

# Is cognitive extension necessarily cognitive augmentation?

Madelief Elisabeth Hunting

S4057937

Supervisor: Marc Slors

7 August 2020

Thesis for obtaining a "Master of Arts" degree in philosophy

Radboud University Nijmegen

I hereby declare and assure that I, Madelief Elisabeth Hunting, have drafted this thesis independently, that no other sources and/or means other than those mentioned have been used and that the passages of which the text content or meaning originates in other works - including electronic media - have been identified and the sources clearly stated.

Place: Zwolle Date: 7 August 2020

# Is cognitive extension necessarily cognitive augmentation?

## Abstract

The research on extended and integrated cognition is all about how resources outside of our brain can improve cognition tremendously. I argue that cognitive extension is not necessarily cognitive augmentation. Although extended cognition tends to be conceptualized as something that always enhances cognition that is not necessarily the case. Various cognitive processes do not necessarily augment. However, we barely hear anything about extended or integrated cognitive processes that are not augmentations. The problem might lie within the fact that these non-augmenting extended or integrated cognitive processes are not accounted for by the complementarity principle, which specifically is about augmenting or adding functions to the cognitive process. Thus, I propose a reformulation of the complementarity principle, namely the transformation principle. This new principle can account for any extended and integrated cognitive process regardless of its normative status while preventing cognitive bloat. To support my claim, I provide several examples and focus on social self-deception, anger issues and psychiatric disorders to show that these are examples of extended cognition that are not accounted for with the complementarity principle. My core point is this: if you find extended or integrated cognition plausible, then attenuated extended or integrated cognition should be part of it.

For the past couple of decades, philosophers have researched how cognition can extend beyond the brain. In this debate, neural or even organismic boundaries are seen as arbitrary stopping points of cognition. Instead, elements like tools, other people and even social institutions can be part of the cognitive process. Research on this took off when Clark and Chalmers wrote a paper in 1998, arguing not only that surroundings can help the cognitive process, but also that sometimes the extra-cranial can be a constituent part of cognition. Their famous example of extended cognition involves Otto and Inga, who try to find the Museum of Modern Art: Otto and Inga agree to meet at the Museum of Modern Art at noon. Inga remembers the way. Otto finds the museum thanks to his external memory: a notebook in which he wrote directions since he cannot find the way on his own due to incipient Alzheimer. Both 'remember' the museum to be on 53rd street. The only difference

between Inga's memory and Otto's memory is the location of that memory. Otto's memory is in his notebook, Inga's memory is in her brain. Clark and Chalmers claim Otto's notebook extends his mind because it plays the same functional role for him as Inga's biological memory does for her.

The recurring theme in the field of extended cognition is that resources outside of our brain improve cognition tremendously. It barely matters what kind of extended cognition is at play, whether it be truly extended or merely situated, scaffolded or integrated (or some other variant). For instance, our memory can be expanded through notebooks and lists, as the example of Otto demonstrates (Clark and Chalmers 1998, Rowlands 1999; Michaelian and Sutton 2013). And parents can emotionally and attentionally regulate their babies. Krueger (2013) and Varga (2016) show how physical interventions by caregivers can regulate attention and emotion in ways that the baby would not be able to do on her own. For example, parents can soothe their child through gestures, facial and bodily movements, vocalizations and so on. Thanks to them, infants can achieve a lot more, cognitively, than on their own. The parent's actions “enhance and extend the infant's cognitive competence” (Krueger 2013, p.40). Furthermore, Gallagher (2013) shows (with his enactive approach to the extension of cognition) that cognition can be a dynamic process between agents and their environments. Social practices, like social interactions but also institutional structures and norms, can support, constitute and extend cognitive abilities. For example, a judge can only give the right judgements because they are coupled in the right way to the institutional structure of the legal system. Moreover, these legal judgments are only intelligible within the legal system. Also, when considering not just individual agents in systems, but systems as a whole, one can find cultural-cognitive ecosystems. Systems can be comprised of people, manuals and tools, connected in such a way that they can do things like navigating ships through treacherous seas or fly a plane (Hutchins 1995). Other examples are public language, computers, calculators, infrastructures of communication, and various

'arts of memory', techniques that aided recall which were popular during the medieval and Renaissance period (Donald 1991, Sutton 2010). In all these examples, extension is augmentation. The sky seems to be the limit when you look at the literature on extended cognition.

However, my question is: Is cognitive extension necessarily cognitive augmentation? I think not. We already know about various cognitive processes that do not necessarily serve us well. After all, people's cognitive processes do not always achieve their goals. Nor is cognition always efficient, appropriate or moral. From deceiving yourself that this will be your last drink for the night (Deweese-Boyd 2017), to being accidentally racist when hiring a new employee because of implicit bias (Brownstein 2019) or confabulating some story about why you did what you did last night (even though you do not remember much) (Hirstein 2009). Human life is riddled with these types of cognitive processes. Even the most famous example of extended cognition stars someone whose cognition is failing him, namely Otto. His cognition is deteriorating; his neural memory is failing him. Even though attenuated cognitive processes are common, we barely hear anything about extended cognitive processes that are less than desirable. The reason is that, in the abundance of extended cognition research, the focus is usually cognition that goes wrong inside the head. Up until now, if a cognitive process goes awry and involves some resource outside of the brain, that outside element is usually not considered cognitive at all, but rather some non-cognitive, causal influence. Even though those resources might be just like Otto's notebook, a baby's parents or the legal system.

Consider a malfunctioning navigation system in a car or a messed up digital calendar. In the case it functions, some philosophers would claim that they are part of extended cognitive processes (Gallagher 2013, 5; Heersmink 2016, 437). So, this should also be the case when they malfunction. And while we can decouple from those cognitive artefacts easily, there are extended cognitive processes we cannot or do not get rid of that easily even though they are not beneficial. For example, anger

issues (Cooper 2017) due to the environment, worsened psychiatric disorders because of certain artefacts (Hoffman 2016) and social self-deception (Dings 2017), where socially extended cognitive processes enable the self-deceptive process. I will demonstrate that these types of examples are extended or integrated cognitive processes. That is why my research is not a refutation of the field of extended or integrated cognition, but rather an addition. My claim here is simple: if you find extended or integrated cognition plausible, then attenuated extended cognition should be part of it.

To get some sense of attenuated extended cognition, first I will show how the literature on extended cognition fixates on augmented cognition. It is important to note here, due to limited space, this paper only aims to analyze attenuated integrated and extended cognition. However, the ideas discussed in this paper may apply to more theories of cognition, like distributed, scaffolded, situated cognition and others.

Next, I argue that some of the classic examples of augmented cognition might not be augmentations at all. Then the problem will become clear: these examples are not accounted for by the complementarity principle, even though, intuitively, they would fall within the field of extended cognition. I shall offer some frameworks to differentiate extended cognitive processes that are augmentations from those that are attenuations of cognition. And I will propose a reformulation of the complementarity principle, namely the transformation principle, which includes both attenuated, augmented and neutral types of extended and integrated cognition. I shall provide several examples to support this new principle, but I will focus on social self-deception (Dings 2017), anger management and anorexia to show that they are examples of extended cognition that are not accounted for by the complementarity principle even though they should be. The transformation principle fixes this lacuna. Lastly, I will refute some objections, mainly by explaining

why philosophers have been preoccupied with augmenting extended cognition in the first place.

### **1. Preliminaries**

Before we can dive into the main argument, it will be helpful to consider some preliminaries. What exactly is extended or integrated cognition and what makes something cognitive? Roughly, proponents of extended and integrated cognition assume that cognition can (on occasion) extend, or be integrated with, factors beyond the brain of a cognitive agent to include elements from their environments.

The challenge all versions of extended cognition must overcome is how to determine whether environmental elements are part of cognition or not. Not everything we interact with is part of our cognition. It would stretch the concept of cognition unrealistically far, rendering it pointless. This is the issue of delineation between the cognitive and the non-cognitive, which is best explained with an example: Otto's notebook grants Otto information, but if everything that grants us information is part of the cognitive system, then the internet and the eight o'clock news are also part of an individual's cognitive system. This would cause 'cognitive bloat' (Rupert 2004) or an overextension of an individual mind. Consider an example: suppose Trisha hears from the BBC that Trump resigned. Does her belief in the proposition "Trump resigned" include the news, the news anchor or the television she watches the news on? What about the miner who mined copper for the electric wiring in the television? It does not make sense to include the miner into Trisha's cognitive process and it is very questionable whether the news, the news anchor and the television are part of her cognition. There has to be a distinction between cognitive systems and their surroundings. How else can one tell the difference between causal influences and constitutional parts of cognition? The miner, the news or the notebook? If it is impossible to tell the difference, one falls victim to a so-called causal/coupling fallacy (Adams & Aizawa 2001). Where a causal relationship between cognitive artefacts and agent is being confused for a coupled system

between them.

The original method to tell the difference between a causal influence and a constitutional part of cognition is the parity principle (Clark and Chalmers 1998):

“If, as we confront some task, a part of the world functions as a process which, were it done in the head, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world is part of the cognitive process.” (1998, p.8)

In other words, the process outside the head must be functionally the same as a process inside the head. Furthermore, the external resource needs to be accessible, endorsed and reliable, like Otto and his notebook. In other words, there must be 'glue and trust' between the agent and a part of the world.

However, there are some problems with the parity principle. One of them is that the parity principle does not apply to the classical example of Otto. Grabbing a notebook, leafing through it and reading is not the same as the process of remembering in the brain, which involves little or no bodily movement and no reading. To deal with this problem, Sutton proposed an alternative principle, the complementarity principle (2010, 194). This principle centres on integration rather than parity. There are a lot of formulations of this principle, but they are all a variation on the same theme. Here, I will use a recent version<sup>1</sup> by Heersmink (2015):

“On a complementarity view, artefacts or other resources do not just augment existing brain functions by externally replicating them, but add new

---

<sup>1</sup> The transformation principle is similar to Menary's 'transformation thesis' (Menary 2007). Menary states that his thesis is about integrated cognition as “[a] coordinated process [that] allows the organism to perform cognitive tasks that it otherwise would be unable to; or allows it to perform tasks in a way that is distinctively different and is an improvement upon the way that the organism performs those tasks via neural processes alone” (p.28).

functionalities to existing ones by integrating them (in varying degrees) into a plastic neural architecture.” (581-582)

This means that tools and resources can augment or add new capabilities to the brain, but they need not mimic or replicate any inner state or process, which makes the complementarity principle more inclusive than the parity principle.

