

Smith – The cost of treating knowledge as a mental state

Brian T. Miller

September 19, 2018

Central argument of the article:

- Williamson claims that knowledge is a mental state (hereafter KSM)
- That thesis is costly, in terms of being inconsistent mental state internalism
- It's *more* costly than a mere commitment to content externalism

Objection from Internalism

Mental state internalism mental states are determined (individuated?) by the internal physical state of one's body and brain

Content internalism the contents of mental states are determined by the internal physical state of one's body and brain

Knowledge is factive, so whether or not I know that p depends on whether p is true, which often depends on the state of things outside my body

So, if KSM then Mental State Internalism is false.

But many already reject Content Internalism. Assuming that mental states are partially individuated by their contents, those people already reject Mental State Internalism.¹ In that case, is how objectionable is KSM?

Basic point: it depends on your motivations for accepting internalism; for some, the degree of deviation from internalism matters a lot

If [a reductionist theory of mind] was one's reason for accepting internalism, then perhaps it would have to be accepted to the letter – one could tolerate no divergence from it. But there are far more pedestrian thoughts that seem to point us in an internalist direction. Here is one: It shouldn't be possible to influence or interfere with a person's state of mind without causally interacting in some way with that person's sense organs or body. If this was the reason one was inclined towards internalism, then one might be quite open to qualifying or compromising the view in various ways.

¹ Why only 'partially'? I might both fear that p and believe that p, and those are distinct mental states. So difference of content isn't necessary for difference of mental state. Plausibly, it's necessary: my belief that p and my belief that q are distinct mental states.

Natural kind externalism

One motivation for rejecting internalism: Twin Earth thought experiments

[There's] a remote planet, exactly like the Earth, but completely devoid of water – that is, H₂O. Instead of water, a different chemical compound – XYZ – flows in the rivers, fills the lakes and oceans, falls as rain and quenches the thirst of the Twin Earthers. XYZ has all of the same macroscopic properties as water – it's clear, odourless, boils and freezes at 100 and 0°C at sea level etc. Furthermore, the Twin Earthers, who speak like us, even call it 'water' (and 'eau' and 'mizu' etc.). Finally, we might imagine that it's prior to 1750 and, without access to microscopes and the like, no one on Earth or Twin Earth would be capable of distinguishing these two substances. According to Putnam, while those on Earth and those on Twin Earth might associate exactly the same ideas and descriptions with the term 'water', when the former use the term it refers to H₂O, and when the latter use the term it refers to XYZ.

The example purports to show that *word meanings* are not determined solely by the agent's physical state. Similar considerations might show that contents of mental states aren't so-determined either.

Suppose that's right: what happens when you're transported from Earth to Twin Earth? How does that affect the contents of your mental states?

Plausible suggestion: after a while, your 'water' thoughts (beliefs, fears, hopes) come to be about XYZ rather than H₂O.

One way this could happen: your causal interactions with XYZ change your physical state in some way, rearranging your neural pathways or something

- that's perfectly consistent with Mental State Internalism

Another way it could happen: you interact causally with XYZ, but that results in no changes to your physical state²

- that's not perfectly consistent with Mental State Internalism

² Weaker: your internal state might change, but those changes don't result in the change in mental contents

Let 'switching' refer to the second way: change in content of mental state with no change in internal physical state

Plausibly, Natural Kind Externalism allows for switching in cases where the immediate environment changes. This switching is:

slow it takes a while after reaching Twin Earth before you mental states are about XYZ rather than H₂O

proximal requires causal interaction with the nearby environment

rare Twin Earth teleportation cases probably don't happen all that often

Social Externalism

...imagine an individual who is unaware that arthritis is a condition of the joints and complains to her doctor of arthritis in her thigh. When she says 'I have arthritis in my thigh' she expresses a (necessarily) false belief. Now consider another individual, with exactly the same history and internal states, who happens to be part of a linguistic community in which the term 'arthritis' is used to refer to a broader class of rheumatoid ailments which happens to include the very complaint she has in her thigh.

According to Burge, when this second individual says to her doctor 'I have arthritis in my thigh' she is not expressing a false belief about arthritis, a condition exclusively of the joints – neither she nor anyone in her community has beliefs about arthritis. Rather, she is expressing a true belief about the broader category that her linguistic community recognizes [call that more widespread ailment 'tharthritis']. If we accept this verdict, we are led to a view that we might call social externalism – the content of certain beliefs can depend upon facts about how terms are used in the believer's linguistic community.

Putting that conclusion another way: the meaning of the word 'arthritis' is determined by social factors external to my body: its use in the broader community. Similarly, exactly what my *arthritis* thoughts are about is determined social factors external to my body. So, if mental states are partially individuated by their contents, then whether in an *arthritis* mental state or a *tharthritis* mental state is determined by factors external to my body, so mental state externalism is true.

Comparing Social Externalism with Natural Kind Externalism:

- Social Externalism has much broader scope: its considerations apply to nearly any thought/ word
- As before, Social Externalism plausibly permits switching:
 - travel from US to UK, word meaning change (individuating words by sounds and spellings here)
 - plausibly, this changes the meanings of words spoke and concepts deployed by the traveler
- here, switching is
 - *slow*: you might need to spend a while in the UK before your *biscuit* thoughts come to be about cookies

- *proximal*: crucial facts are about your community of speakers/thinkers, and those issues are typically facts about what's physically nearby
- *common*: plausibly, this happens any time you travel to a new community and stay long enough

Demonstrative Externalism

...when one uses a demonstrative expression, one's assertion does not merely have descriptive content ... rather, its content may literally involve the perceived object being demonstrated. Suppose I see an apple on the table and, pointing towards it, utter the words 'That apple is overripe'. On the present view, if it had been a different apple sitting on the table, then the content of my assertion would have involved a different object and, thus, would have been a different content, even if the new apple looked exactly the same to me and I was in exactly the same internal state.³

