Extending Introspection

Lukas Schwengerer – Lukas.Schwengerer@uni-due.de

Penultimate Draft for the collected volume *The Mind-Technology Problem - Investigating Minds, Selves and 21st century Artifacts*. Please cite the published version.

Abstract

Clark and Chalmers propose that the mind extends further than skin and skull. If they are right, then we should expect this to have some effect on our way of knowing our own mental states. If the content of my notebook can be part of my belief system, then looking at the notebook seems to be a way to get to know my own beliefs. However, it is at least not obvious whether self-ascribing a belief by looking at my notebook is a case of introspection the same way that knowing my non-extended beliefs is. Traditionally this sort of introspection is thought to be privileged and special in ways that the extended introspection case seems not to be. There is nothing privileged about looking at my notebook. Anyone could do it. The aim of the paper is to find out how to understand extended introspection and whether there is something privileged and special about knowing one's own extended beliefs. Moreover, the notebook case has close analogs using 21st century technology. It seems possible to know our beliefs that are extended to smartphones, wearable technology or a cloud-based data store. First, I present the case of extended introspection. I then discuss whether it should be understood as ordinary introspection or as mind-reading. Both seem to be bad fits, which finally prompts an original account for extended introspection based on epistemic rules.

The Case of Extended Introspection[1]

Let us start with a story about Otto:

> Otto suffers from Alzheimer's disease, and like many Alzheimer's patients, he relies on information in the environment to help structure his life. Otto carries a notebook around with him everywhere he goes. When he learns new information, he writes it down. When he needs some old information, he looks it up. For Otto, his notebook plays the role usually played by a biological memory. In short, Otto has beliefs extending to the notebook. One day I meet Otto in the streets of New York. Out of curiosity I ask him "Otto, I know where the museum is, but do you believe the museum is on 53rd street?" Otto looks at his notebook, finds the right entry, and answers "I believe the museum is on 53rd Street."[2]

[1] Special thanks to Giada Fratantonio, Grace Helton, Jesper Kallestrup, Aidan McGlynn and two anonymous reviewers for feedback on earlier versions of this paper. Large parts were presented at the University of Edinburgh WIP seminar and the conference 'Minds, Selves and 21st Century Technology' at the New University of Lisbon. Further thanks go to the audiences at these events.

[2] This is different from the extended self-knowledge case by Carter and Pritchard (Forthcoming), which has the self-ascription written down in the notebook: "[…] For example, when he learns new information about his own mental states (i.e., beliefs, feelings, desires, etc.) – information about his mental states which would be lost in biological storage – he writes it down in the notebook. Likewise, when he needs some old information about his mental life, he looks it up. For Otto*, his notebook plays the role usually played by a biological memory in

At face value in this story Otto self-ascribes a belief state by looking at the notebook. But this leads straight into conflicting intuitions:

- On one hand Otto appears to simply detect his belief. The story is set up in a way that he has beliefs extending to the notebook, so the way to detect these beliefs is to look at the notebook. And intuitively, what else is introspection other than directly detecting your own beliefs?
- On the other hand Otto appears to base his self-ascription on evidence that I, standing next to Otto, can use just as well to ascribe the belief to Otto. And this is certainly not what we intuitively think introspection is like.

With these conflicting intuitions at hand, we need to search for a proper way to understand extended introspection. Otto definitely self-ascribes a mental state by looking at the notebook. But what kind of self-ascription are we dealing with here? This is the core question I am going to address in this paper. I start by stepping back and taking a quick overview on the notion of extended belief as the basis for extended introspection. I then consider the standard options for self-ascribing mental states: introspection and self-directed mind-reading. I discuss these in turn and show that both are not a great fit. This leads into a dialectic dilemma with no clear way out. I then propose a unique account of extended introspection based on epistemic rules, inspired by Alex Byrne's (2005) account of introspection. On my proposal extended introspection turns out to be reliable, because self-verifying under extended belief condition. Moreover, I show in the final section that my solution for the case of Otto and his notebook is also applicable to 21st century versions of the case, relying on smartphones, wearable technology and the internet.

---

preserving a mental narrative." I will use 'extended introspection' instead of 'extended self-knowledge' to avoid confusion.

Extended Belief

My story about Otto introspecting extended beliefs builds on the Otto case presented by Clark and Chalmers (1998).[3] The central idea behind ascribing Otto plus notebook an extended belief lies in the fact that the notebook plays the role usually played by biological memory. This idea features in the argument as the parity principle:

> *Parity Principle*: If, as we confront some task, a part of the world functions as a process which, were it to go on in the head, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world is (so we claim) part of the cognitive process (Clark & Chalmers, 1998, p. 8).

The notebook does the job of the biological memory, because Otto consistently uses it just like we usually use a biological memory. Whenever he gets new information he writes it down. Whenever he wants to remember something he looks it up in the notebook. And the notebook is with him all the time. So why not straight up accept the notebook as Otto's memory?

The difference in location is not a well-motivated rationale to dismiss the similarities, if we follow Clark and Chalmers. However, this needs a further, more detailed look. Even if we follow the parity principle, we need to say something more on the notebook's role in Otto's cognitive processes. Not every notebook makes a good case for something that "we would have no hesitation in recognizing as part of the cognitive process," that is in this case, as part of memory. I personally use notebooks every once in a while. But if you were to observe my sporadic looks at the notes, you would surely hesitate to call my notebook part of my memory. Moreover, sometimes I write notes so illegibly that there is no way to decipher the content later on. Other times I find myself disagreeing with my notes – "I can't have meant that," I say to myself. And finally, often my notebook is just not with me, whereas my biological memory is consistently with me. Or at least I cannot remember the cases in which it was not. In short: My notebook's role is very differently from biological memory. We have to

[3] Clark (2010) explicitly allows that the 'reading in' of information from devices might count as introspection depending on how one classifies the overall case. He states that "[f]rom our perspective the systemic act is more like an act of introspection than one of perception" (p. 56).

be careful not to trivialize extended beliefs in a way that obscures their differences with non-extended beliefs.

Clark and Chalmers are fully aware of this concern. Therefore they propose four conditions that external aids have to satisfy to play a role in an extended belief (Clark & Chalmers, 1998; Clark, 2010).

External aids must be:

(i) consistently available,

(ii) readily accessible,

(iii) automatically endorsed, and

(iv) present because they were consciously endorsed in the past.

They are slightly skeptical about the previous endorsement condition, claiming that this requirement is debatable. The appeal of (iv) is that by getting rid of it one might lose grip of the difference between remembering and relearning. However, insofar as (iv) requires conscious endorsement, which is an internal condition, (iv) seems to rob the extended belief thesis of its core motivation. The cognitive work would not be extended enough, so to speak (cf. Bernecker (2014, p. 5)). For my purposes condition (iv) is not important, so I leave the debate open. However, conditions (ii) and (iii) will become important later on. Fortunately, they are generally accepted as requirements for extended beliefs.

