lincoln's death is announced at a town meeting, everyone is present but a single member of the community, who is alone in her cabin, bed-ridden with pneumonia. As soon as the rest of the community becomes aware of Lincoln's death, the view is no longer widely shared, so our bed-ridden protagonist ceases to know that Lincoln is president. Just like the attitude switching we've observed for previous versions of the mental state thesis, it occurs quickly and distally. Incredibly, even this extremely non-factive approach cannot seem to avoid distal switching. I'll add that not all switching licensed by the widely shared mental state thesis will occur this quickly. Social forces may also slowly change what a community takes to be knowledge over time, resulting in slow, proximal switching like we see with social externalism. Nonetheless, quite unlike social externalism, instantaneous distal attitude switching is still very much on the table. Finally, while certainly still common when compared to natural kind or demonstrative externalism, it does appear that switching is less common than on the factive or factivoid mental state theses. While difficult to

¹¹ I'm thankful to multiple anonymous reviewers for calling attention to these kinds of cases.

say definitively, I would speculate that changes to the state of reality that can result in the individual switching we've observed previously occur more frequently than group-level changes in what a community considers to be knowledge. At a minimum, switching is certainly easier on factive and factivoid approaches—less needs to change for switching to occur.

5.3 Epistemological Costs

Taking knowledge to be a widely shared mental state is easily the most epistemologically radical of the approaches we've considered. There will now be a significant gap between knowledge and truth, and while we might hope that social processes and other group-level factors might in some way be selective for truth, this is invariably a very, very non-factive approach to knowledge. Everything widely considered to be knowledge by some social group, regardless of its connection with truth, will be knowledge. Moreover, this approach excludes a large class of mental states ordinarily considered to be knowledge, all those low-level, individual, and idiosyncratic states that aren't widely shared or socially endorsed. In a particularly stinging blow for the conservative epistemologist, it would likely exclude at least much of our putative perceptual knowledge. This would represent an epistemological shift orders of magnitude greater than the odd revision to intuitive judgement required by the previous two accounts. Nevertheless, to be fair to the widely held mental state approach, while undeniably coming at a considerable epistemological cost, there may be a plausible rationale for accepting this cost. Kusch has suggested that this might represent the next step in the evolution of our concept of knowledge (2009, footnote 10). Appealing to Craig's idea of "objectivization," on which the concept of knowledge is thought to have developed out of a proto-concept that applied only to the single individual (1990), Kusch maintains that relativizing knowledge to epistemic communities might represent the next stage of objectivization.

Notice, moreover, that there are again likely to be downstream costs for loosening the truth-knowledge connection. As with the representational adequacy account, there is a worry that the role of "knows" and knowledge attribution in communication and learning from others may be degraded. And here the concern is even more pressing. If knowledge is largely delineated by our concept of knowledge, the connection between knowledge itself and how we talk/think about knowledge in everyday contexts cannot be downgraded without raising further complications. Additionally, notice that here the truth-knowledge connection is so weak that it begins to undermine the very distinction between knowledge and belief. Rather than one telling how the world is and the other telling how the world is thought to be, now the distinction is far less substantial. Belief straightforwardly enjoys a wider scope than knowledge, but beyond this, the distinction is increasingly marginal. It may be that sometimes a group might believe that p without believing that it knows that p , but if the central difference between group belief and group knowledge is that the later requires some kind of explicit or implicit group representation of that belief as being knowledge, then I'm not sure whether the distinction would even continue to be of significant interest to epistemology.

6 Diagnosis

Let's recap. We began with the observation that taking knowledge to be a factive mental state leads to a particularly severe version of externalism about the mind, on which externally mediated attitude switching is common, fast, and can occur distally. As it is common to attribute the externalism requirement to factivity, this raised the question of whether taking knowledge to be a non-factive mental state might allow us to avoid, or at least scale back, externalism-related costs. We considered three different proposals for weakening factivity that have attracted support in recent years, adopting each to a thesis about knowledge as a particular kind of non-factive mental state. In each case, the relationship between externalism, factivity, and the mental state thesis proved far more complex than is typically acknowledged (see Table 2). Nevertheless, one central observation is clear—The widespread assumption that the externalist costs of the mental state thesis are attributable to factivity is mistaken. Even a radically non-factive approach like the widely shared mental state thesis is still committed to attitude externalism, retaining distal switching and only marginally downgrading speed and frequency. There are two main lessons I want to draw from this.

