TECHNOLOGIES OF THE EXTENDED MIND: DEFINING THE ISSUES

Peter B. Reiner¹ and Saskia K. Nagel²

¹National Core for Neuroethics, University of British Columbia, Canada.

²Department of Philosophy at the University of Twente, Netherlands.

To appear in: *Neuroethics: Anticipating the Future*, J. Illes, ed., pp. 108-122 (2017)

ABSTRACT

Living in the modern world entails substantial interaction with information technologies. The ways in which people interact with these devices—how they enter into daily practices—has become so profound that they qualify as *technologies of the extended mind*. This chapter distinguishes between these devices acting as cognitive support versus becoming bona fide extensions of our minds, and argues that this latter, new reality has substantial implications for the field of neuroethics. As exemplars, the implications of these technologies of the extended mind for concepts of autonomy and privacy of thought, as well as for the debate regarding cognitive enhancement, are investigated. The chapter calls for a new framework for thinking about neuroethics for technology that takes into account not just the effects of technology upon the brain, but one that also includes a more expansive concept of the mind.

Key words: Technologies of the extended mind, autonomy, privacy of thought, cognitive enhancement, technological embeddedness

INTRODUCTION

One of the primary concerns of the field of neuroethics has been the sanctity of the mind. Whether the worry is over others gaining access to a person's most private thoughts or manipulating memories, increasing cognitive abilities beyond species-typical functioning or the authenticity of modern life, the recurring ethical issues center around the question of whether people are masters of their own destinies. The progress that has been made in exploring this fascinating terrain has been substantial, as evidenced by the growing maturity of the field. Yet even while making great strides in defending the boundaries of what is neuroethically acceptable, many people seem to be yielding the sanctity of their minds to the convenience of modern devices. People do not rely only upon their 3-pound brains to navigate the world that surrounds them. Increasingly, a substantial portion of the population is blending their cognitive space with the algorithmic devices that are nearly always at hand. It is time that we recognize these devices as technologies of the extended mind (TEMs) (Fitz & Reiner 2016; Nagel et al. 2016).

THE EXTENDED MIND HYPOTHESIS

The intellectual forerunner to the concept of TEMs is the extended mind hypothesis (EMH) (Clark & Chalmers 1998), which suggests that cognition extends beyond the brain into the world at large. Importantly, the EMH specifies that for external cognitive processes to qualify as part of the mind, they must be active parts of the mind. Thus, the EMH goes beyond merely suggesting that human cognition relies on external structures for scaffolding and support. Rather, the EMH maintains that at least some of the physical vehicles that realize our cognitive processes lie outside of the bounds of the skull.

The classic example from the original paper by Andy Clark and David Chalmers—the case of Otto and Inga—illustrates the issue nicely (Clark & Chalmers 1998). Otto and Inga live in New York City. One day, Inga hears about an exhibition at a museum that she recalls is on 53rd Street and heads out the door, intent on seeing the artwork. Her neighbor Otto has been having trouble remembering things. In order to overcome this deficit, he has made a practice of storing important information in a small notebook that he carries with him in his breast pocket. When he hears about the exhibition, he consults his notebook, finds that the museum is on 53rd Street, and, just like Inga, sets off for the same destination. The similarity between the two situations should now be obvious: the cognitive function of storing information is mediated by neurons in one case and pen and paper in the other.

From this example, Clark and Chalmers develop the parity principle, asserting that if a process that unfolds in the external world would readily be classified as part of the cognitive toolkit when it goes on in the head, then it is, at least for that point in time, part of the cognitive process. Using the parity principle as a guide, Clark and Chalmers assert the equivalence of neuronal memory and

paper memory as information storage strategies in the case of Otto and Inga.

The claims of the EMH are radical, and it remains a highly contentious theory in philosophy of mind (Adams & Aizawa 2001, 2010; Rupert 2004, 2013; Menary 2010). However, one need not fully accept the philosophical premise to appreciate that the concept resonates with a key feature of modern life: the growing sense that computers and smartphones (and soon “the Internet of Things”) function as sophisticated extensions of the modern cognitive toolkit, even more so than Otto's dog-eared notebook. Moreover, rather than relying upon the parity principle to guide thinking in this regard, we will provide further specification on the features and uses of these algorithmic devices that serve to qualify them as TEMs.