Clark (2008) and Menary (2007) introduce additional conditions of two-way interactions. The idea is that the interactions between the external aid and the person have to be connected in such a way that they continuously affect the other. Clark calls this *continuous reciprocal causation*, Menary *cognitive integration*. An old-fashioned, paper-based book, for instance, would not fit this criterion because the reader does not affect the book sufficiently. For my purposes the conditions (i) to (iii) above are enough, so I can stay neutral on any further requirements. I will also remain neutral on whether the analogy to memory works as well as Clark and Chalmers want it to. The use of notebooks can fail more frequently and in different ways than our biological memory, which might

undercut the parallel drawn by the argument for the extended mind (cf. Rupert (2004), Sterelny (2004)). However, for the present purpose of discussing extended introspection, I will assume that the case can be made in favor of Clark's and Chalmers's picture. Hence, my approach can be seen as a discussion under a conditional: What should we think of introspection of extended beliefs, *if there are extended beliefs and they can be roughly characterized by conditions (i) to (iii).*

There is a final worry to address concerning the focus on the simple Otto case. The case is highly idealized. Why should we care about this case at all? Perhaps looking at other, more realistic scenarios requires different conditions or dimensions of integration (as discussed in Sutton (2006), Sutton et al. (2010), Sterelny (2010), Menary (2010), and Heersmink (2015))? After all, today's technological gadgets are more complex than the simple notebook. My first response here is to point out that the case does not strike me as unrealistic, even though it involves an uncommon situation. Moreover, there are sufficiently similar cases featuring individuals and smartphones or smartwatches that satisfy the same conditions (i) to (iii). For these cases my discussion should still be applicable. I am going to discuss one paradigmatic case more in depth towards the end, but I am now going to address the idea that similar cases are plausible featuring 21st century technology. Fortunately, such a scenario has already been provided by Paul Smart (2018). He presents the following case Otto++:

> Otto++ is a neurologically impaired individual who is biologically identical to Otto. Otto++ has just purchased a shiny new smartphone and augmented reality glasses. Otto++ spends some time configuring his phone by installing a variety of apps. He then carries his phone with him wherever he goes. In order to ensure that he has access to relevant information, Otto++ installs an app that enables him to record important pieces of information. The app, however, does not store information locally on the device. Instead, it relies on a semantically enabled, cloud-based personal data store that stores information in a linked data format. In order to access his personal data store, Otto++ installs an app that enables him to quickly retrieve important items of information using an intuitive graphical user interface. He also links his phone to his augmented reality glasses so that relevant information from his data store can be presented within his visual field. One day, while on a trip to New York City, Otto++ decides he would like to visit MoMA. He automatically says the word "MOMA" out loud. His phone executes a semantic query against his personal information repository and retrieves information about MoMA. A set of directional indicators appear within Otto++'s visual field, alongside some descriptive information about MoMA (Smart, 2018, p. 279).

Smart argues that Otto++ still satisfies conditions equivalent to my (i) to (iii). To show that we have to unpack the story a little. One important difference to the notebook case is that Otto++ does not store information on any particular device. Otto++ stores the information on a cloud-based personal data store that can be accessed with various different devices. He carries some of these devices with him, but he could in principle also use a friend's smartphone if needed. Moreover, the information is stored in a specific format[4] such that it can be retrieved quickly and easily based on queries. This is meant to counteract the difficulties of accessing a specific piece of information in a large database. These features combined allow Otto++ to have the information constantly available and readily accessible, at least to a degree similar to biological memory. Hence, the case passes (i) and (ii). Condition (iii) appears to be satisfied as well. There is nothing in the case that prevents Otto++ from automatically endorsing the information retrieved from his personal cloud-based data store. We can even stipulate that (iv) is satisfied. Simply suppose that information can only be uploaded to the data store by Otto++ himself and (iv) is met. Given these similarities Smart suggests that "[…] the case of Otto++ is sufficiently similar to the original case […]" (2018, p. 279). I agree. Hence, we can imagine an Otto++ case for extended introspection as well. Otto++ apparently can come to know what he believes by suitably attending to his cloud-based data store. Moreover, I seem to be able to come to know what Otto++ believes by suitably attending to Otto++'s cloud-based data store.

My plan is to focus on the simpler case of Otto and his notebook and then discuss whether introspection in the Otto++ case can be explained by the same strategy as the simpler case. The focus on the notebook case makes the discussion a little easier, without significant drawbacks. Nevertheless, it is important to point to potential problems in generalizing from the simple case to more complex cases, such as Otto++. I will respond to these problems towards the end. However, I will not discuss speculative cases involving belief constituting cognitive enhancements that are directly connected to the brain. I find these scenarios still rather alien and it is difficult to have clear intuitions about them.

[4] For details on the format see Smart (2018, pp. 273-274).

Extended Introspection as Introspection

To find out whether the extended introspection case is a genuine case of introspection, we need to define our criterion for such genuine introspection. A good starting point for doing so is to look at our intuitive judgments about introspection, and most importantly what they are contrasted to. The two important points of contrast are first, knowledge of the non-mental world; and second, knowledge of other people's mental states. Introspection appears to be different to both. Self-knowledge seems to some extent *privileged* and *peculiar* (Byrne, 2005). It is privileged insofar as one's beliefs about one's own mental states are more likely to amount to knowledge than one's beliefs about other's mental states or the external world. Moreover, they are peculiar, insofar as they are formed by a special method or way of knowing. Call this the *cognitive access view* of self-knowledge. Cognitive access accounts come in different shapes. For instance, they can accept a peculiar detectivist method of introspection (e.g. Armstrong (1968), Nichols and Stich (2003), Goldman (2006), Macdonald (2014))[5] or an empiricist transparency story as in Byrne (2005), and Fernández (2013).

The cognitive access view is not the only description of self-knowledge available. For instance, s*elf/other parity* accounts of self-knowledge argue that the specialness of self-knowledge is overstated. Self-beliefs are largely[6] formed by the same processes we use to attribute mental states to others (cf. Carruthers (2011), Cassam (2014)). Moreover, a*gentialist* (sometimes called *rationalist* (Gertler, 2011)) positions (e.g. Burge (1996), Moran (2001), Bilgrami (2006)) understand privilege and peculiarity not in terms of better access, but with a particular first-personal connection between agents and their mental states. For Moran (2001) this particular connection is also the basis for the transparency of beliefs, that is, that one can know whether one believes that *p* by attending to the question of whether *p* is true. Agentialist accounts accept that self-belief is special because it is about *my* mental states, and I am responsible for *my* beliefs. Self-knowledge in this conception is important as a precondition for critically reflecting on one's own mental states.