Furthermore, because the integration of artefacts is always a matter of degree, this definition is more nuanced than the parity principle with regards to whether the cognition is truly extended by or merely integrated with cognitive artefacts (2015, 595). As a result, Trisha is not connected to every single artefact when watching the news, which would have overloaded her cognitive capabilities. Rather, every single cognitive artefact can have a different degree of integration, from some integration to no integration at all. Heersmink has collected eight dimensions from a variety of papers (among them are Sutton 2006, 2010; Wilson and Clark 2009; Menary 2010a; Sterelny 2010) to measure the degree of integration between agents and cognitive artefacts. The eight dimensions are: (i) “the kind and intensity of information flow between agent and scaffold, (ii) the accessibility of the scaffold, (iii) the durability of the coupling between agent and scaffold, (iv) the amount of trust a user puts into the information the scaffold provides, (v) the degree of transparency-in-use, (vi) the ease with which the information can be interpreted, (vii) the amount of personalization, and (viii) the amount of cognitive transformation” (Heersmink 2017, 433-434). I will pivot back to these later in section three with a more in-depth explanation.

Because of these eight dimensions, the complementarity principle works for both extended cognition and integrated cognition. However, there is a difference between the extended and integrated. Both views endorse the idea that cognition is not just an insular brain process, but always something that is made in tandem with the world. However, extended cognition goes a step further than integrated cognition and claims that cognitive artefacts can co-constitute cognition. It usually takes a synchronic view, where artefacts become material components of the mind at a point

in time. Cognitive integration, on the other hand, emphasizes a diachronic way of looking at cognition; that is, it understands cognition as a process (Menary 2013). Cognitive artefacts are active components or practices. Some examples of these cognitive practices are: (standard protocols for) coordinating actions amongst a group, solving a problem by talking through it, brainstorming together, or physical manipulation of tools, guided by spoken instructions (2013, 29). These practices are dynamic and embedded in a culture (Menary, 2012; Menary & Kirchhoff, 2013). Integrated cognition aims to explain enculturation, or how culture is endowed to us by cognitive practices and how these transform basic biological capacities, allowing us to complete cognitive tasks in ways unenculturated brains will not allow (Dehaene, 1997, 2009; Menary, 2010a, 2010b, 2013), almost like a kind of reverse parity principle. Also, because the practices are integrated with and not constituting cognition, there is no problem of cognitive bloat.

Whether a cognitive process is integrated cognition or extended cognition depends on the degree of integration. If the integration between resource and agent is dense, it is an extended cognitive process, like Otto's notebook. If the density of the integration is shallow, the artefact is part of an integrated (or embedded) cognitive process. While the paradigm cases are clearly deeply integrated or barely at all, there are a lot of in-between cases, since there is no clear tipping point (Heersmink 2014, 595). Alternatively, if one does not agree with the extended cognition thesis but does think the thesis of cognitive integration is valid, it can also be considered a deeply integrated cognitive process. Or vice versa, when one does not agree with the integrated cognition thesis but does think the thesis of cognitive extension is valid, the process can be considered extended. This does not change the point of my paper. The most important takeaway here is that cognitive integration is necessary but not sufficient for extended cognition. This also means that if cognitive integration can be attenuating, this applies to extended cognition as well. For the sake of brevity, from

here on out, I shall refer to both extended and integrated cognition as 'extended cognition'.

## **2. Normative status**

Now let us return to my main point. I do not have a radical view. I only claim that cognitive extension does not necessarily mean cognitive augmentation. I will give some examples which will demonstrate attenuating cognitive extension and how one can tell the difference between attenuating and augmenting extended cognitive processes. By making the normative status of a cognitive process explicit, it will become clear that the literature on extended cognition is biased towards augmentation (and in some ways even excludes attenuating extended cognition). Then, with the example of a kettle, I will demonstrate how to find where problems in an extended cognitive process can arise.

The examples of extended cognition in the last section fit neatly with Heersmink's theory of augmentation and enhancement when I present them as I just did. Despite this, there are also ways to present the examples in a less flattering light, as cases of attenuating cognitive extension. Let us turn some of the classical cases of augmenting extended cognition on their head.

For instance, on the one hand, when calculating a sum with pen and paper or writing a to-do list, one enhances memory and abstract thinking. However, on the other hand, these cognitive artefacts can become 'cognitive clutter': a messy desk full of books, notes and tools can be overwhelming. An extended cognitive system can function worse than it does when it is non-extended. More bells and whistles do not always help an agent (Wilson 2014, 23).

There are other ways cognitive artefacts can hurt the function of our memory too. When an agent regularly couples with a GPS device to navigate, she reduces the inborn ability to do it herself (Javadi et al. 2017). While in the short term the

cognitive process improves, the attenuation of cognition is usually long term. It becomes harder to remember how to navigate. It is a 'use it or lose it' situation. Worse, the same ability to extend memory can similarly hinder the ability to change and evolve. Eichhorn (2019) argues that forgetting life-events plays a crucial part in personal development because it allows people to move on from the past. Not only on a personal level but also socially. This can be seen on social media and platforms like YouTube, where people are pinned to statements they made years ago. Even though they do not endorse the opinions they had anymore, they are still online and for everyone to find. Other types of cognitive extension can be harmful in other ways as well, Menary's example of group brainstorming (2013, 29) can also be an example of attenuated cognition instead of augmentation. Although great ideas can form during brainstorming sessions, the dangers of groupthink are also present (Paulus 1998). A phenomenon where the desire for harmony causes irrational or dysfunctional decision-making outcomes.

Even Plato (c. 360 B.C.E.) already worried about these types of cognitive processes in the Phaedrus:

“If men learn [to write], it will implant forgetfulness in their souls; they will cease to exercise memory because they rely on that which is written, calling things to remembrance no longer from within themselves, but by means of external marks. What you have discovered is a recipe not for memory, but for reminder. And it is no true wisdom that you offer your disciples, but only its semblance, for by telling them of many things without teaching them you will make them seem to know much, while for the most part they know nothing, and as men filled, not with wisdom but with the conceit of wisdom.” (274c-275 b)

Plato argues that books and writing will not augment memory, but rather attenuate internal memory. While it may not be as cut and dried as Plato suggests, it does

show that it is possible cognitive artefacts can be attenuating rather than always augmenting.

Now, the examples of the notebook (or writing in general), to-do lists and GPS devices suddenly do not fit into the complementarity principle anymore, since it demands that external artefacts augment and add abilities. These artefacts seem to do the opposite. This begs the question: what exactly is meant by augmenting cognition? What does Heersmink mean when he says cognitive abilities can be added? I think cognition can only be judged to be augmented relative to a goal or normative standard of, for example, 'normal functioning'. Questions to consider here are: Is the extended cognitive process augmented from the perspective of the individual with the extended cognition? From the perspective of the group they are in, or society as a whole? Morally? Culturally? Economically? Evolutionary? All of the above? Heersmink does not specify. Instead, he only provides examples (of which I have mentioned some) of augmented cognition<sup>2</sup>. The examples surely seem like augmentations at first glance, but only at first glance.

Heersmink has not provided a normative standard by which we can decide whether an artefact augments something or not. So, how can we tell if extended cognitive processes are augmentations or not? And how can we fit attenuated extended cognition into the whole extended (and integrated) cognition framework? To fill this normative lacuna, two things need to happen: first, I will distinguish the cognitive process from its normative status (either attenuating or augmenting, in this case) and then I will amend the complementarity principle (and the corresponding eight dimensions)<sup>3</sup>.

---

<sup>2</sup> Heersmink has thought about some of these questions since the publication of this paper. In 2016 he published a paper about the moral aspects of cognitive artefacts. However, he has not changed his mind regarding the complementarity principle, which remained unchanged in his June 2020 seminar on the extended mind and artifactual autobiographical memory.

<sup>3</sup> Another candidate that I considered to include attenuated extended cognition in the field instead of the complementarity principle was distributed cognition, but that theory answers different questions. Distributed cognition is merely a perspective, while my claim is stronger: there either is a cognitive system that extends

First, I will consider several kinds of normative standards one could use to differentiate augmented extended cognitive processes from attenuated ones. For example, let us look at the moral consequences of an extended cognitive process such as socially extended self-deception. Self-deception is minimally defined as “(i) a process that originates in (ii) a motivation or intention, which leads to (iii) a self-deceived end state (which can be the formation of a novel belief or the maintenance of an existing belief or other attitude)” (Dings 2017, 17). In the case of solitary self-deception, these events take place within a person. By contrast, socially extended cognition happens when the cognitive artefacts that are part of the cognitive process are other people, social practices or institutions like customs of social media. The case of social self-deception involves setting up social reality in such a way that the agent only needs to make a small step to form a self-deceptive belief (Dings 2017, 20). Simply put, social self-deception occurs when others are, or one other person is, indispensable in the cognitive process of self-deception. This means that the process necessarily involves (a lack of) behaviour, body language or facial expressions from others. The self-deceiver uses all of this (intentionally or unintentionally) to deceive himself. For instance, it can involve influencing others to do or say (or refrain from doing or saying) something, which can set up self-deception, by using the influenced people as sources of evidence. Dings notices this on social media<sup>4</sup>. He illustrates this online deception with the example of David. David posts only his best selfies from a vacation in Spain, posts them on social media and gets great feedback. He might not be handsome, but because of the response he receives and the customary practice of providing compliments on social media, he forms the belief that he is.

---

beyond the brain or a person's cognition is integrated with other artefacts. Also, the complementarity and transformation principle focus on one agent (even if it is not necessarily the locus of control in the cognitive process), contrary to distributed cognition, which is used for supra-individual systems like groups of people. This also affects the way it is normatively evaluated and there is not enough space here to say something sensible about that.

<sup>4</sup> Research has shown (Caspi & Gorsky 2006) people agree there is a lot of deception online, but they do not partake in online deception. This might be explained by self-deception. People do not know they are deceiving themselves (and, as a side-effect, sometimes deceive others).