TECHNOLOGIES OF THE EXTENDED MIND

Before we consider the circumstances under which a device qualifies as a TEM, it is useful to explore what we mean when we use the word *mind*. A nuanced description of the term is beyond the scope of this chapter, and there exists an entire subdiscipline of philosophy that addresses the issue, but a few words of clarification are in order, in particular to place the notion of TEMs in context.

The *Oxford English Dictionary* defines mind as the element of a person that enables them to be aware of the world and their experiences, to think, and to feel: the faculty of consciousness and thought. Yet this definition is not satisfying, specifically because it dodges the elephant in the room: the mind–body problem. While we will use the term mind liberally in this chapter, it is not our intent to slip into some version of substance dualism in which there is brain-stuff and mind-stuff. But one specific distinction between brain and mind is in order: as we view it, brain is a *thing* while mind is a *concept*. Mixing those ontological levels is what often leads to confusion as to the relation between brain and mind.

One way of thinking about the issue is to say that the mind represents the full set of cognitive resources that we deploy in the service of thinking. Here we construe thinking to include what Keith Stanovich calls reflective, algorithmic, and autonomous thinking (Stanovich 2009). One can quickly see how this definition is friendly to the EMH, for once one uses the term full set of cognitive resources, one opens the door to things other than the brain contributing to mind. Such formulations represent direct challenges to the hard neuroessentialist perspective which, one of us has argued, suggests that there is no need to include anything other than the brain (Reiner 2011). Yet, as we shall see below, once one begins to give due consideration to things outside of the brain—in particular TEMs—it becomes difficult to conclude that they are not being deployed in the service of thinking. In this view, mental processes and mind cannot be fully reduced to cognitive processes, as they also refer to the plethora of affective, motivational, and social resources that can influence thinking and go beyond it in various

ways. While many current TEMs aim at the extension of cognitive processes, this is by no means a principal constraint; to the contrary, design might explicitly focus on an extension of affective processes as well.

Having provided some conceptual clarity over what constitutes mind, we now address the question of what sorts of devices are TEMs. It is not the case that every algorithmic function carried out by devices external to the brain qualifies them as a TEM, but rather that there is a relatively seamless interaction between brain and algorithm such that a person *perceives* of the algorithm as being a bona fide extension of a person's mind. This raises the bar for inclusion into the category of algorithms that might be considered TEMs. It is also the case that algorithmic functions which do not qualify as TEMs today may do so at some future point in time and vice versa.

By way of illustration, consider the use of GPS in a smartphone, an example we have previously described in Nagel et al. (2016). Imagine a 35-year-old man John who has lived in Manhattan for the past decade. John has rarely driven a car since he moved to New York, as he now relies heavily on the subway or taxis to get about town. But John recently had a baby and needs a bit of extra income, so he asks his brother whether he thinks he could become a driver for Uber, a company that enables non-professionals to act as taxi drivers using their own vehicles. His brother, who has been making a fair bit of money doing just that over the past year, is quite encouraging and even offers John the use of his car in the evenings when it is idle. John signs up and within a week finds himself behind the wheel of his brother's car, answering calls for Uber drivers to come and ferry people all over New York City.

In the normal course of events, Uber drivers are highly reliant on GPS. On his very first day, John, an avid user of computer technology, is fascinated by how easy it is to use the Uber app on his smartphone and have it show him the best route to his passengers' destinations. It is particularly helpful because otherwise John would often find himself lost, as many of his fares lead him to visit neighborhoods with which he is unfamiliar. Of course, John has heard stories that sometimes GPS can lead you to the wrong place, so he remains alert to his environment in order to be certain that he delivers his passengers to their destination without a hitch.

Is John's GPS functioning as a TEM? It is certainly carrying out computational work that is external to John's brain. But it is probably most appropriate to consider the GPS in John's smartphone as cognitive support, for neither the algorithmic calculations nor John's reliance upon them are seamlessly integrated with John's mind.

Now imagine that a week or two has passed. John has taken dozens of passengers to their destinations. Even though he now knows the city a bit better than before, he always uses the GPS in his smartphone, and it has not let him down even once. At this point, when the address and

route flashes up on the screen, he doesn't give it a moment's notice before following it to the destination suggested by his smartphone. The GPS now functions very much as a TEM, for John has integrated its algorithmic output into the working of his mind.