[5] These lists are not exhaustive.
[6] 'Largely' because they usually limit the parity to propositional attitudes.

For this paper I will restrict myself to the cognitive access view due to limited space. Hence, I understand privileged access for introspection as a person being more reliable in self-ascribing mental states than in ascribing mental states to other people. Moreover, I take peculiar access to denote some sort of peculiar way of knowing one's own mental states, compared to knowing other people's mental states ('mind-reading').

With the features of privileged access and peculiar access as the defining features for introspection we can start looking at whether extended introspection presents us with these features. When Otto self-ascribes his belief by looking at the notebook, does he have privileged and peculiar access?

The first step to investigate his privileged position is to consider whether Otto is reliable. Clearly he is in the way the story is set up. Otto has, by stipulation, the extended belief that the museum is located on 53rd street. When I ask him, he looks at the notebook and avows that he believes the museum is on 53rd street. And in doing so, he makes a correct statement. It is true that he has this belief. Even more so, it seems very difficult for Otto to be wrong about himself in this case. Given that the notebook is consistently available, readily accessible, and automatically endorsed, Otto is highly reliable in looking at the notebook and self-ascribing the belief. In nearly all nearby possible worlds in which Otto looks at the notebook to self-ascribe the belief, he will read the notebook correctly, understand what is written in the notebook and successfully attribute the belief to himself. If this was not the case, then Otto would fail to have the appropriate connection to the notebook and not have an extended belief. But I stipulated the extended belief conditions to be satisfied. Hence Otto has to be reliable.

However, is Otto *more* reliable than other people looking at his notebook? Privileged access requires more than just reliability. It requires being more reliable than other people in attributing mental states to Otto. Consider the following take on Otto:

> Otto suffers from Alzheimer's disease, and like many Alzheimer's patients, he relies on information in the environment to help structure his life. Otto carries a notebook around with him everywhere he goes. When he learns new information, he writes it down. When he

> needs some old information, he looks it up. For Otto, his notebook plays the role usually played by a biological memory. I look at Otto's notebook and read that the museum is at the 53rd street. Therefore, I judge that Otto believes the museum is at 53rd street.

Is my belief that Otto believes the museum is at 53rd street reliably formed? This depends on how we individuate the process. If the process in question is "looking at notebooks" in general, then it probably is not. However, given that I stipulated Otto's extended belief, it seems the answer is a clear yes if we individuate the process based on looking at *Otto's* notebook. That is, if I ascribe Otto beliefs based on looking at *his* notebook, I will be highly reliable. As long as I can read Otto's handwriting and understand his language, I will end up with true beliefs about Otto's extended beliefs.[7] Therefore I seem to be equally reliable as Otto in ascribing extended beliefs to Otto. Otto is not privileged.

Perhaps I am a little too quick here. One might bring up at least two objections. First, one may claim that there is a sense in which Otto is privileged. Otto has his notebook with him all the time, whereas I don't have the same kind of permanent access. So he has better access after all! To this I respond that this is not the right sort of privilege that I am talking about. It is important to distinguish two kinds of privileged access here. First, one can have 'accidentally' privileged access because one is more often in possession of the relevant evidence, but that evidence (or sufficiently similar evidence) is in principle accessible for other people. Second, one can have 'essentially' privileged access, when the evidence (or sufficiently similar evidence) is not in principle accessible for other people. Otto seems to have accidentally privileged access to his extended belief, but no essentially privileged access. And it is the latter that is of interest here. If I were to follow Otto all the time and look at his notebook constantly there would not be an extended belief of his that he knows but I do not.

Second, one can argue that one needs to have reasons to base judgements of Otto's beliefs on the notebook. One needs to know not only that this is Otto's notebook, but also that Otto is related to the notebook in a way that satisfies the conditions for extended belief. I can accept the former, but deny the latter. One needs some reason to ascribe a belief based on the notebook, but knowing that

[7] And there seems little reason to give Otto any privilege with reading and understanding the notebook, or at least nothing that could not be removed by modifying the case slightly.

this is Otto's notebook is enough.[8] After all, Otto himself does not need to know that the extended belief conditions are met. He just needs to look at the notebook and self-ascribe the belief. The same applies to me ascribing mental states to Otto by looking at his notebook. I need to have some reason why this notebook relates to Otto's beliefs, but those reasons need not entail any of Otto's beliefs. These reasons are only necessary to prompt me to ascribe a belief on the basis of the notebook. They need not be reasons that guarantee the truth of my ascription.

What about peculiar access? Is there anything special about Otto's self-ascription? At first glance what Otto does and what I do when we ascribe a belief based on the notebook seems not that different. We both look at the very same thing and use it as a basis for a belief ascription. It is the same notebook. Sure, he has access to the notebook constantly, but that does not change the fact that when we both look at it to ascribe a mental state, we do the same thing. Otto might use the notebook for much more, but for this single belief-forming process it is hard to find a difference. That the notebook constitutes his first-order belief does not influence the second-order belief formation. Moreover, there is no reason to assume more inferential work being done by me than Otto. Sure, I need to recognize that this is Otto's notebook. But so does Otto. I don't see any reason why I would need any additional inferential step that Otto does not need. I can treat whatever is written in the notebook as direct evidence for Otto's belief, just as Otto does. Gertler (2007) uses this observation to argue against the existence of extended beliefs. As I will show later this is not the route that I want to take. However, Gertler is right that this symmetry between Otto and me is bad news for peculiar access and overall bad news for extended introspection as a subspecies of genuine introspection.

[8] This includes knowing something about the function of a notebook. Some knowledge of the function is necessary to distinguish one's notebook from a scrap of paper lying around. That it is my notebook and not a random scrap of paper provides a prima facie reason to attribute to myself the content. Thanks to Grace Helton for pointing out this issue.