First, the knowledge-reality connection goes far deeper than factivity. While it is typical for both belief-first and knowledge-first accounts to wrap up the link between knowledge states and external reality in a factivity constraint, it would be entirely mistaken to attribute knowledge's external metaphysics to factivity. Rather, factivity is but one way of capturing a more fundamental feature of knowledge. As far as I'm aware, no non-factive account of knowledge has been proposed that would even come close to fully decoupling knowledge states from the state of reality, and it's not at all clear how one might describe something even recognizable as a theory of knowledge without some substantial connection to the external world. Notice, however, that this coupling with reality need not have much to do with truth. It could be a matter of being able to successfully interact with the external world, or even just being in the same representational state as those around you. This observation is particularly important for knowledge-first epistemology, highlighting major constraints on what a plausible theory of knowledge as a mental state might look like. Even if knowledge isn't strictly factive, as a growing contingent of epistemologists now think, there may still be external demands on the class of non-factive states that might be plausibly considered knowledge states.

Second, because the knowledge-reality connection is not merely a consequence of factivity, there is no easy path to anything like an internalism-friendly account of knowledge as a non-factive mental state—or even a more moderate externalism, like content externalism or proximally constrained attitude externalism. As we saw above with the factivoid approach, we can even describe non-factive accounts that are more or less indistinguishable from their factive counterpart in switching profile. And, crucially, the externalist demands of mental state theses for knowledge do not scale down proportionally with the degree to which we weaken factivity. Moving from the minimally non-factive factivoid approach to

Table 2 Summary of the versions of the mental state thesis for knowledge considered here. “Revisions” refers to potential changes in what cases are considered knowledge, referenced to the factive mental state account. (+) denotes a potential addition to the body of putative knowledge, and (–) an omission. Frequency is relative to the total body of knowledge on a given account and does not reflect differences in what cases constitute knowledge

Theory of knowledge	Switching profile				Epistemological profile	
	Speed	Range	Frequency	Type	Truth connection	Revisions
Factive mental state	Fast	Distal	Most common	Attitude	Strongest	–
Factvoid mental state	Fast	Distal	Common	Attitude	Stronger	Very truthlike, quantitative representations (+)
Representationally adequate mental state	Slow-Fast	Distal	Moderately common	Attitude	Weaker	Misrepresentations adequate for successful action (+)
Widely shared mental state	Slow-Fast	Distal	Moderately common	Attitude	Weakest	Low-level content (–) Individual content (–) Idiosyncratic content (–) Widely shared misrepresentations (+)

the radically non-factive shared-mental-state approach results in only the slightest weakening of externalist demands—fast proximal attitude switching becomes marginally less common, accompanied now by slow, distal switching. As alluded to at the start of the paper, this largely negligible and entirely incommensurate response to weakening factivity confirms that factivity does not bear primary responsibility for the strong externalism associated with taking knowledge to be a mental state. If knowledge’s external-world connection were entirely, or even just primarily, a matter of factivity, then weakening factivity should at least appreciably weaken the associated externalism—even if it doesn’t eliminate it outright. Because this doesn’t prove to be the case, we can put to rest the widespread assumption that the unusually strong externalism associated with the mental state thesis is all down to factivity.

7 Conclusion

Here I have ventured to at least sketch out the beginnings of a merger between knowledge-first epistemology and the nascent non-factive approach to knowledge. We observed a number of ways in which extant proposals for weakening factivity can easily be adapted to theories of knowledge as a non-factive mental state. This allows us to describe versions of knowledge-first epistemology that might be particularly attractive to those a bit outside the boundaries of mainstream epistemology, especially pragmatists and sociologists of knowledge. Surprisingly, however,

none of these non-factive approaches lived up to the promise of a knowledge-first epistemology that's compatible with internalism about the mind, or even an appreciably more moderate version of externalism. This is the central contribution of the paper—Contrary to a widely held attitude, the thesis that knowledge is a factive mental state does not result in attitude externalism primarily in virtue of factivity. Instead, there is a much deeper connection between knowledge states and external reality, one that even non-factive approaches to knowledge cannot seem to plausibly dissolve. Accordingly, while there may still be potential for the future development of a non-factive knowledge-first epistemology, this unfortunately will not likely include the option to avoid some strong version of externalism about mental attitudes.

Acknowledgements I'd like to thank Paul Silva for inadvertently providing the impetus for this paper, as well as two anonymous reviewers for their incredibly helpful comments.

Funding Open Access funding provided by University of Turku (UTU) including Turku University Central Hospital. This research was supported by Academy of Finland postdoctoral grant no. 339681.