Much as neuroessentialism challenges traditional views of how people view themselves (Reiner 2011), conceiving of the mind as a blend between brain and algorithm poses a challenge to a range of existing worldviews. At one end of the spectrum, if one is troubled attempting to reconcile concepts of ensoulment with mechanistic explanations of how the brain works (Bering 2006; Farah & Murphy 2009; Preston et al. 2013), the realization that an algorithm is now an extension of one's mind makes the issue even more problematic. At the other end of the spectrum, the very same mechanistic explanations of brain function that lead to the position that "we *are* our brains" (Greene & Cohen 2004; Roskies 2007; Reiner 2011) are somewhat destabilized by the emerging reality of TEMs. For these reasons, people may very well resist the disruptive concept of TEMs. However, the intrusion of algorithms upon daily life seems relentless, and in much the same way as John's GPS transitioned from external computational algorithm to bona fide TEM, we envision this concept becoming an emergent meme, with an ever-growing share of the population perceiving of their devices as TEMs, even if they do not explicitly conceptualize them as such.

THE NEUROETHICAL ISSUES

The neuroethical issues engendered by the EMH were first brought to light by Neil Levy in a prescient paper published in 2007 (Levy 2007). At the time, he allowed that the EMH was an obscure debate, but nonetheless suggested that it had substantial implications for the field. He points out that part of what makes neuroethics a discipline is the claim, sometimes explicit but nearly always implicit, that there is something different about intervening in the brain, that such interventions are different than traditional means of altering mental states. Levy suggests that if one accepts the EMH, the exceptionalism that is normally offered to worries about intrusions, manipulations, and surveillance of mental states in the brain naturally flow to the extended mind as well. He sums the situation up nicely in his concluding paragraph:

These reflections on the *prima facie* parity between environmental manipulations and new technologies may seem to reduce the importance of neuroethics. On the contrary, I think that it dramatically increases it. It might seem that the extended mind thesis entails that neuroethics is less important inasmuch as, in its light, it becomes apparent that neurological interventions into the mind—that is, interventions that target neurons, neurotransmitters or brain structures—are not after all so special: they represent merely the latest means of doing something that is quite ubiquitous in human cultures. Although it is true that the extended mind thesis may dampen some

of the hype surrounding these technologies, it should be seen as dramatically expanding the *scope* of neuroethics, not detracting from its importance. Neuroethics focuses ethical thought on the physical substrate subserving cognition, but if we accept that this substrate includes not only brains, but also material culture, and even social structures, we see that neuroethical concern should extend far more widely than has previously been recognized. In light of the extended mind thesis, a great many questions that are not usually seen as falling within its purview—questions about social policy, about technology, about food and even about entertainment—can be seen to be neuroethical issues. (Levy 2007)

We agree with Levy that the EMH has profound implications for our thinking about neuroethics. Three issues are particularly worthy of further elaboration: the threat to autonomy posed by manipulation of our TEMs, the threat to privacy of thought by peering into TEMs, and the relevance of TEMs to questions regarding cognitive enhancement. Each has changed substantially in the decade since Levy's paper, a decade in which adoption of smartphones as ubiquitous computing has moved from concept to reality. This development is worth highlighting, as it seems to be more than just another technology, but rather has been a major influence that shapes our technological surroundings. Other influences are and will be relevant, but the prevalence of smartphones—nearly half the adult population on the planet has one (Pew Research Center 2016)—and our readiness to rely on them is certainly remarkable. These features make it a useful example with which to explore some of the neuroethical issues associated with TEMs.

AUTONOMY

The concept of autonomy is foundational to modern thinking about who we are as sentient beings. Particularly since the Enlightenment, Western societies have largely accepted the position that we—and only we—have the right to determine the course of our lives. Autonomy underpins many of the most celebrated political and social movements of modernity. In human rights, it is autonomy that beckons us to consider the needs and desires of the individual as having primacy. In politics, autonomy is the firmament upon which democratic governance rests. Autonomy shows no signs of losing its appeal, continually forcing us to modify our practices. This is best exemplified by the medical realm where modern bioethical principles have invoked the concept of autonomy to produce a sea change in the way that physicians and their patients interact (Beauchamp & Childress 2012).