Extended Introspection as Mind-Reading

Mental state ascriptions are commonly assumed to be either introspectively formed, or based on mind-reading.[9] I just argued that extended introspection does not fit the criteria of privileged and peculiar access, so it does not look like it goes into the introspection category. One should then expect a better fit with mind-reading. And if so, calling it 'extended introspection' might actually be misleading. To see whether this is true I want to start with a rough and ready characterization of mind-reading. I understand mind-reading here as a capacity to attribute mental states to human beings based on behavioral observations and evidence of the situation/environment. To illustrate this take this simple mind-reading story by Jordi Fernández (2013, p. 4):

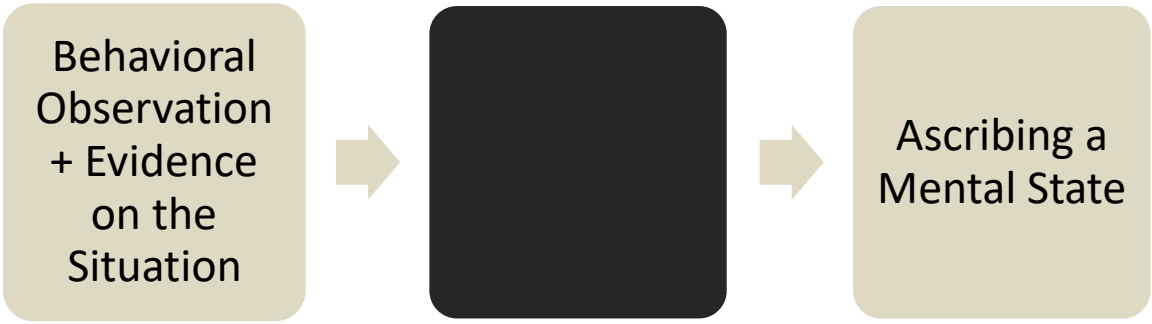
> "Suppose that one of the things that you believe about me is that I want Barcelona FC to win the UEFA Champions League. Suppose, furthermore, that your belief is justified. What could justify your belief? Perhaps you heard me express that desire, or you observed me screaming at the TV while we watched one of the Champions League games, or you noticed my mood when I read in the news that the team was not doing so well in that competition."
> (Fernández, 2013, p. 4)

Whatever can justify your belief about his mental state has to be something you observed. You cannot directly access his mental state, but rather you need to base it on the evidence you gather by perception. You can listen to his testimony and you can see him get emotional when watching the game. Perhaps even facial expressions showing his mood can be sufficient if you know enough about the situation he is in right now. Crucially, the mental state attribution is based on things other than the mental state attributed. Plausibly they are not completely unrelated – his emotional reaction is connected to the desire that Barcelona FC wins – but they are different things. This is not to say that mental states cannot play any role in mind-reading processes. Rather, the mental state that should be attributed at the end of a process cannot itself be the input of the very same process. If I ascribe mental state M_1 by process P_1 at t_1 I can later on use M_1 in process P_2 at t_2 to ascribe M_2. For instance, I ascribe the desire for Barcelona FC to win the Champions League to Jordi Fernández based on his

[9] Some argue for a distinct method of knowing other people's mental states that is not inferential mind-reading (cf. Spaulding (2015)). I will not consider these options.

behavior while watching a football game. Then later on I see him celebrating after the game finished. I can now use the previously ascribed desire to figure out that he is happy that Barcelona FC won.

Mind-reading approaches that start with behavioral observation plus evidence of the situation come in two varieties: As theory-theory accounts (e.g. Gopnik and Wellman (1994; 2012), Gopnik and Meltzoff (1997)) and as simulation accounts (e.g. Goldman (2006)). Both differ in what is done with the observational input, but for my purpose the focus is on the input itself. So I can work with a very simplified black-box model.



The model can be self-directed, so it is possible to self-ascribe a mental state based on behavioral observation and a grasp of the current situation. A standard example for this type of case is Wright's explanation of a scene in Jane Austen's Emma:

> Emma has just been told of the love of her protégée, Harriet, for her — Emma's — bachelor brother-in-law, a decade older than Emma, a frequent guest of her father's, and hitherto a stable, somewhat avuncular part of the background to her life. She has entertained no thought of him as a possible husband. But now she realizes that she strongly desires that he marry no one but her, and she arrives at this discovery by way of surprise at the strength and color of her reaction to Harriet's declaration, and by way of a few minutes' reflection on that reaction. She is, precisely, not moved to the realization immediately; it dawns on her as something she first suspects and *then* recognizes as true. It *explains* her reaction to Harriet. (Wright, 1998, pp. 16-17)

With self-directed mind-reading in the mix, how can we tell a mind-reading story of Otto's self-ascription? Otto simply looks at the notebook and avows that he believes the museum is at 53rd street. There is no strong, colorful reaction to the notebook that Otto then can interpret in such a way that makes it possible to attribute a belief. The only reaction is the assertion that he believes

that the museum is at the 53rd street. But that reaction already presupposes what the mind-reading process wants to get at. It is useless as a basis for a mind-reading process. What other behavior can be considered as the basis for Otto? Perhaps he can base his mind-reading on previous instances of looking at the notebook. In the past he read the notebook and acted accordingly. But is this enough as a basis? I doubt it. While he might get to the general conclusion that usually he acts according to what is written in the notebook, there is no way this general claim can lead him to the specific attribution of a belief *that p*. Where should the propositional content come from, if his basis is a general claim?

The obvious amendment is to let the content written in the notebook play a role in the explanation. So the general observation of Otto's past behavior plus the fact that *p* is written in the notebook are the input for the mind-reading process. The output is then the self-ascription that he believes that *p* – that the museum is at 53rd street in this case.

But this is no good either. There are red flags for both parts of the input. First, it does not seem obvious whether Otto can use his past behavior at all as an input. Otto has Alzheimer's after all. How can we let any representation of his past behavior play a crucial role in the belief-production if it is unclear whether his memory supports any such representation? If Otto needs a notebook to remember where the museum is, he likely won't be able to remember how he behaved in the past. We can imagine him writing down all his behavior in the past as well, but that seems unnecessary in our story about Otto's self-ascription. The story is complete without him also checking the notebook for his previous behavior.

Second, if we accept that *p* written in the notebook plays a role as input, we might be already stepping away from mind-reading. Remember that the mental state attributed at the end of the mind-reading process cannot be already the input. If we let the notebook play such a pivotal role as input, we are in danger to abandoning this general rule, because the notebook stating that *p* is part of the extended belief that Otto wants to self-ascribe. This means that Otto would effectively use his

extended belief to self-ascribe the very same belief. While this perfectly fits the story, it does not tell a mind-reading tale anymore. Instead we have to deal with Otto directly detecting his own extended belief. He bases his second-order belief on his first-order belief. Suddenly the mind-reading approach no longer seems appropriate.

The proponent of the mind-reading solution can attempt to save his account by implementing a different move. It is not the present behavior that Otto interprets, but rather his past behavior. That there is something written in the notebook is evidence of his own past action of writing down the location of the museum. When Otto recognizes what is written in the notebook, he recognizes that this is something that he wrote and hence, that he believes. In this case he bases his self-ascription on the evidence of his prior action. The first problem with this solution is that it heavily relies on Otto having written that *p* into the notebook earlier. However, as I noted before, the previous endorsement condition is not all that necessary according to Clark and Chalmers. And if there is no previous endorsement condition, then there can be cases in which the writing in the notebook is no evidence of Otto's past actions. Second, in some cases Otto will not be able to tell whether he wrote *p* into the notebook. Just think of notes on a smartphone. There is no handwriting that can be recognized as something Otto himself wrote. Nevertheless, he can have beliefs extended to the notes on the smartphone and be unable to rule out that someone else wrote them. It seems that he does not need to know who put the notes in there. If this is correct, then he cannot use these notes as results from his past actions. Hence, he cannot use them as a basis for self-directed mind-reading.