Social self-deception can attenuate cognition in various ways. It can be “a threat to moral self-knowledge, a cover for immoral activity, and a violation of authenticity”. (Deweese-Boyd 2017). Furthermore, it can facilitate harm to others and oneself (Linehan 1982); undermine autonomy (Darwall 1988; Baron 1988); corrupt your conscience (Butler 1726); violate authenticity (Sartre 1943); weaken the capacity for compassionate action (Jenni 2003); violate an epistemic duty to properly ground self-ascriptions (Fernández 2013); or violate a general duty to form beliefs that “conform to the available evidence” (Nelkin 2012)<sup>5</sup>.

There is a philosopher that specifically focusses on the normative status of extended cognitive processes (or in his case, situated affectivity). Slaby (2016) specifically zooms in on the consequences for individuals and collectives when considering extended cognition (or situated affectivity in his case). He says: “[Situated affectivity] is either conducive to an individual or collective and enables flourishing, ...[or it is creating] unhealthy dependencies, [that] tie us to oppressive routines, sustain inequality, destroy[s] communal bonds or lead[s] to affective, and other mental habits that are detrimental to us or our kin” (Slaby 2016 p.11; see also Honneth, 2012, ch. 9 and 10).

Other consequentialist ways to categorize cognitive processes are to make a difference between neural, bodily and environmental consequences, long-term or short-term effects, et cetera<sup>6</sup>. This paper will not be the end all be all regarding the

---

<sup>5</sup> Of course, self-deception can be an augmentation as well, when regarding other standards. David might feel better about himself and see himself in a better light thanks to his selfie self-deception. Since self-deceptive and inflated views of ourselves and our prospects (‘positive illusions’) can cause psychological and physical wellbeing and social advancement (Taylor & Brown, 1994; McKay & Dennett, 2009). However, that does not mean that extended cognition cannot be attenuated.

<sup>6</sup> There are surely more ways than the ones I have just demonstrated. One can judge the intention of the agent initiating the extended cognitive process to find the normative status or use utilitarian, deontic or virtue ethics and more. In the future, evolutionary or teleofunctionalist frameworks might even be used to differentiate good from bad cognition, since most of the extended cognition literature is based on functionalism. Teleofunctionalism gets its norms from evolutionary arguments. For example, a heart works well because it pumps blood. But most, if not all, cognitive processes are too complex to know whether they are truly fitness-enhancing or decreasing. Most cognitive processes can be explained in both ways, as I have demonstrated in this section and the footnote before this one.

normative status of extended cognitive processes since this is a notoriously contested subject. The point of this section is to show that not all extended cognitive processes need to be augmentations and highlight various ways in which it is or is not. For present purposes, it does not matter which framework one utilizes, but only that there are normative standards by which you can measure the functioning of extended or integrated cognition. This means readers can apply any framework they fancy to establish the normative status of an instance of extended cognition because my approach does not depend on a specific normative framework<sup>7</sup>. The only two conditions that the framework needs to fulfil are that it is in line with the current findings on the subject matter and that it is specific enough to bring into sharp relief the elements of the phenomenon that are most pertinent to addressing one's particular inquiry.

Now, the analysis of attenuated extended cognitive processes can continue by further specifying the cause of the attenuation. In other words, where in the extended cognitive process there is a 'failure', as Cooper (2017) calls it. The failure can be in the brain or body of the agent or the cognitive artefacts. Alternatively, there might be a mismatch between the two.

To explain failures, Cooper employs the example of a kettle. If one plugs in a kettle but it fails to boil water, why is the kettle not working? In a multiple-component system, the problem might be with the kettle, but it could also be a problem of the

---

<sup>7</sup> However, what is important when considering a normative framework, is whether one uses an extended or integrated view on cognition. Is the cognitive process, a person, or a system augmented or attenuated? When analyzing with a personal view of cognition, it might seem that artefacts allow us to complete cognitive tasks better in one way or another, but what happens is that one cognitive task is transformed into another. Norman (1991) illustrates this well: "When pilots use a checklist rather than their biological memory, they make less mistakes. According to the system view, pilot-plus-checklist is an enhanced memory system because it is more reliable. However, according to the personal view... the task is transformed from retrieving items from biological memory into retrieving items from the list. On this view, the list is not an enhancer because it does not improve biological memory but merely transforms a task". This of course is but one framework by which to decide whether something is enhanced or not. When considering integrated cognition, the person's mind might not be augmented, but the integrated cognitive process can still be done better than non-integrated with regards to some external consequence.

socket (or maybe even composition of the atmosphere or the water). To locate the problem, one could try plugging the kettle into another socket and see if it works there. Or one could try to plug another kettle into the original socket. If the second kettle boils water without a problem, this could mean the original kettle is broken. But this is not necessarily the case. There could be a problem of 'misfit'. Maybe the original kettle has a two-prong plug but the socket has three holes. This would mean both the socket and the kettle could work, but just not together. To know for sure whether the original kettle is functional, you would need to find a socket in which it works.

In the same vein, this procedure can be used to find the cause of the attenuation or failure in an extended cognitive system. Take anger issues in prison (Cooper 2017, 299-300). Zelda learnt to control her anger by walking away. She usually uses her body (kettle) and her environment (socket) to regulate her emotions (boil water). But this does not work when she is spending a sentence in prison. Every time Zelda encounters Anna during lunch, Anna really rubs her the wrong way. They get into fights regularly. One could say Zelda has anger issues, but she might not have had those up until she got locked up. Before, Zelda walked away when she felt anger bubbling up, but that is not possible when she is imprisoned. There is nowhere to go. It is the system (or integration) of prisoner plus prison that is the problem here. Or, as Cooper states, when people function satisfactorily in a better environment, concluding that the problem is within the person might not be fair (p.303).

However, this does not mean that the environment is always the problem. Just like there is nothing wrong with either a two-pronged plug or a three-pronged socket: They just do not fit together. The problem can be a mismatch between agent and environment; It is a system failure. While this may be true, Cooper warns not interpret this too permissively. Saying there is nothing wrong with an electric kettle that does not work on electricity but does boil water by placing it in a fire, does not seem right. However, plugging the kettle into another socket is a reasonable

environmental shift. There is some limit to the range of 'test' environments that are appropriate<sup>8</sup>. Therefore, placing the kettle in the fire is 'cheating'.

In short, there is a difference between the extended cognitive *process* and the normative *status* of that process. There are several normative frameworks through which we can gauge the functioning of a given cognitive extension (or integration) as augmenting or attenuating. If the process is attenuated, one can evaluate the parts of the cognitive process and if the parts 'match' to find out where the cause of the attenuation is.

### **3. The transformation principle**

To account for all normative types of extended or integrated cognitive processes and not just augmentations, I propose a reformulation of the complementarity principle, the 'transformation principle':

Artefacts, people or practices can change existing brain functions by externally replicating them, but also add new functionalities to existing ones, or lead to dysfunctionalities, by integrating them (in varying degrees) into a plastic neural architecture.

Like all definitions of extended cognition, this definition too needs to overcome the challenge of the coupling-constitution fallacy. When are external factors part of an extended cognitive process and when do they merely hinder or augment a cognitive process without being part of it? There are several ways to go about this, Clark's (2010, 46) glue and trust criteria can apply here. Kaplan's mutual manipulability criterium (2012) works too<sup>9</sup>. But here I will stick to Heersmink's eight dimensions of

---

<sup>8</sup> Cooper does not specify what the limit of the range of 'test environments' is. But this lacuna can be filled with a normative framework of choice.

<sup>9</sup> I refer the reader to Kaplan's paper (2012) if they want to know more about the mutual manipulability criterium.

cognitive integration because they work and contain the glue and trust criteria. The dimensions still function because augmented extended cognitive processes have the same structure as attenuated extended cognitive processes, as I have demonstrated in the last section. Some adjustments are necessary though because it was originally only formulated for augmented cognition. To illustrate this, I will highlight one example of attenuated cognition each for every dimension and in addition, I will analyse David's case of social self-deception for all dimension.

Before we can begin, a quick and dirty analysis of social self-deception with the transformation principle is necessary. David deceives himself by believing he is handsome, by posting only his best selfies on social media and believing feedback from others that he is in fact, handsome. Extra-cranial factors, like social media and the people on it, can change David's existing brain functions by externally replicating them and integrating them into his cognitive system, which changes both his neural architecture and his cognitive processes. In David's case, he can self-deceive internally (in his brain only), but also externally through other people on social media. If David extends his cognition or integrates it with cognitive artefacts, he can gain new functionalities, like more ways to deceive himself. But he can also gain dysfunctionalities. Here, the function of gaining self-knowledge or the functionality of having a relatively accurate self-image is distorted. That is not to say our self-knowledge is perfect or even good, just that it can become (more) distorted when it is extended by or integrated with cognitive artefacts.

Now, on to the in-depth analysis with the dimensions of cognitive integration (Heersmink 2015). I will extensively cover the dimensions of information flow, trust, individualization and transformation. The other four dimensions, durability,

reliability, procedural and informational transparency will be covered in a footnote since they are less relevant or fine as they are<sup>10</sup>.

The first dimension is information flow, which agents can attain through epistemic actions with an artefact, person or practice. Epistemic actions “improve cognition by changing the input to an agent's information-processing system ... making certain cognitive tasks easier, faster, more reliable or possible at all”. (Heersmink, 2015, 583) Similarly, I want to add, cognitive artefacts, people and practices can also slow down tasks, make them harder, less reliable or even impossible. David's social self-deception is a multi-step process with several types of cognitive resources (a phone, a camera, software, people, etc.) and interactions (a website and people, etc.), both social and solo, where information can flow reciprocally. Furthermore, every step is crucially dependent on the steps before it and the steps can improve the next iterations of this social self-deception. Or, in other words, David can perfect his photo-game through the feedback he gains.