A fundamental feature of the concept of autonomy is that the autonomous individual should not be unduly influenced when making decisions (Frankfurt 1971; Dworkin 1981). One need look no further than the ideal of the rugged individualist who navigates their environment relying on nothing more than his or her wits to see how this works. In this common trope, the autonomy of decision-making by the individual is held as

sacrosanct, and any infringement is considered a violation. It turns out that this picture of humans as self-sufficient rational actors who live their lives independently of others is at substantial variance with how people actually live in the real world. Rather, decisions are regularly influenced by the input of others, whether it is in the form of the books and newspapers that we read, the opinions of people that we listen to, or other features of the social and physical environment around us. Recognizing this, feminist and communitarian scholars have updated the concept, developing variants on what is commonly known as relational autonomy (Nedelsky 1989; Friedman 2003; Christman 2004; Meyers 2005; Mackenzie 2010). The relational account of autonomy suggests that when people make decisions, they often admit input from friends, family, colleagues, or professionals. That this occurs without demerit complicates the calculus for evaluating when an external influence might be considered due and undue. Indeed, “representing these two sorts of effects with roughly accurate proportionality is, however, a formidable project [since] matters of degree are notoriously difficult to specify philosophically” (Friedman 2003).

If it is a struggle to determine what influences are due and undue in the context of input from other people, the task is even more complicated when we consider the influence of TEMs upon these very same decisions. Before we do so, it is worth considering the general features of algorithms that might modify the degree to which an influence is perceived as violating autonomy. We have suggested that three factors are key: (1) the persuasiveness of the algorithm in the decision-making process, (2) the seriousness of the decision, and (3) the ability for the algorithm to learn about user preferences (Nagel et al. 2016).

Persuasiveness seems to be a central feature of technological autonomy violations. Decision-making can be influenced across a spectrum ranging from minimally to highly persuasive, and can even turn into being coercive (Fogg 2002; Verbeek 2006, 2009). If the ability to thoughtfully engage in the decision-making process and to reflect on the situation is ensured, the influence of the technology will not be perceived as a particularly problematic autonomy violation, as self-control does not seem to be harmed.

The seriousness of the decision also varies across a spectrum that differs depending upon the level of potential harm or benefit an individual may experience as a result from a particular choice. The lower the assumed potential harms or benefits, the lower the perceived seriousness of the decision. Imagine for a moment the grave real-world situation of an individual grappling with the decision of whether to undergo a course of chemotherapy that will briefly extend his or her life by some weeks but will compromise the quality of those weeks substantially. Many such patients consult—and trust—online resources to help them evaluate the relative risks and benefits. Under such circumstances, even a

small misstep in influence can result in a substantial autonomy violation.

The ability to learn about user preferences is important, as it makes a great difference as to whether a system only follows a set of preprogrammed instructions or is able to monitor individual behaviors and preferences and learns from them. If we return to the situation of the individual confronting the decision of whether to embark upon a course of chemotherapy, we can easily imagine two versions of the online resource. In the static version where the information that is found online is offered by an unknown designer, the opportunity for the designers' preferences to be unwittingly smuggled into the algorithmic calculation is high, and the possibility that an autonomy infraction may occur not insubstantial. A more dynamic version would have the underlying algorithm learn something about the preferences of the user and then provide advice accordingly. Since the algorithm has learned at least something about the user's worldview, the advice that it offers might better be tailored to their second-order desires, and the likelihood of an autonomy violation diminished.

Having laid the groundwork for our thinking about the issue, we turn our attention to an example in which the external technology is a TEM. We will use a simple and relatively trivial example to illustrate the relevant issues, but the reader can readily imagine how modifying persuasiveness, seriousness, and learning user preferences can affect the degree of perceived autonomy violation.

To begin, consider an algorithm that is not a TEM, for example, the GPS app on John's smartphone in the example described in the previous section. Imagine that on the first evening that John used the GPS function (i.e., before it has transitioned into being a TEM), John passes a billboard advertising a bakery. Pushing the bounds of current practice (but only slightly), let's also imagine the smartphone calls attention to the billboard and alerts John to the fact that the bakery is just up ahead; sealing the deal, the smartphone "asks" if John wishes to drop in to get something sweet. John is tempted, and although he likes to manage his intake of sugar, he decides that there would be little harm. A moment later he stops at the bakery, purchases a donut, and soon is smiling as he tastes the sugary treat. In this scenario, the GPS program has clearly influenced John, sufficiently so that he altered his second-order desires about his food intake. Many people would call such an influence undue and, despite the relative triviality of the infraction, the situation represents an autonomy violation of sorts.