I showed that both the introspection and the mind-reading approach fail to capture extended introspection appropriately. What are we supposed to do now? I believe there are four options available, with little to go in favor of each.

1) We can change our account of introspection, such that privileged- and peculiar access are less important.

2) We can change our account of mind-reading, such that we can allow the ascribed belief to already be a part of the input in some sense.
3) We can get rid of the idea of extended belief altogether, and stick to our guns for introspection and mind-reading.
4) We can propose that extended introspection needs its own, distinct account of producing self-knowledge.

It is a difficult dialectic position to be in. Moreover, if one looks at the individual debates, then one can find independent motivation for every single of these options. One can accept that introspection does not come with this sort of privileged- and peculiar access in general with Carruthers (2011), Cassam (2014), or Schwitzgebel (2008) and go for (1). One can adapt ideas from direct social perception of mental states such as presented in Krueger (2012) or Spaulding (2015) and go for (2). One can get rid of extended minds with Gertler (2007) or Adams and Aizawa (2010) and go for option (3). I, however, want to opt for (4). The reasoning is largely defensive. If one chooses option (4), one need not change any independently motivated position. Therefore, (4) is the least invasive way to go. I can avoid the internal debates of introspection, mind-reading, and extended mind for the bargain of accepting a new source of self-knowledge: Extended Introspection.

Extended Introspection Sui Generis

Perhaps it should not surprise us that both introspection and mind-reading do not work for our story about Otto. Why expect a phenomenon to fit an explanation that was modeled after very different cases? So let's start looking at the case and build an account from the bottom up.

The main idea is to describe what is going on in the Otto case and then transform this description into a general principle or rule. This approach is not entirely original. Rather, it is a staple in the epistemologist's toolbox to build epistemic rules out of cases, where an epistemic rule is simply a rule of belief formation.

For instance, we can start with the following story taken from Alex Byrne (2005, p. 93):

> Mrs. Hudson might hear the doorbell ring, and conclude that there is someone at the door. By hearing that the doorbell is ringing, Mrs. Hudson knows that the doorbell is ringing; by reasoning, she knows that there is someone at the door.

This case is straightforward. Mrs. Hudson believes that someone is at the door, because the doorbell rings. We can transform this into a rule that Mrs. Hudson follows:

DOORBELL If the doorbell rings, believe that there is someone at the door.

It is easy to see that this rule fits the case. Mrs. Hudson's belief formation can be described as her following this conditional, whereas following the conditional means that she forms the consequent belief *because* she recognizes that the antecedent condition holds. Generalizing this, Byrne (2005, p. 94) states that

S follows the Rule R ('If conditions C obtain, believe that *p*') on a particular occasion iff on that occasion:

a. S believes that *p* because she recognizes that conditions C obtain

Which implies:

b. S recognizes (hence knows) that conditions C obtain

c. Conditions C obtain

d. S believes that *p*

DOORBELL happens to be a good rule, that is, it usually produces true beliefs. On the other hand, you can think of bad rules that produce false beliefs most of the time. For instance, "If you are hungry, believe that it is sunny outside" is a rule that will not generate true beliefs in general. There are simply too many instances in which you are hungry, but it is not sunny outside. In other words, the rule is unreliable.

I can now make use of epistemic rules to get a grasp of what goes on in the Otto case. The case was the following:

> Otto suffers from Alzheimer's disease, and like many Alzheimer's patients, he relies on information in the environment to help structure his life. Otto carries a notebook around with him everywhere he goes. When he learns new information, he writes it down. When he needs some old information, he looks it up. For Otto, his notebook plays the role usually played by a biological memory. I ask Otto whether he believes that the museum is on 53rd Street. Otto looks at his notebook and answers "I believe the museum is on 53rd Street."

Just as I described Mrs. Hudson's belief formation, I can now describe Otto's belief formation as a two-step process. Otto looks at his notebook, and then self-ascribes a belief because of what is written in the notebook. He recognizes that the notebook says that *p*, and transitions, by reasoning, to the belief that he believes that *p*. He thereby fits the following rule:

NOTEBOOK If your notebook says that *p*, believe that you believe that *p*.

Otto believes that he believes that *p*, because he recognizes that his notebook says that *p*. This looks quite similar to Byrne's (2005) general rule BEL: If *p*, believe that you believe that *p* (Byrne, 2005, p. 95). There is a sense in which it is just an instance of BEL, because Otto recognizes that *p* by looking at the notebook. However, I do not opt to use BEL here. My rationale is twofold. First, BEL shows a very strong asymmetry between a first person formulation and a third person version. We already saw that the case of extended introspection does not fit with this asymmetry to such an extent. NOTEBOOK on the other hand captures the close similarity to a third-person rule that fits our initial intuition that extended introspection is not quite as privileged and special as genuine introspection. Second, NOTEBOOK can provide additional insights to the Otto case. Both reasons will become apparent in this section.

So far NOTEBOOK looks just like DOORBELL. However, it is more than the simple DOORBELL rule. NOTEBOOK is a very special rule, if one supposes that the conditions for extended beliefs are satisfied for Otto. Remember conditions (ii) and (iii) that external aids have to satisfy for an extended belief. The external aid has to be (ii) readily accessible, and (iii) automatically endorsed. (ii) plays an important role in making it possible to follow NOTEBOOK. Epistemic rules in general do not say whether they can actually be followed. DOORBELL, for instance, gives you a conditional that provides a transition from recognizing the doorbell ringing to a belief that will likely be true, if the antecedent

holds. But you might not be able to recognize that the doorbell is ringing. You could be deaf, or simply listening to music on headphones on a volume that makes it impossible to hear the doorbell. Nothing guarantees that a good rule is one that you can follow. However, this is different for extended believing Otto and the NOTEBOOK rule. Otto is guaranteed to be able to follow NOTEBOOK reliably with regard to recognizing the antecedent.[10] The argument for this is rather simple:

1) Otto has a belief extending to the content of his notebook. (Assumption)
2) Otto's belief can only be extended if the content of the notebook is readily accessible. (Conditions for Extended Belief)
3) Otto can readily access the content of the notebook (from 1, 2)
4) If the content of the notebook is readily accessible, then Otto can reliably recognize that the notebook says that *p*, if it says that *p*. (Spelling out Accessibility)

---

5) Otto can reliably recognize that the notebook says that *p*, if it says that *p*. (from 3, 4)

The argument shows that Otto is able to reliably recognize that the antecedent of the NOTEBOOK rule holds, if it holds. He can do so in virtue of the extended belief condition that makes the notebook readily accessible. The only step in the argument that is not an assumption or already independently argued for is (4), but I take this to be intuitively true. What else could it mean to be able to readily access a notebook, if not that I can reliably recognize that the notebook says that *p*, if it does say that *p*?[11]

That Otto can reliably recognize that the notebook says that *p* is not enough to guarantee that he can reliably follow NOTEBOOK completely. He further needs to be able to follow the conditional and form

[10] It is important to highlight the difference between a rule being reliable and one being able to follow the rule reliably. The rule is reliable (good) if it mostly produces true beliefs. On the other hand, one can follow an inferential rule reliably if one usually is in a position to follow it. One can reliably follow a reliable rule, but one can also reliably follow an unreliable rule. The same goes for being unable to reliably follow a rule.