The dimension of information flow is also relevant in certain attenuated cognitive processes like those that can be found in psychiatry. The transformation principle is

---

<sup>10</sup> For the sake of completeness, I will review the other four dimensions here. [i] Durability is about the duration of processes and the quantity. David's social self-deception is probably a repeated cognitive process with a variable duration. [ii] Procedural transparency is reached when using a cognitive artefact does not demand conscious effort thanks to training and habituation. The use of social media is probably very habitual for David. Imagine him automatically reaching for his phone, selecting the app for social media when he is on the couch, toilet or when he is waiting, mindlessly scrolling and liking posts. [iii] Informational transparency concerns the ease with which the agent understands the information presented by the artefact. Considering David's social media use is a repeated process, the apps, programs or sites and probably even people used by him are quite informationally transparent to him. [iv] Reliability is about the physical reliability of the cognitive artefact. Through its weight, size and design, a cognitive artefact can fulfil its function. For example, Otto's notebook is usually more reliable than a phone, since it has no battery and cannot get viruses (Heersmink 2015, 586). Similarly, David's social self-deception is not a very reliable process, because a lot of steps can go wrong. First, he needs to have access to a device that can make photos and a device that can store and upload the photos to the internet. Those devices can lose power or get viruses. Additionally, he needs other people that see the photos and respond in a very specific way. However, something that makes it a tad more reliable is that elements like photos and social media, which are necessary for the deception, can be quite portable. Furthermore, part of this dimension is dependent on the cognitive profile of the agent. The need for reliable access to the artefact can vary. Otto, for instance, is very reliant on his notebook as an extension of his memory. In the same vein, research (Burrow & Rainone 2016; Coulthard & Ogden 2018) has shown that regular selfie posters, like David, may depend on social media feedback for their self-esteem. This would increase the need for reliable access to social media for David. Even if it is not the most healthy way to regulate self-esteem.

a natural fit since it is known that mental health is affected by the psychosocial and physical environment of the mentally ill person, both in good and bad ways. The complementarity principle cannot account for this because it does not allow attenuated extended cognition, like the more inclusive transformation principle does. There already have been some steps to argue for the extended nature of psychiatric disorders: both in diagnosis and treatment. In particular, Hoffman (2016) uses the parity principle to argue for extended psychiatric disorders, but as discussed in section one, there are some problems with this principle. But since the transformation contains the parity principle and expands upon it, the examples still apply. Consider Mona, who is suffering from anorexia. Her fear of gaining weight, a criterion for anorexia is “partially realized by aspects of [her] environment like fashion magazines, computer screens containing media images equating beauty with thinness, or verbal expressions of the values of close companion” (p.1167). So, information flow to her can cause symptoms of her anorexia. Furthermore, her diary can make her feel worse by retaining negative affective states, which can be part of a psychiatric diagnosis. It is easy to imagine that Mona's diary is a part of her extended memory of her self-image: “The diary is not just a causal prop; as a nearby, accessible, reliable, and trusted resource, it serves to retain her standing feelings” (p. 1167). Without her diary, she may even forget those feelings, however temporary, and more easily cultivate positive feelings towards herself. In an analogous way “a mirror may realize a preoccupation with a bodily defect by retaining that preoccupation. Without the mirror, the preoccupation may dissipate more easily” (p. 1167), just like hand sanitizer may retain a compulsion to wash your hands obsessively, by being a constant reminder of this obsession<sup>11</sup>. These cognitive artefacts are part of the cognitive process because they change information flow, are trusted and can be individualized, transformed and in turn transform the

---

<sup>11</sup> I would like to add another interesting avenue of research: addictions. Since addiction is, at least partly, cognitive, it could be extended as well. This might explain why alcoholics frequently relapse after rehab when they return to the environment where they were addicted in the first place. The place where all the cognitive artefacts are that constituted (some) cognitive processes from their addiction.

individual<sup>12</sup>. If extended cognition is taken seriously in psychiatry, it could prompt revisions in the DSM, thereby changing how psychiatrists diagnose mental illness (Hoffman 2016, 1168). Furthermore, it could encourage therapists to take external resources more seriously, as possible harbours of negative mental states. This could reduce suffering in many and change their self- and social image.

The second dimension is trust. Trust in cognitive artefacts, people or practices is usually given because we think, among other things, that they are truthful, helpful or beneficial<sup>13</sup>. People do not want to connect with cognitive resources that are not to be trusted (that is, attenuating) and if they find out they are, they will disconnect as soon as possible. One might even say attenuated extended processes might be a theoretical possibility, but not an empirical reality. The reason why some might get this idea is that we usually focus on the intentional aspect of extended cognition. Otto uses his notebook to improve his memory because he wants to go to the museum, a group gets on a ship and steers it to a place they want to go to, two people collaborate to write a book together, a father wants to regulate his baby's emotions, etc. Few of us would intentionally want attenuated (extended) cognition. If we knew about it, we would probably change these processes to make them better.

Nevertheless, not wanting attenuated extended cognitive processes does not mean that they do not exist. It does not mean that cognitive artefacts cannot be unreliable, unsafe and malevolent, but still trusted and glued to the agent. Most of our cognitive processes and abilities are not about what we explicitly want at all, most are tacit.

Nobody actively sees tables and chairs instead of random undifferentiated shapes.

We cannot self-deceive on purpose every step of the way. An anorexic person cannot

---

<sup>12</sup> Hoffman warns that this is not a one size fits all explanation of extended psychiatric disorders. Not all individuals will respond in the same way to cognitive artefacts: "Indeed, one could imagine the opposite effect: the absence of the mirror, for example, might fuel or ground one's preoccupation with appearance. But it remains plausible that external artefacts may at least sometimes count as bona fide realizations of affective states, and that these extended affective states are relevant to psychopathology" (p.1167).

<sup>13</sup> We can conclude something is trustworthy through implicitly endorsing it or explicitly evaluating it.

help but keep looking at a mirror, activating her belief that she is too fat. Nobody wants to have anger issues in prison or believe sexist statements about their gender on purpose and limit their autonomy.

Slaby (2016) thinks the resistance to accepting non-intentional extension comes from the natural inclination to understand the mind in terms of operative processes alone; one individual with her inbuilt mental capacities, using an external resource to augment it. This is what he describes as the user/resource model, which leaves little room to explain the normative part of cognitive processes and non-intentional extension (or mind invasion as he calls it). By differentiating between the cognitive process and its normative status, the natural inclination to use the user/resource model is easier to circumvent.

So, while it is true that we usually only connect to cognitive artefacts because we think that will be beneficial – thus limiting the potential extension – that does not mean that in practice it always turns out to be the case. That is why this dimension is key to understanding attenuated extended cognition because an attenuated extended cognitive process is usually not something that we should trust. If the agent finds out his extended cognitive process is attenuated in some way, the logical step is (trying to) remove or change the attenuating part. This is not always what happens though: a GPS device is usually trustworthy and helps us with navigation, but it also attenuates people's internal navigational skills (Javadi et al. 2017). Will people change their navigation style if they find this out? I know I have not. In other cases, people trust artefacts, people or practices undeservedly: David is deceiving others to deceive himself, the information he gives and everyone who believes it is not trustworthy. He just picks his best picture, maybe photoshops it a bit and get compliments that might be undeserved if his followers knew what he looked like offline. However, David does put trust in his online followers as if they are trustworthy authorities to deceive himself that he is, in fact, handsome.

Another option is that trust might not even be a relevant dimension to the extended cognitive process, not even implicitly given trust. Consider the prison example, where Zelda cannot control her anger as she usually does with her body and environment, because she cannot walk away when those feelings bubble up. It does not matter whether she trusts the environment she is in; She is stuck there.

The third dimension is individualization. Which is about adjusting the informational properties of a cognitive artefact to make it more effective and efficient for realizing its cognitive function or, I would like to add, less effective and efficient. The fourth dimension of transformation goes one step further; this happens when cognitive artefacts change the agent. For instance, Otto individualized his notebook extensively (to the point that only he can use it) and in turn, the notebook transforms his behavioural and cognitive routines. When a cognitive artefact has such a prominent role in one's habits, usually both the agent and the artefact transform. While Heersmink focusses his attention on artefacts changing to fit our needs and agents changing for the better, I would like to supplement this with the idea that these artefacts can similarly change themselves and agents for the worse.

To return to Dings' example, David probably individualized his social media quite a lot: He can change his personal page. He can add and delete 'friends'. He can mute or block people and change the frequency or type of notifications. He can actively game the site's algorithms to show him more or less content from particular users. He might have installed many different add-ons and extensions to enhance his experience. In turn, David's social media use transforms him. With time, he probably learns how to make better pictures and find his perfect angle. He will probably pick up photoshop skills or maybe he will learn to write engaging captions for his pictures. In the long-term, social media might become a tool for mood-regulation, which of course changes his cognitive routines as well (Burrow & Rainone 2016). He could develop the habit to measure his attractiveness and self-image in terms of, or

based on, likes (Coulthard & Ogden 2018). That is, he comes to think of himself in terms of the system of social media.

The dimension of transformation is especially relevant when considering diachronic cognitive integration and extension on a large scale. Among others, Gallagher (2013) and Heersmink (2015) have examples of transformations that are beneficial to our cognition from mathematics and language (Heersmink 2015, 591) to cognitive practices like the ones ingrained into agents from the scientific community or the legal system (Gallagher 2013). These skills and practices can change what and how we cognize, and they augment our cognitive abilities. For example, mathematics improves our thinking about large numbers and makes it easier to do multiplications or divisions. Likewise, tools, from rulers and GPS devices to smartphones, can train certain embodied interactive skills and thus make certain cognitive processes more efficient (Heersmink 2015, 592). However, not every transformation is beneficial. Slaby (2016) is one of the few philosophers who explicitly centres problematic instances of extended cognition, or, in his case, situated affectivity (or 'mind invasion'). He explains how agents can be drawn into certain affective styles and interaction patterns which can be detrimental to the agent themselves. This can happen, for instance, in the workplace. Among other effects, workspace culture can shape agents in such a way that everyone is required to be 'always-on' and needs to be permanently industrious (p.10), even outside working hours. Other places can cause harmful cognitive patterns too. Cults and their dynamics might be explained through such harmful transformations. And certain social rules and categories connected to, for example, racism and sexism, can transform cognition in a bad way (Menary 2013) as well as undermine autonomy (Cash 2010; 2013). These cognitive structures are already given and are hard to escape since a person is always integrated with her social and material environment and is shaped by it through the process of enculturation.