Now let us imagine that 2 weeks have passed, and the relationship between John and the GPS app has grown more intimate—John now integrates its algorithmic output into the working of his mind while driving his Uber cab. For all intents and purposes, the GPS algorithm is now functioning as a TEM. Of course, because John carries his smartphone everywhere, the device gathers a fair bit of information about his daily activities, and the abilities that this confers on the

smartphone only serve to reinforce the feeling that the suite of apps on the phone are functioning as a more-or-less unified TEM. It requires little in the way of stretching credulity to imagination that the GPS might not only call attention to a billboard alongside the road, but having "consulted" a database that indicates that John frequently shops at an organic grocery store, instead of suggesting that John get something sweet, the smartphone now suggests that he may wish to get a piece of organic blueberry pie. Adding to the persuasiveness of the suggestion, the GPS program on his smartphone—which also monitors his activity levels during the day—reminds John that he went for a 5K run this morning and can probably afford the calories. A few minutes later, John is smiling as he enjoys a forkful of delicious pie. Once again, the GPS program has influenced John, sufficiently that he altered his second-order desires about his food intake. But given that the suggestions are aligned with John's overall desire to eat organic food, and that they include at least a rough calculation on calories burned earlier in the day, does the persuasive influence of the GPS app represent an autonomy violation?

The difference between the two scenarios described above depends to a substantial degree on how John perceives the algorithmic device that accompanies him everywhere. As his relationship to the smartphone becomes closer, as both he and device provide information that the other uses to help them navigate the world around them and John perceives more and more that he relies on the program's advice, the device increasingly becomes an integral part of his extended mind. Given the considerations from relational autonomy, the influence of the smartphone on John's behavior in the second scenario appears to be less "undue" than it was in the first instance. In this view, the more that devices become TEMs, the likelihood that their advice results in perceived autonomy violations diminishes.

However, there is something missing from this description, and that is that the very same algorithm that is an extension of John's mind is also an extension of the mind of an other—in this case the corporate entity that has designed the GPS app. It is not at all out of the bounds of modern economic activity that the corporate entity would be paid for directing John to the bakery. Thus, while one objective of that app is to align John's behavior with his second-order desires, a second objective of another entity is also smuggled into the relationship. In view of such potential conflicts of interest, it becomes harder to accept the premise that as our devices become TEMs, autonomy violations become less likely; if anything, they become more insidious. Serving two masters is certainly part of the issue, but it is also the case that the privacy implications are impossible to ignore.

PRIVACY OF THOUGHT

One of the most controversial topics in neuroethics is the worry over *mind reading* (Wolpe et al. 2005; Illes 2006,

2007; Greely & Illes 2007; Parens & Johnston 2014). Frances Shen has summarized the sentiments of many by suggesting that using advances in neuroscience to access the most private of spaces—our minds—would have tremendous privacy implications (Shen 2013). Yet despite intriguing insights obtained with functional magnetic resonance imaging (Haynes & Rees 2006; Mitchell et al. 2008; Naselaris et al. 2009; Rissman et al. 2010), it remains difficult to obtain more than a rudimentary profile of the thoughts of individuals by measuring brain activity (Farah et al. 2009). In contrast, it seems that a great deal of what might be termed the inner life of the mind may be accessible via the technologies on our desktops, and even more so, in our pockets. Cognizant that this is a discourse that has encountered more than its fair share of hype, it seems prudent to consider whether access to our TEMs may represent intrusions upon privacy of thought.

Such considerations come in the context of the well-known fact that online privacy is generally under threat. Illegal breaches and over-sharing of digital information have grown from the occasional to everyday events. While any intrusion on privacy may be unwelcome, some transgressions are more problematic than others. When others gain access to information that reflects one's innermost thoughts, the incursion may go beyond privacy to intrude upon privacy of thought. Such intrusions have particular normative significance in neuroethics, and thus our analysis focuses upon this subset of privacy concerns.