[11] In earlier versions I was tempted to read the ready access condition stronger than merely reliable access. However, that would be against Clark and Chalmers (1998) intention of providing a parallel to biological memory. Clark (2010) uses reliable access instead of ready access to avoid this confusion. Thanks to Brie Gertler for pointing this out.

a belief according to the conditional. This is not worrisome at all. As long as Otto is able to reason, he is able to follow a conditional just fine.

So far I established that Otto can reliably follow NOTEBOOK, given that he has beliefs extending to his notebook. However, I still need to provide reasons why NOTEBOOK is actually a good epistemic rule. Why should NOTEBOOK generate true rather than false beliefs? Here I take another condition of extended belief to play the pivotal role. This time it is condition (iii), the automatic endorsement condition. The idea is that whenever Otto looks into his notebook and reads that *p*, he automatically endorses that *p* and thereby is guaranteed to believe that *p*. This can be used in an argument as follows:

1) Otto automatically endorses that *p*, if he reads that *p* in his notebook. (Condition of Extended Belief)
2) Endorsing that p entails believing that *p*. (Spelling out Endorsement)
3) Otto reads that *p* in his notebook. (Assumption)
4) Otto endorses that *p*. (from 1, 3)
5) Otto believes that *p*. (from 2, 4)

---

6) If Otto reads that *p* in his notebook, he believes that *p*. (from 3, 4, 5)

This conclusion shows that NOTEBOOK is actually a good epistemic rule. It is good, because the consequent belief will always be true when Otto follows the rule. Whenever Otto follows NOTEBOOK, he starts by looking at the notebook which says that *p*. Otto recognizes that *p* and automatically endorses it. This endorsement guarantees that he believes that *p*. So when Otto follows the conditional and forms the belief that he believes that *p*, he will be correct. Following NOTEBOOK is infallible, because the mere act of following the rule guarantees the second-order belief to be true by securing the first-order belief.

A crucial step in the argument is (2), which depends on the notion of endorsement in play. Clark and Chalmers (1998) do not provide much information in this regard. However, Clark (2010) says that endorsement means that "It should not usually be subject to critical scrutiny (unlike the opinions of other people, for example). It should be deemed about as trustworthy as something retrieved clearly from biological memory" (Clark, 2010, p. 46). I take this to entail belief, insofar as it is equal to regarding *p* as true. If Otto recognizes that the notebook says that *p*, he holds *p* to be true without additional, critical scrutiny. He will use *p* as a premise in practical and theoretical reasoning, the same as if he would hold *p* to be true based on any other source. This should be uncontroversial, given that I assume that the extended mind thesis is true.

Behind this argument lies a general observation of the Otto case. If Otto self-ascribes a belief by looking at the notebook two usually[12] unrelated factors coincide. On one hand there is the ground for a belief. For instance, I can form a perceptual belief based on a perceptual seeming. I form the belief that the sun is shining, because I have a visual experience of the sun shining. This is my evidence that I base my belief on. However, on the other hand there is an external fact that makes the belief true. My belief that the sun is shining is true, if the sun is in fact shining. This truthmaker is different from the basis of my belief. In the Otto case things seem different. The very same thing that Otto bases his second-order belief on also makes this second-order belief true. He looks at the notebook and believes that he believes that *p* because of the notebook saying that *p*. At the same time the notebook makes it the case that he believes that *p*, thereby making his second-order belief true. This makes it difficult for Otto to be wrong about himself, if he self-ascribes by looking at the notebook. The external aid is both the basis and the truthmaker for his self-ascription. Here even the most determined sceptic cannot find a gap in which to insert his knife – as long as the sceptic is on board with extended beliefs in general.

I established that NOTEBOOK is an epistemic rule that Otto can reliably follow, and moreover a good, truth-conducive rule. However, where does NOTEBOOK put extended introspection with respect to

[12] Even though not always. One might argue that the same thing happens if I form beliefs about my qualia.

privileged and peculiar access? To answer this I want to look at third person equivalents to the NOTEBOOK rule. After all, the prior intuition was that there is nothing special about Otto looking at the notebook compared to me looking at the notebook. So perhaps there is a similar epistemic rule for me. And I believe there is, let's call it O-NOTEBOOK.

O-NOTEBOOK If Otto's notebook says that *p*, believe that Otto believes that *p*.

Under the assumption that Otto has beliefs extended to the notebook, O-NOTEBOOK also looks like a very good rule. If I look at Otto's notebook and recognize that it says that *p*, then it will be true that Otto believes that *p*. It will be true, because by the assumption of extended beliefs whatever is written in the notebook constitutes dispositional beliefs of Otto, just the same way something stored in biological memory would. However, there are some differences to NOTEBOOK. First, O-NOTEBOOK plus the assumption that Otto has extended beliefs does not guarantee that one can reliably follow the rule. Whereas Otto can reliably follow NOTEBOOK in virtue of the extended mind conditions, there is no condition that guarantees me any access to Otto's notebook. Hence, I might not be able to recognize the antecedent of the conditional in a large number of cases.

Second, Otto can ascribe *occurrent* beliefs by using NOTEBOOK, whereas O-NOTEBOOK cannot do the same. The idea here is that whenever Otto looks at his notebook to follow his NOTEBOOK rule he thereby endorses the content right at that moment. That is, the endorsement, and thereby the belief that *p*, plays an active role in Otto's cognitive machinery at the moment of him following NOTEBOOK. Moreover, it is plausible that Otto will be consciously aware of his belief that *p*, when he follows NOTEBOOK. On the other hand, if I look at Otto's notebook, there is no way for me to tell whether Otto believes that *p* occurrently or dispositionally. It is possible that I ascribe to Otto the belief that *p* by looking at his notebook, while at that moment Otto himself does not look at the notebook at all and is thinking about something completely unrelated to *p*. In this case, I can still correctly ascribe a belief that *p* to Otto, but only a dispositional belief. The difference can be spelled out by expanding both epistemic rules:

NOTEBOOK* If your notebook says that *p*, believe that you occurrently believe that *p*.