In short, the eight dimensions of cognitive integration work, with some tweaks, for both augmenting, attenuating and of course neutral cognitive artefacts. Specifically, the dimensions can be applied to cases of self-deception, anger management, some psychiatric disorders like anorexia and phenomena like harmful workplace cultures and enculturation.

#### **4. Summary and conclusion**

The idea that cognitive processes occur through elements outside the head is not new. That our cognition is not perfect is also old news. However, there has been little attention for extended and integrated cognition which does not augment. The complementarity principle does not even allow non-augmenting extended cognition. This is a problem because many instances of extended cognition are not necessarily augmentations. I have addressed this issue by showing that extended cognitive processes can be attenuating and why. Then I adjusted the complementarity principle into the transformation principle, which is more neutral with regards to the normative status of the cognitive process. I showed that there is no cognitive bloat or coupling-constitution problem by analyzing social self-deception with the existing eight dimensions of integration of cognitive artefacts by Heersmink. These dimensions still work with the transformation principle, with some alterations. By broadening the field of extended and integrated cognition, knowledge of the domain can also be applied to a wider variety of (social) phenomena. With this paper, I hope to have made a small contribution towards explaining the extended and integrated nature of attenuated cognitive processes and I hope that, in the future, this framework can contribute to a greater understanding of other, less than desirable cognitive processes that are pervasive in society.

## Bibliography

- Adams, F. and K. Aizawa. (2001). The Bounds of Cognition. *Philosophical Psychology* 14(1), 43–64.
- Baron, M., (1988). What is Wrong with Self-Deception. In B. McLaughlin & A. O. Rorty (eds.), *Perspectives on Self-Deception* (431-449). Berkeley: University of California Press.
- Brownstein, M. (2019). Implicit Bias. *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.). Retrieved 6 May, 2020, from <https://plato.stanford.edu/archives/fall2019/entries/implicit-bias/>
- Burrow, A. L. & Rainone, N. (2016). How many likes did I get? Purpose moderates links between positive social media feedback and self-esteem. *Journal of Experimental Social Psychology* 69, 232-236. <https://doi.org/10.1016/j.jesp.2016.09.005>
- Butler, J., (1726). Upon Self-Deceit. in D.E. White (ed.), 2006. *The Works of Bishop Butler*. Rochester: Rochester University Press.
- Cash, M. (2010). Extended cognition, personal responsibility, and relational autonomy. *Phenomenol. Cogn. Sci.* 9, 645–671. doi: 10.1007/s11097-010- 9177-8
- , (2013). Cognition Without Borders: ‘Third Wave’ Socially Distributed Cognition and Relational Autonomy. *Cognitive Systems Research* 25–26, 61–71.
- Caspi, A. & Gorsky, P. (2006). Online deception: Prevalence, motivation, and emotion. *CyberPsychology & Behavior*, 9(1), 54-59.  
<http://dx.doi.org/10.1089/cpb.2006.9.54>.
- Clark, A. and Chalmers, D. (1998). The extended mind. *Analysis* 58, 7–19. doi: 10.1093/analys/58.1.7
- Clark, A. (2010). Memento’s revenge: The extended mind, extended. In Menary, R. (Ed.), *The extended mind*. Cambridge, Massachusetts: MIT press.
- Cooper, R. (2017). Where’s the problem? Considering Laing and Esterson’s account of schizophrenia, social models of disability, and extended mental disorder. *Theoretical Medicine and Bioethics* 38, 295–305. <https://doi.org/10.1007/s11017-017-9413-0>
- [Coulthard, N.](#) & [Ogden, J.](#) (2018). The impact of posting selfies and gaining feedback (‘likes’) on the psychological wellbeing of 16-25 year olds: An experimental study.

*Cyberpsychology: Journal of Psychosocial Research on Cyberspace* 12-2) 4.  
<https://doi.org/10.5817/CP201824>

Darwall, S. (1988). Self-Deception, Autonomy, and Moral Constitution. In B. McLaughlin and A. O. Rorty (eds.), *Perspectives on Self-Deception*. Berkeley: University of California Press.

Dehaene, S. (1997). *The number sense*. New York: Oxford University Press.  
—, (2009). *Reading in the brain*. New York: Penguin Viking.

Deweese-Boyd, I. (2016). Self-Deception. *The Stanford Encyclopedia of Philosophy*. Retrieved 6 May, 2020, from <https://plato.stanford.edu/archives/fall2017/entries/self-deception>

Dings, R. (2017). Social strategies in self-deception. *New Ideas in Psychology* 47, 16-23.  
<http://dx.doi.org/10.1016/j.newideapsych.2017.03.0110732-118X>

Donald, M. (1991). *Origins of the Modern Mind*. Cambridge, MA: Harvard University Press.

Eichhorn, K. (2019). *The End of Forgetting: Growing Up With Social Media*. Cambridge, MA: Harvard University Press. <https://dx.doi.org/10.4159/9780674239333>

Fernández, J. (2013). Self-deception and self-knowledge. *Philosophical Studies* 162-2, 379–400.

Gallagher, S. (2013). The socially extended mind. *Cognitive Systems. Research* 25–26, 4–12. doi: 10.1016/j.cogsys.2013.03.008

Heersmink, R. (2015). Dimensions of integration in embedded and extended cognitive systems. *Phenomenology and the Cognitive Sciences* 14(3), 577–598.  
—, (2015). Extended mind and cognitive enhancement: Moral aspects of cognitive artifacts. *Phenomenology and the Cognitive Sciences* 16 (1), 17-32. doi:10.1007/s11097-015-9448-5  
—, (2017). Distributed Cognition and Distributed Morality. Agency, Artifacts and Systems. *Science and Engineering Ethics* 23, 431–448. DOI 10.1007/s11948-016-9802-1  
—, (2020). The extended mind and artifactual autobiographical memory. In *The Mark of the Mental (MOM) Project. IUSS Philosophy Seminars*. Lecture conducted from Università Vita-Salute San Raffaele, Pavia, Italy.

Hirstein, W. (ed.). (2009). *Confabulation: Views from Neuroscience, Psychiatry, Psychology, and Philosophy*. Oxford: Oxford University Press.  
<https://dx.doi:10.1093/acprof:oso/9780199208913.001.0001>

- Hoffman, G.A. (2016). Out of our skulls: How the extended mind thesis can extend psychiatry. *Philosophical Psychology* 29 (8), 1160–1174.  
<http://dx.doi.org/10.1080/09515089.2016.1236369>
- Honneth, A. (2012). *The I in the We: Studies in the Theory of Recognition*. Cambridge: Polity.
- Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.
- Javadi, A., Emo, B., Howard, L. *et al.* (2017). Hippocampal and prefrontal processing of network topology to simulate the future. *Nature Communications* 8, 14652.  
<https://doi.org/10.1038/ncomms14652>
- Jenni, K. (2003). Vices of Inattention. *Journal of Applied Philosophy* 20(3), 279–95.
- Kaplan, D. M. (2012). How to Demarcate the Boundaries of Cognition. *Biology and Philosophy* 27, 545–70.
- Krueger, J. (2013). Ontogenesis of the Socially Extended Mind. *Cognitive Systems Research* 25, 40–46.
- McKay, R. and Dennett, D. (2009). The Evolution of Misbelief. *Behavioral and Brain Sciences*, 32(6): 493–561.
- Menary, R. (2007). *Cognitive Integration: Mind and Cognition Unbound*. London: Palgrave Macmillan. doi: 10.1057/9780230592889
- , (2010a). Dimensions of mind. *Phenomenology and the Cognitive Sciences* 9(4), 561–578.
- , (2010b). Cognitive integration and the extended mind. In R. Menary (Ed.), *The extended mind* (pp. 227–243). Cambridge, Mass.: MIT Press.
- , (2012). Cognitive practices and cognitive character. *Philosophical Explorations* 15(2), 147–164.
- , (2013) Cognitive integration, enculturated cognition and the socially extended mind. *Cognitive Systems Research* 25(26), 26–34.  
<http://dx.doi.org/10.1016/j.cogsys.2013.05.002>
- Menary, R. & Kirchhoff, M. (2013). Cognitive transformations and extended expertise. *Educational Philosophy and Theory* 46, 610–623.  
<http://dx.doi.org/10.1080/00131857.2013.779209>
- Michaelian, K. & Sutton, J. (2013). Distributed cognition and memory research: History and current directions. *Review of Philosophy and Psychology* 4(1), 1–24.

- Nelkin, D. (2012). Responsibility and Self-Deception: A Framework. *Humana. Mente Journal of Philosophical Studies* 20, 117–139.
- Norman, D. (1991). Cognitive artifacts. In J. M. Carroll (Ed.), *Designing interaction: Psychology at the human-computer interface* (pp. 17–38). Cambridge: Cambridge University Press.
- Paulus, P.B. (1998). Developing Consensus about Groupthink after All These Years. *Organizational Behavior and Human Decision Processes* 73 (2–3), 362–374.  
<http://dx.doi.org/10.1006/obhd.1998.2767>.
- Plato (R. Hackforth, Trans.). (1952). *Phaedrus*. Cambridge: Cambridge University Press. (Original work published c. 370 BCE).
- Rowlands, M. (1999). *The body in mind: Understanding cognitive processes*. Cambridge: Cambridge University Press.
- Rupert, R. (2004). Challenges to the thesis of extended cognition. *Journal of Philosophy* 101, 389–428.
- Sartre, J-P. (1956). *Being and Nothingness*. Paris: Gallimard; (H. E. Barnes, Trans.). New York: Washington Square Press. (Original work published 1943).
- Sterelny, K. (2010). Minds: Extended or scaffolded? *Phenomenology and the Cognitive Sciences* 9(4), 465–481. <http://dx.doi.org/10.1007/s11097-010-9174-y>
- Slaby, J. (2016). Mind Invasion: Situated Affectivity and the Corporate Lifehack. *Frontiers in Psychology* 7 (266). <http://dx.doi.org/10.3389/fpsyg.2016.00266>
- Sutton, J. (2006). Distributed cognition: Domains and dimensions. *Pragmatics & Cognition* 14(2), 235–247.
- , (2010). Exograms and interdisciplinarity: history, the extended mind, and the civilizing process. In R. Menary (ed.), *The Extended Mind* (pp. 189–225). Cambridge, MA: MIT Press. <http://dx.doi.org/10.7551/mitpress/9780262014038.003.0009>
- Taylor, S. and Brown, J., (1994). Positive Illusion and Well-Being Revisited: Separating Fact from Fiction. *Psychological Bulletin*, 116: 21–27.
- Varga, S. (2016). Interaction and extended cognition. *Synthese* 193, 2469–2496.  
<http://dx.doi.org/10.1007/s11229-015-0861>
- Wilson, R. & Clark, A. (2009). How to situate cognition: Letting nature take its course. In P. Robbins & M. Aydede (Eds.), *The Cambridge handbook of situated cognition* (pp. 55–77). New York: Cambridge University Press.