The manifesto establishing the right to privacy was drafted in 1890 by Samuel Warren and Louis Brandeis (Warren & Brandeis 1890). As influential as it has been in legal circles, it is the historical sweep of the article that draws one's awe. As Warren and Brandeis incisively explain, at one time liberty meant freedom from physical restraint, but as society increasingly recognized the inner life of individuals, the right to life came to mean the right to enjoy life, and protection of corporeal property expanded to include protection of the products of the mind such as literature and art, trademarks, and copyrights. Warren and Brandeis noted that "recent inventions and business methods call attention to the next step which must be taken for the protection of the person." In 1890, their worry was the intrusiveness of photographers, but more generally their article suggested that as technology changes, societal norms might require updating. In the modern world, in which information all-too-readily moves online, we appear to be overdue for just such a review.

Legal arguments support such a view. Pointing out that under the third-party doctrine "an individual does not have a reasonable expectation of privacy with respect to information he voluntarily discloses to a third party, like a bank or a telecommunications carrier," Wittes and Chong suggest that the third-party doctrine is ill-suited to the modern situation in which devices generate data about their users (Wittes & Chong 2014). In a chilling demonstration of how much information is contained

within metadata—the data about the numbers we call, or where our phone is at a particular moment, who we e-mail and the subject lines of those messages, our search queries and the websites we visit—Ton Siedsma, a researcher at the Dutch digital rights foundation Bits of Freedom, allowed an app on his cellphone to collect such metadata for 1 week (Tokmetzis 2014). The limited analysis of this information revealed not only Ton's circles of friends and the like, but also the types of information included in his search queries. That Ton might have an interest in bicycles is fairly innocuous, but what if he was searching for information about antidepressants or gender reassignment surgery, and these were topics that were in his head, but he had not—and more importantly did not wish to discuss with others?

Just such a situation is described by Wittes and Lui when alerting us to the fact that "the privacy that consumers value in practice is not always the privacy that activists devoted to privacy value on their behalf." To illustrate the matter, they point out that many people acquire medical information, salacious images, and more online because "they would rather be tracked online by commercial vendors than have to face parents, teachers, doctors, or the stern-faced old lady at a news stand" (Wittes & Liu 2015). In a similar vein, Frances Shen reminds us that what he terms the "privacy panic script" unfolds first with fear mongering regarding the power of technology, followed by the suggestion that institutions will use the technology in devious ways, and that current laws are insufficient to protect the average citizen (Shen 2013). These observations suggest that when advocating in favor of privacy, humility may be in order.

Given these varied interests, we suggest that explicitly recognizing some of our devices as TEMs may help us in arriving at a more reasonable state of affairs, especially if we arrive upon a clearer understanding of what we mean by privacy of thought. The philosopher Michael Lynch, writing as *Amicus Curiae* in support of the plaintiffs in *American Civil Liberties Union v. James Clapper*, the lawsuit that challenged the legality of the National Security Agency's bulk phone metadata collection program traces the notion of privacy of thought back to Descartes and Locke, who both held that:

what identifies your thoughts as your thoughts is that you have "privileged access" to them. This means at least two things. First, you access them in a way I cannot. Even if I could walk a mile in your shoes, I cannot know what you feel in the same way you can: you see it from the inside, so to speak. Second, you can, at least sometimes, control what I know about your thoughts. You can refrain from telling me the extent of your views and your feelings. (Lynch 2013)

In the context of TEMs, these criteria seem appropriate: if you have a private means of accessing information on your device that is not available to others, and you have the means of controlling that access, it seems as if you have demonstrated that, in your view, it is private. The

problem arises, of course, when those very same devices are connected to others by virtue of the ever-present Internet.

To see how this may play out, it is worth recalling the infamous case of the discount retailer Target and the pregnant teen. By combining metadata that was packaged, bought, and sold (online, of course) with analysis of something as innocent as the shopping behavior of young women, Target was able to predict with high accuracy when women were in their second trimester of pregnancy (Hill 2012). The commercial objective was to target these potential customers with ads for diapers, baby clothes, and the like, but when Target did so, customers felt so invaded that they complained (in this case, it was a father who only discovered his daughter's pregnancy through this route). So Target modified their strategy; instead of sending ads just for diapers, etc., they mixed those ads in with ads for everyday items unrelated to pregnancy just to prevent the perception that they had invaded their customer's privacy.