This describes *Demonstrative Externalism*: contents of certain beliefs involving demonstrative concepts can depend upon the observed objects being demonstrated

Presumably, switching cases occur anytime the referent of the demonstrative – then thing you're 'pointing at' (metaphorically speaking) changes. This switching is

- *very fast*: content of thought changes the instant the environment changes
- *rare*: how often does someone surreptitiously switch the apple on the table that's your referent of 'that apple'?
- *proximal*: typically the objects that might be switched are in your nearby environment, though they needn't be (MS imagines switching cases conducted via Skype)

Disjunctivism

Clarification: MS is here discussing *metaphysical* disjunctivism rather than *epistemic* disjunctivism

Core thesis: veridical perception and non-veridical perception (hallucination, illusion) are distinct mental states, even when they're subjectively indistinguishable. There is no unified category of 'perceptual experience', only a composite of distinct mental state types

³ Similar considerations motivate *indexical* externalism, where indexicals include 'I', 'here', 'now', etc.

This very similar to Demonstrative Externalism, which says that my thought about *that apple* is has a different content from a subjectively indistinguishable thought about an apple that does not in fact exist

As long as we allow that the agent can be in the same internal physical state whether perceiving veridically or hallucinating/ suffering illusion, Metaphysical Disjunctivism implies Mental State Externalism.

Here switching is possible: imagine a veridical perception of the cup on the table, and then the cup vanishes and is replaced by a super-realistic hologram. The internal physical state of the agent has not changed, but the mental state has switched from *veridical perception* to *illusion*. This switching is:

- *very fast*: instantaneous
- *rare*: presumably it's rare that objects of our perception vanish and yet we have indistinguishable experiences
- *proximal*: changes in mental state require changes in the perceived environment

Back to the thesis that Knowledge is a Mental State

MS's claim: KSM requires a much greater departure from Knowledge State Internalism than anything we've considered

As before, whether I know depends in part on what's going on outside of my body: I can't know that p unless p is true, and the truth of p often depends on things outside of my body.

As a result, switching is possible when facts external to my body change. What's the nature of that switching? Case:

I currently believe, and know, that my local fishmonger has a blue shopfront. Suppose that, while I'm 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 shop front had not been touched.

Here the switch from knowing that the shopfront is blue to not knowing it is:

- *very fast* the change to my mental state occurs the instant the external world is changed (when the shopfront is painted)
- *very common* no weird thought experiments to imagine here
 - not just common as with Social Externalism, which occurs every once in a while. Given KSM switches happen many times a day, when some proposition we know stops being true. So this type of switching is *extremely common*
- *distal* I might be very, very far away when the shopfront is painted. Possibly many lightyears away.

This combination is new, more of a departure from Mental State Internalism than other view considered above.

Further departures result from also accepting *safety* (as TW and many others do)

safety one's belief is *safe* iff it could not easily have been the case that one falsely believed P via M at t⁴

- Say that one's belief is safe iff there is no close possible world in which one believes P via M at t and P is false.

⁴ proposition P, via method M, at time t.

So about that instantaneous switching...

...return to the case of Lincoln's assassination. Booth...could have easily decided to shoot Lincoln a few minutes earlier than he did [i.e.] there are close possible worlds in which he does shoot Lincoln a few minutes earlier than at the actual world. If this is right then, on the evening of April 14th 1865, the belief that Lincoln is president of the US would have ceased to be *safe* at least several minutes before it ceased to be *true*... If knowledge requires safety, and is a mental state, then this is the point at which the mental states of people the world over would have switched – *prior* to the event that would appear responsible for the switch. What we seem to confront here is a case of *backwards* switching, where an event triggers a mental state switch *in the past*.

MS goes one to consider a possible response that's worth thinking through:

Perhaps what these reflections about safety show is that the assassination itself is *not* what truly triggers the mental state switch – rather other events are responsible, such as Booth approaching Ford's theatre,

his crouching outside the presidential box, pistol in hand etc. And these events, we might insist, really do take place before the mental state switch does. The trouble with this redescription, though, is that if Booth had not gone on to assassinate Lincoln, it's not clear that there would have been *any* mental state switching at all, even if these other events had still taken place.

BTM:

- It's hard to say anything precise here because we don't have a criteria for the closeness of possible worlds.
- important point: TW (who is very careful about these things) discusses safety not in terms of possible worlds, but of *cases*. Cases are just like worlds except
 1. they're centered on a time, place, and individual, so 'I'm here now' doesn't require any *further* specification of the referents of its indexicals in order to have a truth value
 2. they're time-slices of worlds, rather than temporally extended worlds. So it makes sense to compare the closeness of the actual world at one moment to the actual world to a millionth of a second later to some moment in another world or in the actual world at a very different time: each is simply a case

Understood in this framework, Smith's diagnosis of the situation seems misguided. It's never (too strong?) the *event* that changes the mental state, or at least never the event alone. If knowledge requires safety, then knowledge requires nearby cases to have certain properties. Those he cases nearby to case c will likely include other temporally nearby time-slices of the same possible world, and they will include cases from other possible worlds. Each of these types of cases – non present but actual, non-actual but present, and neither present nor actual – can be relevant to whether or not my belief is safe.

What's the lesson? Smith seems to think there's a kind of backwards-causation worry. But the problem is bigger than that: there's a more general 'how could all of that other stuff possibly be relevant to what I know' worry! We've got both backwards and forwards causation, and effects on my actual, contingent mental states from non-actual possibilities. That's pretty strange!