Wilson, R.A. (2013). Ten questions concerning extended cognition, *Philosophical Psychology* 27 (1), 19-33. DOI:10.1080/09515089.2013.828568



# PhD Project Proposal

## Scaffolding autonomy with social media

### Summary

Autonomy is something to strive towards. But research in the past decades has shown that we are not as in control as we would like to think. We are dependent on social and material factors outside of our control. Relational autonomy is the idea that we cannot and do not always need to control those things to be autonomous agents. This project will develop the idea of relational autonomy further by investigating how the literature on extended cognition can improve theories of relational autonomy. Extended cognition is a tradition in philosophy according to which the brain is merely an arbitrary stopping point for the investigation of cognition. The body, people, practices and external structures can become part of the cognitive process.

I will make a taxonomy that combines relational autonomy and extended cognition. By considering that a part of autonomy could be extended cognition, intuitions and research about why some influences on autonomy are a threat to it, can be explained better and more systematically. I will demonstrate this in a case study on social media because it is an influence that, according to research, manipulates our cognition often and effectively and thus influences our autonomy. Furthermore, by knowing the threats of social media to autonomy, I can determine which skills are most important to enhance autonomy.

The first stage of this project aims to develop a taxonomy of the kinds of relational autonomy and how they relate to types of extended cognition. Stage two will apply the taxonomy. It investigates the extent to which we can explain existing research about the autonomy undermining effects of social media, by invoking ideas from

extended cognition. Based on these findings, three will identify which skills can counteract the negative effects of social media on autonomy and cognition.

## **Keywords**

Relational autonomy, extended cognition, social media, cognitive competencies, scaffolded autonomy.

## **The problem**

The first aim of this project is to develop a taxonomy of the diverse kinds of autonomy that includes the lessons learnt in extended cognition. The second aim is applying those insights to better explain research about how we become more *and* less autonomous because of the cognitive consequences of social media.

Most philosophers, psychologists and other scientists agree that we are not self-sufficient, hyper-rational and unaffected by others or emotions. That is why proponents of relational autonomy argue that people's identities, needs, interests and choices are shaped in relation to others and our social, historical and physical surroundings.

The problem is that there has been little attention to how extended cognition fits into theories of relational autonomy. Extended cognition and relational autonomy have a common denominator: the situatedness of agents. Extended cognition states that cognitive processes are not just carried out in the brain, but our body and environment can be part of the cognitive process too.

I hypothesise that extended cognition and relational autonomy can benefit, enrich and reinforce each other. Investigating connections between them on a wide scale is new and worthwhile because it can give insights into the research and intuitions about non-neural factors in cognition that affect autonomy. In a case study on social media, I will apply and test my taxonomy. By invoking ideas from extended

cognition, I will be able to better explain research and intuitions about how we become more *and* less autonomous because of social media.

People feel like they are missing a limb if they forget their phone and thus their access to social media. Even though this limb is partly controlled by people, algorithms and organizations that can manipulate us with precision outside of our conscious awareness. One only needs to look at the Cambridge-Analytica scandal to know the far-reaching consequences. Where vast amounts of data from Facebook users were harvested without consent and employed to exploit their decision-making vulnerabilities with political advertising. This incident is not an aberration and while propaganda (and advertising) is not new, the scale and intimacy are (Susser et al. 2019).

Currently, the idea that autonomy undermining and enhancing factors can be part of our cognition, is barely examined in the research on relational autonomy. My project will give clarity about the connection between relational autonomy and extended cognition. One of the places where insight into this connection is useful is social media because it better explains research on the cognitive effects of social media on autonomy.

## **Philosophical and scientific background**

### **Relational autonomy**

My description of autonomy will be broad-brush, risking caricature for the sake of clarity of line. In the past thinking of autonomy has been dominated by the idea that we are independent and rational reasoners, unaffected by social relations, stripped from 'distorting' influences like emotions. Nowadays, research has shown that we are not the independent and rational individuals we think we are. The idea that autonomy does not necessarily require hyper rationality or self-sufficiency, that we are shaped by our body, our social environment and our history is called relational autonomy (Stoljar 2018). This feminist critique of old conceptions of autonomy

argues that we can be autonomous by relating to those influences in the right way. Relational autonomy concentrates on subjects like freedom and capabilities, threats to autonomy (e.g. coercion, paternalism, deception, manipulation, force) and the limits of freedom and autonomy (e.g. in contexts of punishment, education, health).

Minimally, with relational autonomy, one is autonomous if they are a 'normal adult' without cognitive impairment (Narayan 2002), though most philosophers go beyond that. There are numerous ways in which to cash out the consequences of a relational view on autonomy. While some might see autonomy as a unitary concept and think that the various conceptions exclude each other, some thinkers see autonomy as a multi-dimensional concept and they have integrated various theories on autonomy into one framework. One of those thinkers is Mackenzie (2014). Because she has a good overview of the ways one can be autonomous, I will use her work as a foundation to compare the diverse types of extended cognition to and to make my taxonomy. She distinguishes three types of autonomy: self-governance, self-authorization and self-determination. Self-determination involves having the freedom and opportunities to make and enact choices. This includes what to value, who to be and what to do. It is about external, structural conditions for individual freedom (e.g. Dworkin 1998; Raz 1986; Oshana 2006).

Self-governance is about the skills and capacities to make and enact choices that fit with our identity. This axis identifies internal conditions for autonomy. For example, willpower, but also reason-responsiveness and authenticity (e.g. Beauchamp & Childress 2012; Wolf 1990; Benson 1991; Meyers;1989 Frankfurt 1988; Christman 2009).

Lastly, self-authorization relates to concepts like accountability, self-trust and social recognition, which means regarding oneself as having the authority to be self-determining and self-governing (e.g. Benson 2005; Westlund 2009; Mcleod 2002; Anderson 2014).

Furthermore, there are three levels on which these types of autonomy occur: local, programmatic and global. Respectively, from deciding in concrete situations, to managing major life issues (e.g. concerning finances, work or family) and lastly, being in control of your whole life. A person might be very autonomous in one domain or level while being subjected to domination in another.

While relational autonomy does acknowledge that autonomy can be diminished or enhanced by external factors (like social media) and internalized oppressive norms, I think it can become a better theory by taking the lessons learned in extended cognition into account, where external factors can become part of our cognition.

### **Extended Cognition**

Research on cognition has undergone the same kind of transformation as research on autonomy. We used to think cognition was a rational and atomistic process, even though it is not always rational (Fine 2008) or something that only happens inside a head. The idea that cognition can be extended beyond the brain is gaining ground. This tradition states that cognitive processes are not just carried out by our brain, but our body and environment can be part of it too. Cognition does not begin in a head and extends from there, it also “works from the outside in; tools, technologies and institutions often shape our cognitive processes... and may even elicit plastic changes in neuronal structure” (Gallagher 2013, 7). This idea started with Clark and Chalmers (1998) and the parity principle which states that if something outside of the head is functionally the same for an agent as a brain process, it is cognitive. The second wave in the field uses the complementarity principle, which focusses on the integration of cognitive artefacts into neural structures, rather than parity (Heersmink 2015). The third wave, enactive cognition, adds to this that cognition arises due to dynamic interactions between acting organisms and their environment (Gallagher 2013). There are many variations on extended cognition. For example, cognitive integration (Menary 2013) considers cognitive practices and the history of agents when trying to explain how cognition can be extended.

Yet, there are some lacunae in the research on extended cognition. The first is that most of the research is about individuals, using tools deliberately and rationally, to improve their cognitive process. However, cognitive processes need not be deliberate or rational, one agent need not be the locus of control and extension is not necessarily enhancement (Slaby 2016, Hunting; unpublished). The second lacuna is that little research has been done into the connection between extended cognition and relational autonomy, apart from Cash (2010;2013), Anderson and Heath (2010) and Vierkant (2014). Cash has researched how socially distributed cognition, responsibility and relational autonomy relate to each other and Anderson has investigated scaffolded autonomy in various papers, specifically extended willpower and procrastination (just like Vierkant). Their work is a starting point for my wider aim of a taxonomy of how the types of extended cognition relate to the dimensions of autonomy as distinguished by Mackenzie.

Which types of extended cognition map unto which dimensions of autonomy? I want to test my answer to that question by looking at a case study of social media. There is some research on extended cognition and the web (Smart 2017; Heersmink & Sutton 2018), but not on social media specifically. When does social media extend cognition? To what extent does extended cognition explain existing intuitions and research about the autonomy undermining and enhancing effects of social media?

## **The Research Project**

### **Year 1: Autonomy and cognition**

This stage aims to develop an integrated account of relational autonomy and extended cognition. While the debates surrounding them have been mostly in isolation from each other, the two concepts share a common philosophical background. Namely the idea that agents are not isolated from their surroundings. Where and how do these concepts overlap and connect? There is some research on it by Cash (2010; 2013), who investigates relational autonomy, responsibility and

socially distributed cognition, and Anderson, who has, among other things, argued for scaffolded autonomy and extended willpower. Their accounts are focused on parts of relational autonomy and extended cognition. I aim to expand upon their work and make a systemic overview of interconnections between relational autonomy and extended cognition. By bringing extended cognition into the research on relational autonomy, the way we see 'external' factors that influence autonomy will change and thus the relational part of autonomy will be emphasized.