Similar brinkmanship can be seen in the relationship between Facebook and its users. In a longitudinal study of 5076 Facebook users between 2005 and 2011, Facebook users increasingly exhibited privacy-seeking behavior, decreasing the amount of personal data that they shared publicly. At the same time, the amount of personal information they revealed privately to so-called connected profiles increased. The unintended result was that disclosures to "silent listeners" on the network—Facebook itself, third-party apps, and ultimately advertisers—grew. The authors conclude that these findings "highlight the tension between privacy choices as expressions of individual subjective preferences, and the role of the environment in shaping those choices" (Stutzman et al. 2012).

From these examples we can see that the networked nature of our TEMs makes them particularly vulnerable to intrusions upon privacy of thought. While people may take steps toward increasing the privacy of their online behavior, the corporate entities at the other end of the phone, as it were, are continually mining these very same actions in an effort to enhance their competitive position in the world of commerce. What is lost in the conversation is that along the way, people may *perceive* of their devices as extensions of their minds. We suggest that this implicit tagging of Internet-connected algorithmic devices as TEMs raises the bar for privacy, qualifying the information within the amalgam of our brains and our devices as private thoughts.

COGNITIVE ENHANCEMENT

Discussion of the propriety of using advanced neurotechnologies to enhance cognitive abilities has been among the most protracted of debates in neuroethics (Parsons 1998; Farah et al. 2004; Greely et al. 2008). Ethical arguments from different biopolitical positions (Reiner 2013) have been waged over a number of issues, but the cardinal concerns of safety, pressure, distributive

justice, and authenticity have dominated (Fitz et al. 2014). In many ways, the debate is more pragmatic than philosophical, and so we will restrict our comments regarding the implications of TEMs to the cognitive enhancement debate to a short outline of the relevant issues.

The debate over cognitive enhancement is somewhat compromised by the observation that the agents that are available today—drugs and electrical stimulation devices—do not yield the sorts of enhancements that people seek: effect sizes tend to vary from minimal to none (Farah 2015; Ilieva et al. 2015). As we have recently suggested, this may be because attempts to bolster cognitive function from within run up against the hard limits of neurobiological reality (Fitz & Reiner 2016). Indeed, we suggest that a better strategy is to rely upon our TEMs to enhance cognitive function, not only because it is more effective but because we are already well down the road toward "commingling our cognitive space with technology."

The question that remains open, and one that merits further normative and empirical work in the field, is the degree to which considering TEMs as cognitive enhancers settles debates that have gone on for over a decade now. The answer is likely to be a mixed bag—the ubiquity of TEMs might help to dispel the distributive justice concern, but concerns over safety, pressure, and authenticity now morph into new territory. We predict that exploring these issues with a new technological lens may assist the field in clarifying how we should move forward, even if it involves playing catch-up to how we live today.

TEMS NECESSITATE A NEW FRAMEWORK FOR NEUROETHICS

The many important issues that have made neuroethics such a vibrant and fascinating field—ranging from questions on handling of incidental findings, brain reading to detect deception, brain intervention of offenders, to enhancement of healthy people by means of pharmacological agents or electrical brain stimulation—have largely been focused on the brain. Recognizing that TEMs qualify as extensions of the mind provides the impetus for "dramatically expanding the scope of neuroethics," as Neil Levy has argued (Levy 2007). This expansion is even more interesting the more one recognizes that the role of technologies is not just one of cognitive support but is fundamental to the modern way of being in our technological culture.

In philosophy, it has long been well accepted that humans cannot be understood and often do not conceive of themselves as isolated beings. Rather, people are *socially embedded*, existing as part of a collection of other beings who shape us and our image of ourselves. This idea of our deeply social nature can be found already in the writings of Hegel and Marx, and has been extensively developed by the critical psychologist Klaus Holzkamp (Holzkamp 1985). Importantly, in this view social embeddedness goes beyond the direct private

relations between and among individuals, reaching out to the various social institutions that surround and influence us.

Analogously, the modern world with its increasingly technological surroundings defines us as *technologically embedded* beings. Consider that today the trajectory of a life begins surrounded by reproductive technologies and often ends with a different array of medical technologies easing us into death. Just try to imagine for a moment which technologies are around you right now—from the lightbulb above you, to the heat in the corner, to the smartphone in your pocket. Perhaps you took a drug this morning to help normalize some bodily function, or maybe you will use a car or bus or train to move from place to place, or even have a video-chat with a colleague on a different continent. Indeed, humans rely on a vast array of technologies, and one might argue that they depend upon them—some of us more, some less. Such is the modern predicament (Scialabba 2011) that it is hard to hear silence, that many live in fear of the content of their food, that it has become a common experience to feel estranged from the natural world. No matter how one *evaluates* being surrounded by and dependent upon technologies, it is fair to say that we are increasingly embedded by technologies.