Declarations

Conflict of interest The author declares no conflicts of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Antognazza, M. R. (2020). XII—The distinction in kind between knowledge and belief. *Proceedings of the Aristotelian Society*, 103(3), 277–308.
- Bloor, D. (1976). *Knowledge and social imagery (reprint 1980)*. Routledge & Kegan.
- Bricker, A. M. (2018). Visuomotor noise and the non-factive analysis of knowledge. PhD Thesis. The University of Edinburgh.
- Bricker, A. M. (2020). The neural and cognitive mechanisms of knowledge attribution: An EEG study. *Cognition*, 203, 104412–104412.
- Bricker, A. M. (2022a). Knowledge is a mental state (at least sometimes). *Philosophical Studies*, 179(5), 1461–1481.
- Bricker, A. M. (2022b). Knowing falsely: The non-factive project. *Acta Analytica*, 37(2), 263–282.
- Buckwalter, W., & Turri, J. (2020a). Knowledge and truth: A skeptical challenge. *Pacific Philosophical Quarterly*, 101(1), 93–101.
- Buckwalter, W., & Turri, J. (2020b). Knowledge, adequacy, and approximate truth. *Consciousness and Cognition*, 83, 102950–102950.
- Burge, T. (1993) Mind-body causation and explanatory practice. In *Mental causation* (pp. 97–120). Heil & Mele (eds.). Clarendon.
- Burge, T. (1979). Individualism and the mental. *Midwest Studies in Philosophy*, 4(1), 73–121.
- Burge, T. (1986). Individualism and Psychology. *The Philosophical Review*, 95(1), 3–45.

- Burke, P. (2015). *What is the history of knowledge?* Polity Press.
- Carter, J. A., Gordon, E. C., & Jarvis, B. (2017). *Knowledge first: Approaches in epistemology and mind*. Oxford University Press.
- Craig, E. (1990). *Knowledge and the State of Nature: An essay in conceptual synthesis*. Oxford University Press.
- Evans, G. (1982). *The varieties of reference*. Oxford University Press.
- Fricke, E. (2009). Is Knowing a State of Mind? The Case Against. In *Williamson on Knowledge* (pp. 31–59). Eds. Greenough, P., & Pritchard, D. Oxford University.
- Gilbert, M. (1987). Modelling collective belief. *Synthese*, 73(1), 185–204.
- Greenough, P., & Pritchard, D. (2009). *Williamson on knowledge*. Oxford University Press.
- Hazlett, A. (2010). The myth of factive verbs. *Philosophy and Phenomenological Research*, 80(3), 497–522.
- Kaplan, D. (1989). Demonstratives. In *Themes from Kaplan* (pp. 481–563). Eds. Almog, J., Kaplan, D., Perry, J., & Wettstein, H. K. Oxford University Press.
- Khalidi, M. A. (2015). Three kinds of social kinds. *Philosophy and Phenomenological Research*, 90(1), 96–112.
- Krupenye, C. (2021). The evolution of mentalizing in humans and other primates. *The neural basis of mentalizing* (pp. 107–129). Springer International Publishing.
- Kusch, M. (2009). Testimony and the Value of Knowledge. In *Epistemic Value* (pp. 60–94). Eds. Haddock, A., Millar, A., & Pritchard, D. Oxford University Press.
- Kvanvig, J. (2003). *The value of knowledge and the pursuit of understanding*. Cambridge University Press.
- Ludwig, K. (1994). Causal relevance and thought content. *The Philosophical Quarterly*, 44(176), 334–353.
- McGlynn, A. (2014). *Knowledge first?* Palgrave Macmillan.
- McLaughlin, B. (1989). Type Epiphenomenalism, type dualism, and the causal priority of the physical. *Philosophical Perspectives*, 3, 109–135.
- Nagel, J. (2013). Knowledge as a mental state. *Oxford studies in epistemology*. Oxford University Press.
- Niiniluoto, I. (1999). *Critical scientific realism*. Oxford University Press.
- Pavese, C. (2021). Knowledge and mentality. *Philosophical Perspectives*, 35(1), 359–382.
- Phillips, J., & Bricker, A. M. (2022). Unpublished Data. URL = <https://github.com/adambricker/Shared_knowledge>.
- Phillips, J., Buckwalter, W., Cushman, F., Friedman, O., Martin, A., Turri, J., Santos, L., & Knobe, J. (2021). Knowledge before belief. *Behavioral and Brain Sciences*, 44, e140–e140.
- Putnam, H. (1975). The meaning of ‘Meaning.’ *Minnesota Studies in the Philosophy of Science*, 7, 131–193.
- Smith, M. (2017). The Cost of Treating Knowledge as a Mental State. In *Knowledge First* (pp. 95–112). Eds. Carter, A., Gordon, E., & Jarvis, B. Oxford University
- Soteriou, M. (2020). The disjunctive theory of perception. In *The Stanford Encyclopedia of Philosophy* (Summer 2020 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/sum20/entries/perception-disjunctive/>>.
- Starmans, C., & Friedman, O. (2020). Expert or esoteric? Philosophers attribute knowledge differently than all other academics. *Cognitive Science*, 44(7), e12850–n/a.
- Stjernberg, F. (2009). Restricting factiveness. *Philosophical Studies*, 146(1), 29–48.
- Westra, E., & Nagel, J. (2021). Mindreading in conversation. *Cognition*, 210, 104618–104618.
- Williamson, T. (1995). Is knowing a state of mind? *Mind*, 104(415), 533–565.
- Williamson, T. (2000). *Knowledge and its limits*. Oxford University Press.
- Wray, K. B. (2001). Collective belief and acceptance. *Synthese*, 129(3), 319–333.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.