O-NOTEBOOK* If Otto's notebook says that *p*, believe that Otto occurrently believes that *p*.

NOTEBOOK* is a good rule, whereas O-NOTEBOOK* is not. The former will produce mostly (always) true beliefs, but the latter generates a ton of false beliefs in cases where I follow O-NOTEBOOK* when Otto does not look at his notebook. Hence there is a difference between NOTEBOOK and O-NOTEBOOK insofar as they both use a general notion of 'belief', but have different types of beliefs as truthmakers in general.

With these differences in mind I can confidently say that there is something peculiar and special about extended introspection. But it is only a minor difference. Nothing guarantees that I can reliably follow the third person equivalent to Otto's NOTEBOOK rule. Moreover, even when I can follow O-NOTEBOOK, I cannot employ quite the same method as Otto. However, I can do something in the vicinity, closely resembling Otto's belief formation. And I can be reliable as well; I am just limited in the range of reliable extended mental state ascriptions. I cannot reliably ascribe occurrent states based on Otto's external aids, but I can reliably attribute his extended beliefs in general. The result is somewhere in-between the features that ordinary self-knowledge and mind-reading are said to possess. It is not quite as peculiar as self-knowledge based on usual introspection, but there is still some difference between the first-personal access and the third-personal one.

From Otto to Otto++

So far I have discussed the extended introspection case for Otto and his notebook. Earlier on I remarked that the same strategy that works as an explanation for extended introspection in the case of Otto and his notebook also works for Smart's (2018) case of Otto++, who has information stored online on a personal cloud service that can be accessed via smartphones, smartwatches or augmented reality glasses. Otto++ has sufficient access to the information via these devices to satisfy conditions (i) to (iii). Otto++ poses two interesting challenges. First, can we come up with an

epistemic rule for extended introspection that fits the case? And second, is there still the same kind of privilege involved in this case?

To answer the first question, I propose to use an analog to the NOTEBOOK rule:

CLOUD If your cloud-based data store says that *p*, believe that you believe that *p*.

We have to be liberal in interpreting the phrase "if your cloud-based data store says […]." The idea is that this phrase covers all cloud-based information being presented by any device that is sufficiently connected to that cloud-based store. The antecedent can be satisfied if your smartphone app connected to the cloud-based data store shows you that *p*, but also if the augmented reality glasses tell you that *p*, or a device produces sounds providing you with the information that *p*. In all of these cases the same arguments I provided for the notebook case can be applied in relation to the CLOUD rule. If you have beliefs extended to the cloud-based data store satisfying conditions (i) to (iii), you will be in a position to reliably follow the rule CLOUD. Moreover, you will automatically endorse the content provided by the cloud-based data store and therefore believe that *p* whenever you follow the rule.

Other people will also be able to come to know Otto++'s beliefs by accessing his cloud-based data store. So we get the third-person rule for Otto++:

O-CLOUD If Otto++'s cloud-based data store says that *p*, believe that Otto++ believes that *p*.

However, just as with O-NOTEBOOK, there is no guarantee that I am able to reliably follow O-CLOUD. On the other hand, the conditions for extended belief guarantee that Otto++ can reliably follow CLOUD. Hence, Otto++ is privileged with regard to his cloud-based beliefs just like Otto is with regard to his notebook-based beliefs.

In addition, Otto++ is also privileged with regard to his occurrent beliefs. To see this, consider the rule for occurrent beliefs:

CLOUD* If your cloud-based data store says that *p*, believe that you occurrently believe that *p*.

For the same reasons as in NOTEBOOK* the rule CLOUD* will be a good rule and produce true beliefs, whereas the third-person equivalent O-CLOUD* will not.

O-CLOUD* If Otto++'s cloud-based data store says that *p*, believe that Otto++ occurrently believes that *p*.

Just as in the notebook case, CLOUD* is a good rule, but O-CLOUD* is not. Following O-CLOUD* does not reliably lead to true beliefs. Hence, Otto++ is also privileged with regard to the occurrent nature of his extended beliefs. Overall, Otto++ has privileged access to his extended beliefs similar to Otto in the notebook case, hence I found my answer the first question.

To answer the second question, I want to consider the following scenario. Suppose that the FBI can observe Otto++ constantly by utilizing the technological gadgets Otto++ uses. The augmented reality glasses have accessible cameras built in that are always filming Otto++ and the FBI can use his smartphone to listen in on what he says at all times. Finally, suppose that they use well developed computer programs and well trained agents to infer Otto++'s mental states from the observed behavior. In this scenario the FBI seems to be in an especially good position to know Otto++'s mental states. With all these surveillance tools the FBI might be able to get to know Otto++ even better than Otto++ knows himself. Is this a problem for the claim that extended introspection is privileged?[13] I do not think so. I am willing to grant that someone could come to know *more* about Otto++'s beliefs than he himself knows. And the constant observation via wearable technology seems to be a plausible scenario in which that is the case. However, this does not threaten the privileged access Otto++ has to his extended and non-extended beliefs. We should not understand privileged access as having an especially high number of justified true beliefs about one's mental states. Rather, following Byrne (2005), I suggest that we should understand privileged access as beliefs formed by

[13] Thank you to a reviewer who raised this issue.

introspection being especially likely to amount to knowledge. There is less of a chance for making a mistake when introspecting. This is also the case for extended introspection. CLOUD is a rule that is self-verifying under extended belief condition. If Otto++ follows CLOUD he will end up with a true belief. Moreover, the conditions for having beliefs extended to the cloud-based data store guarantee that Otto++ can reliably follow the rule. Nothing guarantees that anyone else can follow the third-person rule O-CLOUD. One might respond here that O-CLOUD does not seem to be a difficult rule to follow, especially for someone who is able to observe everything that Otto++ does at any given time. But even in that case, at best one could be on par with Otto++ with regard to his dispositional beliefs. Otto++ is still privileged with regard to occurrent beliefs formed by CLOUD*. Otto++'s observable behavior will merely provide you with a reasonable basis to infer Otto++'s occurrent beliefs, but this basis underdetermines his actual mental states. Any behavior that might indicate that Otto++ likely accesses information (e.g. all the mechanical steps that lead to the information being displayed plus his eye movement) could happen without any information being accessed. Otto++ might look as if he accesses information from the cloud when he actually does not. Hence, there is no third-personal path to know whether Otto++ occurrently believes that *p* that can compete with Otto++'s privileged access. Otto++ can use CLOUD* and thereby accurately find out that he occurrently believes that *p*. Hence, there is still room for Otto++'s privileged access.