While autonomy can be about having enough options to choose from, this research targets the cognitive parts of relational autonomy due to the focus on extended cognition. For example, the ability to recognize those options and the willpower and self-trust you need to make a choice.

Because there are many models of extended cognition, this phase investigates how the types of explanations fit with the dimensions of relational autonomy as surveyed by Mackenzie (2014). This will be done with the help of the taxonomy of extended cognition by Slors (2019) and the degree of integration of cognitive artefacts (Heersmink 2015; Hunting: unpublished).

The result of this phase will be a matrix of where and how extended cognition and relational autonomy overlap. For example, the parity principle is probably applicable to local self-governing types of autonomy, because it focusses on specific instances of cognition in individuals. While Menary's theory of cognitive integration centralizes the effect of institutions on cognition, just like self-determination does.

## **Year 2: Autonomy and social media**

The second stage will be a test case on social media. I hypothesise that my taxonomy will be able to better explain existing intuitions and research about the autonomy enhancing and undermining effects of social media because I take extended cognition into account. If I run into problems here, the taxonomy will be adjusted accordingly.

The idea that systems and tools change us, is not new, but the threat social media brings to autonomy might be more pressing. If it is cognitively linked in the same way as other parts of the internet (Smart 2017; Heersmink & Sutton 2018), then social media has become part of our cognition (Dings 2017). What does this intimate connection between us and social media mean for theories on cognition and autonomy?

There is much research on how social media manipulates us (e.g. Klenk 2020; Susser et al. 2019; Yeung 2016; Nagel et al. 2016; Zarsky 2003), from fake news to online echo chambers, digital PR campaigns that manipulates emotions and misleading political rhetoric that influences public opinion. This manipulation is known as 'techno-social engineering' (Frischmann & Selinger 2018) and it aims to change how we think, perceive and act.

Which types of extended cognition and relational autonomy are at play in social media? To test my taxonomy, I will categorize the threats of social media to autonomy. Next, I find out which are cognitive and which could be extended cognition. For example, social self-deception on social media (Dings 2017).

### **Year 3: Autonomy skills**

This stage will apply the insights gained into autonomy, cognition and social media by asking: how can ordinary people counteract the negative effects of social media on autonomy? My tentative hypothesis is that, because of the understanding gained in the connections between extended cognition and relational autonomy, I will gain more insight into which skills are important when one wants to enhance ones autonomy.

First, the skills which can enhance the numerous ways people are autonomous will be identified. Meyers's work on autonomy competencies (1989; 2005) and Nussbaum's (2000) work on capability approaches to autonomy will be applied to structure and interrelate competencies. From reasoning skills like imagination

(Carlson & White, 2013; Smith 2010) and creativity (Hutchins 2010) to perceptual skills like aspect-seeing (Baz, 2000; Kalis 2017) and self-understanding to managerial skill like willpower.

Just like some threats to autonomy are extended cognition, some autonomy enhancing skills could be too. There is some research on extended cognitive skills. For instance, aspect-seeing: the idea, well-known in the extended cognition literature, that perception and action are intertwined (Merleau-Ponty 1996; Noë 2004). This suggests that we might exercise agency by changing the way we see the world (McGeer 2008, 2015; Pettit 2016). Willpower can be extended too. Anderson and Heath (2010) argue that what looks like ‘sheer willpower’ is not a personality trait, but “a feature of persons as they are in particular contexts, with the requisite scaffolding” (p. 249). Our environment can give us the same willpower as ‘simply bucking down’.

#### **Year 4: Thesis**

Here, it is time to summarize and draw up the implications. How do extended cognition and relational autonomy interrelate? What does that mean in a digital society like ours? What can individuals do to enhance their autonomy?

#### **4. Philosophical, scientific and societal impact**

This project doubles down on the idea that we are creatures that are embedded in our social, physical and digital world, which can be both a blessing and a curse. People’s autonomy can be threatened on a massive scale, as has happened with the Cambridge-Analytica scandal. Moreover, the way it happens is more intimate than ever: the manipulation becomes part of our cognition. Research into this is necessary to find out what it means to say social media becomes part of our cognition. And what it means to say social media threatens autonomy.

It enriches the field of extended cognition with the notion that extension is not necessarily enhancement. This allows for theorizing about all types of non-enhancing extended cognition and nuances how we think about cognitive extension.

It also enriches the field of relational autonomy by stressing the importance of factors that, up until now, were considered merely causal influences, while in fact, they are part of our cognition. That means new vulnerabilities and benefits of resources external to the brain can be discovered. The focus on skills to enhance autonomy might help individuals overcome the threats and use the benefits of social media.

This research might also have an impact on law-making regarding social media. What should be expected from the media and lawmakers regarding autonomy and social media? What skills should be expected from or taught to citizens?

## Work programme

<i>Year/Stage</i>	<i>Months*</i>	<i>Programme</i>
1	12	<ul style="list-style-type: none"> <li>- Gather and read the latest literature on versions and taxonomies of extended cognition (Like Slors and Heersmink).</li> <li>- Gather and read the latest literature on versions and taxonomies of relational autonomy (like Mackenzie).</li> <li>- Execute stage one of the research.</li> <li>- Write an article on the connections between relational autonomy and cognition.</li> </ul>
2	12	<ul style="list-style-type: none"> <li>- Gather and read philosophical and empirical work on the threats of social media on autonomy (from the effects of algorithms to hyper nudges).</li> <li>- Select which of those threats are cognitive in nature.</li> <li>- Find out if and how social media is part of our cognition (like social self-deception).</li> <li>- Find and visit the main expert(s) on these topics.</li> <li>- Write and publish an article on autonomy, cognition and social media understood in the developed framework of stage one.</li> <li>- Execute stage two of the research.</li> </ul>
3	12	<ul style="list-style-type: none"> <li>- Research the various cognitive skills individuals can learn to enhance autonomy.</li> <li>- Select the cognitive skills useful to enhance autonomy considering social media.</li> <li>- Investigate which skills could be extended cognition.</li> <li>- Write and publish an article on the (extended) skills for autonomy when specifically regarding social media.</li> <li>- Execute stage three of the research.</li> </ul>
4	12	<ul style="list-style-type: none"> <li>- Synthesize findings and write PhD-thesis.</li> <li>- Present conclusions at international conferences.</li> </ul>

*\*Months include vacations and holidays. The standard amount of vacation days and teaching obligations have been taken into account in the setup of this work program.*

## **Summary for non-specialists (in Dutch or English)**

Iedereen wil autonoom en onafhankelijk zijn. Een deel van autonomie is cognitief, zoals besluiten wat je wil, ontdekken wat goed voor je is en plannen hoe je je doelen kunt bereiken. Maar onderzoek laat zien dat we behoorlijk afhankelijk en irrationeel zijn. Die afhankelijkheid is onontkoombaar door lichamelijke, sociale en materiële factoren. Maar het goede nieuws is dat, hoe onwaarschijnlijk ook, deze afhankelijkheid ons autonoom maakt. Dankzij onze afhankelijkheid van anderen, bijvoorbeeld, leren we niet in ons eentje hoe de wereld in elkaar zit en kunnen we advies inwinnen van vrienden als we zelf niet weten wat we moeten doen of hoe we ons doel moeten bereiken. Zonder dit soort handelingen is autonoom zijn lastig. Sociale media kunnen dit versterken. Zo kunnen we communiceren via Whatsapp, leren we via YouTube, vragen we hulp via Reddit en bepalen we op wie we stemmen via Facebook.

Het is alsof we een ledemaat verliezen als we onze telefoon met alle sociale media apps zijn vergeten. Sociale media beïnvloeden ons en ons denken, want denken gebeurt niet in een vacuüm. We worden beïnvloed door onze omgeving, andere mensen en de technologie die we gebruiken. Sterker nog, in de cognitiefilosofie wordt het idee dat we niet-neurale structuren gebruiken om onze cognitie te faciliteren, steeds populairder. Cognitie kan bijvoorbeeld lichamenlijk zijn, zoals kinderen die alleen kunnen tellen met hun vingers. Na verloop van tijd zijn vingers niet meer nodig om te tellen. Volwassenen gebruiken nog steeds allerlei middelen buiten het brein om na te denken, van notities op de telefoon om dingen te onthouden tot GPS-systemen die ons helpen navigeren. Deze 'extended cognition' heeft verregaande effecten op ons denken. Ik denk dat inzichten van 'extended cognition' kunnen uitleggen waarom sociale media zoveel effect heeft op ons. Maar sociale media hebben niet alleen positieve gevolgen voor onze cognitie en autonomie. In 2018 werden we met de neus op dit feit gedrukt met het Cambridge-Analytica schandaal. Het bedrijf zette data van miljoenen Facebookgebruikers in om

kwetsbaarheden in hun besluitvorming uit te buiten en zo hun stemgedrag tijdens de Amerikaanse verkiezingen te beïnvloeden. Dit soort incidenten gebeuren op grote schaal en zijn zeer effectief, van advertenties tot echokamers en van propaganda tot gefilterde nieuwsfeeds. Mijn hypothese is dat we, door de ideeën van 'extended cognition' te gebruiken, beter kunnen uitleggen waarom we sommige vormen van invloed, zoals die van sociale media, als autonomie-bedreigend zien en andere niet

Om dit te onderzoeken, ga ik een brug slaan tussen autonomie en 'extended cognition'. Beide concepten benadrukken dat onze omgeving ons beïnvloed en dat wij onze omgeving nodig hebben om te denken en handelen, maar er is nog weinig onderzoek naar hoe de twee concepten bij elkaar passen. Door verscheidene theorieën van 'extended cognition' en autonomie te vergelijken en in te delen in een taxonomie, kan ik intuïties en onderzoek over waarom sociale media onze autonomie manipuleert, beter verklaren. Daaruit zal ook blijken welke soort vaardigheden het belangrijkste zijn om sociale media in ons voordeel te laten werken en autonomie te vergroten.

## Bibliography

- Anderson, J. (2014). Vulnerability and Autonomy Intertwined in *Vulnerability: New Essays in Ethics and Feminist Philosophy*, edited by Catriona Mackenzie, Wendy A. Rogers, and Susan Dodds. New York: Oxford University Press.
- Baz, A. (2000). What's the point of seeing aspects? *Philosophical Investigations*, 23-2, 97-121.
- Beauchamp, T and Childress, J. (2012). *Principles of Biomedical Ethics*, 7th ed. New York: Oxford University Press.
- Benson, P. (1991). Autonomy and Oppressive Socialization. *Social Theory and Practice* 17: 385–408.
- , (2005). Feminist Intuitions and the Normative Substance of Autonomy in: J.S. Taylor (ed.), *Personal Autonomy: New Essays on Personal Autonomy and Its Role in Contemporary Moral Philosophy*, (p. 124–142). Cambridge: Cambridge University Press
- , (2005). Taking Ownership: Authority and Voice in Autonomous Agency. In *Autonomy and the Challenges to Liberalism*, edited by J. Christman and J. Anderson. (101–126). Cambridge: Cambridge University Press.
- Carlson, S. M., & White, R. E. (2013). Executive function, pretend play, and imagination. In M. Taylor (Ed.), *Oxford library of psychology. The Oxford handbook of the development of imagination* (p. 161–174). Oxford University Press.
- Cash, M. (2013). Cognition Without Borders: 'Third Wave' Socially Distributed Cognition and Relational Autonomy. *Cognitive Systems Research* 25–26, 61–71.
- Cash, M. (2010). Extended cognition, personal responsibility, and relational autonomy. *Phenomenology and the Cognitive Sciences*, 9(4), 645–671. <http://dx.doi.org/10.1007/s11097-010-9177-8>
- Christman, J. (2009). *The Politics of Persons: Individual Autonomy and Sociohistorical Selves*. Cambridge: Cambridge University Press.
- , (2014) Freedom and the Extended Self. *Ethical Perspectives* 21, no. 2: 225–54.
- Clark, A. and Chalmers, D. (1998). The extended mind. *Analysis* 58, 7–19. doi: 10.1093/analys/58.1.7
- Dings, R. (2017). Social strategies in self-deception. *New Ideas in Psychology* 47, 16-23. doi: 10.1016/j.newideapsych.2017.03.0110732-118X
- Dworkin, G. (1988). *The Theory and Practice of Autonomy*. New York: Cambridge University Press.
- Fine, C. (2006). *A mind of its own*. New York: Norton & Co.
- Frankfurt, H. (1988). *The importance of what we care about*. Cambridge: Cambridge University Press.
- Frischmann, B., & Selinger, E. (2018). *Re-Engineering Humanity*. Cambridge: Cambridge University Press. doi:10.1017/9781316544846
- Gallagher, S. (2013). The socially extended mind. *Cognitive Systems Research* 25–26, 4–12. doi: 10.1016/j.cogsys.2013.03.008

- Heath, J & Anderson, J. (2010) Procrastination and the Extended Will. In Chrisoula Andreou and Mark White (eds.), *The Thief of Time: Philosophical Essays on Procrastination*. (233-252) New York: Oxford University Press.
- Heersmink, R. (2015). Dimensions of integration in embedded and extended cognitive systems. *Phenomenology and the Cognitive Sciences* 14-3, 577–598.
- Heersmink, R & Sutton, J. (2018). Cognition and the Web: Extended, Transactive, or Scaffolded? *Erkenntnis*. doi:10.1007/s10670-018-0022-8.
- Hutchins, E. (2010). *Enaction, imagination, and insight. Enaction: Towards a new paradigm for cognitive science*. 425–450.
- Kalis, A. (2017). Self-control as a normative capacity. *Ratio*. doi:10.1111/rati.12180
- Klenk, M. (2020). Digital well-being and manipulation online. In C. Bur, & L. Floridi (Eds.), *Ethics of digital well-being: A multidisciplinary approach*. Dordrecht: Springer.
- McLeod, C. (2002). *Self-Trust and Reproductive Autonomy*. Cambridge, MA: MIT Press.
- Mackenzie, C (2014). Three Dimensions of Autonomy: A Relational Analysis. In *Autonomy, Oppression, and Gender*. Edited by Piper and Veltman, New York: Oxford University Press.
- McGeer, V. (2008). The Moral Development of First-Person Authority. *European Journal of Philosophy*. 16-1, 81–108.
- , (2015). Mind-making practices: the social infrastructure of self-knowing agency and responsibility. *Philosophical Explorations*, 18-2, 259–281.
- Menary, R. (2013). Cognitive integration, enculturated cognition and the socially extended mind. *Cognitive Systems Research* 25-6, 26–34. doi:10.1016/j.cogsys.2013.05.002
- Merleau-Ponty, M. (1996). *Phenomenology of perception*. London: Routledge.
- Meyers, D. (1989) *Self, Society and Personal Choice*. New York: Columbia University Press.
- , 2005, Decentralizing Autonomy. Five Faces of Selfhood. In Anderson and Christman, *Autonomy and the challenges to liberalism: new essays*. (27–55). Cambridge: Cambridge University Press; 2005.
- Nagel, S.K., Hrinco, V., and Reiner, P.B. (2016). Algorithm Anxiety—Do Decision-Making Algorithms Pose a Threat to Autonomy? Presented at 2016 IEEE International Symposium on Ethics in Engineering, Science and Technology, Vancouver, Canada.
- Narayan, U. (2002). Minds Of Their Own: Choices, Autonomy, Cultural Practices and Other Women. In L. Antony and C. Witt (eds.), *A Mind of One's Own. Feminist Essays on Reason and Objectivity* (p. 418–432). Boulder, CO: Westview.
- Noë, A. (2004). *Action in perception*. Cambridge: MIT Press.
- Nussbaum, M. (2000). *Women and Human Development: The Capabilities Approach*. Cambridge: Cambridge University Press.
- Oshana, M. (2006). *Personal Autonomy in Society*. Aldershot, UK: Ashgate.
- Pettit, P. (2016). Making Up Your Mind: How Language Enables Self-Knowledge, Self-Knowability and Personhood. *European Journal of Philosophy*, 24-1, 3-26.

- Raz, J. (1986). *The Morality of Freedom*. Oxford: Clarendon Press.
- Slaby, J. (2016). Mind Invasion: Situated Affectivity and the Corporate Lifehack. *Frontiers in Psychology* 7-266. doi:10.3389/fpsyg.2016.00266
- Slors, M. (2019). Symbiotic cognition as an alternative for socially extended cognition, *Philosophical Psychology*, 32:8, 1179-1203, doi: 10.1080/09515089.2019.1679591
- Smart, P. (2017). Extended Cognition and the Internet. *Philosophy & Technology* 30, 357–390. doi:10.1007/s13347-016-0250-2
- Smith, M. N. (2010). Practical imagination and its limits. *Philosophers' Imprint*. 10-3, 1–20.
- Stoljar, N. (2018). Feminist Perspectives on Autonomy. *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.). <https://plato.stanford.edu/archives/win2018/entries/feminism-autonomy>
- Susser, D., Roessler, B. & Nissenbaum, H. (2019). Technology, autonomy, and manipulation. *Internet Policy Review*, 8-2. doi: 10.14763/2019.2.1410
- Zarsky, T.Z. (2003). "Mine your own business!": Making the case for the implications of the data mining of personal information in the forum of public opinion. *Yale Journal of Law and Technology* 5-1, Article 1.
- Vierkant, T. (2014). Mental muscles and the extended will. *Topoi*, 33-1, 57-65.
- Westlund, A. (2009). Rethinking Relational Autonomy. *Hypatia* 24-4. 26–49.
- Wolf, S. (1990). *Freedom within Reason*. New York: Oxford University Press.
- Yeung, K. (2016). 'Hypernudge': Big Data as a mode of regulation by design. *Information, Communication & Society* 20. 1-19. doi:10.1080/1369118X.2016.1186713.

## Curriculum Vitae PhD candidate

### a) Education

- **WO Research Master Philosophy at the Radboud University** 2018 – 2020

Specialization in Philosophy of Mind

Courses:

- Philosophy of Mind: Cognitive neuroscience, self-understanding and interpretation
- Philosophy of Mind: Mechanistic explanation and causation in cognitive neuroscience
- Philosophy of Mind: Socially extended cognition
- Analytic Philosophy: Implicit Bias
- Analytic Philosophy: Social perception
- Analytic Social Philosophy: Evolution & Social cognition
- Philosophy & Science: Heidegger & Stiegler on Science, Technology and the Human Condition
- Neurophilosophy
- Media theories of influence
- Persuasion in entertainment media
- Methods and Skills I & II

- **WO Pre-master Philosophy Radboud University** 2017

- **HBO Bachelor of Arts: ArtEZ Creative Writing** 2011 – 2016

Graduated with a collection of science poems 'Fractaal'

And a thesis on autonomy, 'brainwashing' and empathy:

'Hoe normale mensen terroristen worden'

- **VWO Olympus College** 2004 – 2010

### b) Relevant academic experience

#### Presentations

- Extended cognition, enhanced cognition? 2020

- Presentation for WIP, Radboud University

#### Conferences, workshops, extracurricular meetings

- Conference: The Ends of Autonomy 2020

The University of Warwick (Coventry, The United Kingdom)

- Conference: Free Will Methodology 2019  
KU Leuven (Leuven, Belgium)
- Conference: Situated Cognition 2019  
Ruhr-University Bochum (Bochum, Germany)
- Workshops: Explaining Social Cognition 2019  
Organized by K. Abramova  
Radboud University (Nijmegen, The Netherlands)
- Workshops: Affordances, Content Internalism and Externalism 2018  
Organized by M. Slors  
Radboud University (Nijmegen, The Netherlands)
- c) Other relevant positions**
- Hulpstudent 2019 –present
- Tutor VWO: Mathematics
- Mangaschool 2014 – 2016
- Teacher in Creative Writing
- Olympus College: 2008 – 2012
- Tutor VWO: Biology
- Tutor VWO: Mathematics
- Tutor VWO: English
- Tutor VWO: Philosophy
- Fasold Games: Interactive fairy tales 2014 – 2017
- Writer
- Editor
- ArtEZ Creative Writing magazine PERPLEX 2014 – 2016
- Editor
- Contributor
- Uitgeverij Virtùmedia 2014 – 2015
- Writer
- Editor