Conclusion

I showed that a straightforward interpretation of extended introspection as genuine introspection is not appropriate. Furthermore, a self-directed mind-reading story does not fit either. This left us with four distinct options. We can change our accounts of introspection and mind-reading, deny the extended-mind thesis, or propose a sui generis form of extended introspection. I chose the latter. Therefore, I proposed an original account of extended introspection based on epistemic rules. I provided a rule for the Otto example, and argued that this happens to be a rule that Otto can follow and that generates true beliefs. Both are secured by the requirements of extended beliefs. If Otto has extended beliefs, then he can reliably follow the NOTEBOOK rule which leads him to true beliefs

about his mental states. Finally, I showed how this picture fits the ideas of privileged and peculiar access. Otto is privileged, because having beliefs extended to the notebook (plus reasoning) guarantees that he can reliably follow the NOTEBOOK rule. Furthermore, he is in a special position because he can attribute that a belief is occurrent based on NOTEBOOK, whereas a third-person variation of the rule, O-NOTEBOOK, cannot do the same. It is still an open question to what extent this result can generalize to other cognitively integrated artifacts. This is mainly a question of whether (i)-(iii) hold for all extended mental states, or whether some artifact can be integrated enough to count as part of the mind without satisfying (i)-(iii). Plausibly, widespread technology such as smartphones and smartwatches fit these conditions to some extent. In particular I argued that it does generalize to Smart's (2018) Otto++ case, which satisfies (i) to (iii). The other open question is how extended beliefs can be combined with an agential account of self-knowledge. I leave this to future work.

References

Adams, F. R., & Aizawa, K. (2010). Defending the bounds of cognition. In R. Menary (Ed.), *The Extended Mind* (pp. 67-80). Cambridge, MA: MIT Press.

Armstrong, D. M. (1968). *A Materialist Theory of Mind.* London: Routledge & Kegan Paul.

Bernecker, S. (2014). How to Understand the Extended Mind. *Philosophical Issues*(24), pp. 1-23.

Bilgrami, A. (2006). *Self-Knowledge and Resentment.* Cambridge, MA: Harvard University Press.

Burge, T. (1996). Our Entitlement to Self-Knowledge I. *Proceedings of the Aristotelian Society, New Series, 96*, pp. 91-116.

Byrne, A. (2005). Introspection. *Philosophical Topics, 33*(1), pp. 79-104.

Carruthers, P. (2011). *The Opacity of Mind: An Integrative Theory of Self-Knowledge.* Oxford: Oxford University Press.

Carter, A. J., & Pritchard, D. (Forthcoming). Extended Self-Knowledge. In J. Kirsch, & P. Pedrini (Eds.), *Third-Person Self-Knowledge, Self-Interpretation, and Narrative.* Berlin: Springer.

Cassam, Q. (2014). *Self-Knowledge for Humans.* Oxford: Oxford University Press.

Clark, A. (2008). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension.* Oxford: Oxford University Press.

Clark, A. (2010). Memento's Revenge: The Extended Mind, Extended. In R. Menary (Ed.), *The Extended Mind* (pp. 43-66). Cambridge, MA: MIT Press.

Clark, A., & Chalmers, D. J. (1998). The extended mind. *Analysis, 58*(1), pp. 7-19.

Fernández, J. (2013). *Transparent Minds: A Study of Self-Knowledge.* Oxford: Oxford University Press.

Gertler, B. (2007). Overextending the mind? In B. Gertler, & L. Shapiro (Eds.), *Arguing About the Mind* (pp. 192-206). New York: Routledge.

Gertler, B. (2011). *Self-Knowledge.* New York: Routledge.

Goldman, A. (2006). *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading.* Oxford: Oxford University Press.

Gopnik, A., & Meltzoff, A. N. (1997). *Words, Thoughts and Theories.* Cambridge, MA: MIT Press.

Gopnik, A., & Wellman, H. M. (1994). The Theory Theory. In L. Hirschfield, & S. Gelman (Eds.), *Mapping the Mind: Domain Specificity in Cognition and Culture* (pp. 257-93). New York: Cambridge University Press.

Gopnik, A., & Wellman, H. M. (2012). Reconstructing Constructivism: Causal Models, Bayesian Learning Mechanism and the Theory Theory. *Psychological Bulletin, 138*(6), pp. 1085-1108.

Heersmink, R. (2015). Dimensions of Integration in Embedded and Extended Cognitive Systems. *Phenomenology and the Cognitive Sciences, 13*(3), pp. 577-598.

Krueger, J. (2012). Seeing Mind in Action. *Phenomenology and the Cognitive Sciences, 11*(2), pp. 149-173.

Macdonald, C. (2014). In my 'Mind's Eye': introspectionism, detectivism, and the basis of authoritative self-knowledge. *Synthese, 191*(15), pp. 3685-3710.

Menary, R. (2007). *Cognitive Integration: Mind and Cogntition Unbounded.* Basingstoke: Palgrave Macmillan.

Menary, R. (2010). Dimensions of Mind. *Phenomenology and the Cognitive Sciences, 9*(4), pp. 561-578.

Moran, R. (2001). *Authority and estrangement.* Princeton: Princeton University Press.

Nichols, S., & Stich, S. (2003). *Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds.* Oxford: Oxford University Press.

Rupert, R. (2004). Challenges to the Hypothesis of Extended Cognition. *Journal of Philosophy*(101), pp. 389-428.

Schwitzgebel, E. (2008). The Unreliability of Naive Introspection. *Philosophical Review, 117*(2), pp. 245-273.

Smart, P. (2018). Emerging Digital Technologies: Implications for Extended Conceptions of Cognition and Knowledge. In A. J. Carter, A. Clark, J. Kallestrup, O. S. Palermos, & D. Pritchard (Eds.), *Extended Epistemology* (pp. 266-304). Oxford: Oxford University Press.

Spaulding, S. (2015). On Direct Social Perception. *Consciousness and Cognition*(36), pp. 472-482.

Sterelny, K. (2004). Externalism, Epistemic Artefacts and the Extended Mind. In R. Schantz (Ed.), *The Externalist Challenge* (pp. 239-254). Berlin: De Gruyter.

Sterelny, K. (2010). Minds: extended or scaffolded? *Phenomenology and the Cognitive Sciences, 9*(4), pp. 465-481.

Sutton, J. (2006). Distributed cognitions:dimains and dimensions. *Pragmatics and Cognition, 14*(2), pp. 235-247.

Sutton, J., Harris, C. B., Keil, P. G., & Barnier, A. J. (2010). The Psychology of Memory, Extended Cognition, and Socially Distributed Remembering. *Phenomenology and the Cognitive Sciences, 9*(4), pp. 521-560.

Wright, C. (1998). Self-Knowledge: The Wittgensteinian Legacy. In C. Wright, B. Smith, & C. Macdonald, *Knowing our Minds* (pp. 13-45). Oxford: Oxford University Press.