It is no surprise then that some of those technologies engage with our cognitive processes. Even more: that they serve as TEMs by functioning as part of the mind that resides outside of the brain. And as the mind is perceived as perhaps the most intimate part of us, conceptualizing technologies as extensions of the mind deeply challenges the sense of who we are as humans. The questions Who am I? Am I my brain? Am I my mind? Am I my relationships? (Haslam 2004; Rose 2005; Feinberg 2006; Glannon 2009; Burwood 2009; Noë 2009; Brand 2010; Brenninkmeijer 2010; Pardo & Patterson 2010; Reiner 2011) are viewed from a different angle when a person recognizes that technologies serve as a functional part of the self. Ultimately, with this perspective, humans must find a new mode of being, and consequently a new mode of understanding ourselves and others. For if our cognitive processes are not restricted to our biological brains, we need to give serious consideration to their role in our being.

An important result is that the relationship that people have with those devices merits much closer attention than they have received to date. We might go as far as arguing that we need to consider TEMs as new actors on the neuroethical stage. Thus, we call for a new framework for neuroethics that is informed by philosophy of mind which needs to reconsider how humans conceive of who they are today: What is at stake when studying neuroethical questions? What is the “substrate” we are dealing with? If we take the nature of the intertwinement of modern humans with their surroundings seriously—and we argue that we should—we cannot any longer focus on the human brain as isolated or, at most, integrated into a body. The body and the brain are not only in constant interaction with the

environment (Varela et al. 1991; Gallagher 2005; Fuchs 2011; Nagel 2015), they also constantly make use of their surroundings, of significant and non-significant others, and of technologies. Thus, in this new framework, neuroethics is not so much concerned with an organ but ultimately with what makes us human beings, and recognizing the impact of technology upon what makes us “us” today matters for the relevant questions in the field.

REFERENCES

- Adams, F. and Aizawa, K. (2001). The bounds of cognition. *Philosophical Psychology*, 14(1), 43–64.
- Adams, F. and Aizawa, K. (2010). Defending the bounds of cognition. In: Menary, R. (Ed.) *The Extended Mind*, pp. 67–80. Cambridge, MA: MIT Press.
- Beauchamp, T.L. and Childress, J.F. (2012). *Principles of Biomedical Ethics* (7th edn.). New York: Oxford University Press.
- Bering, J. (2006). The folk psychology of souls. *Behavioral and Brain Sciences*, 29(5), 453–462.
- Brand, C. (2010). Am I still me? Personal identity in neuroethical debates. *Medicine Studies*, 1(4), 393–406.
- Brennkmeijer, J. (2010). Taking care of one's brain: how manipulating the brain changes people's selves. *History of the Human Sciences*, 23(1), 107–126.
- Burwood, S. (2009). Are we our brains? *Philosophical Investigations*, 32, 113–133.
- Christman, J. (2004). Relational autonomy, liberal individualism, and the social constitution of selves. *Philosophical Studies*, 117(1), 143–164.
- Clark, A. and Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7–19.
- Dworkin, G. (1981). The concept of autonomy. *Grazer Philosophische Studien*, 12, 203–213.
- Farah, M.J. (2015). The unknowns of cognitive enhancement. *Science (New York, NY)*, 350, 379–380.
- Farah, M.J. and Murphy, N. (2009). Neuroscience and the soul. *Science (New York, NY)*, 323(5918), 1168.
- Farah, M.J., Illes, J., Cook-Deegan, R., et al. (2004). Neurocognitive enhancement: what can we do and what should we do? *Nature Reviews Neuroscience*, 5(5), 421–425.
- Farah, M.J., Smith, M.E., Gawuga, C., Lindsell, D., and Foster, D. (2009). Brain imaging and brain privacy: a realistic concern? *Journal of Cognitive Neuroscience*, 21(1), 119–127.
- Feinberg, T.E. (2006). Our brains, our selves. *Dædalus*, 135(4), 72–80.
- Fitz, N.S. and Reiner, P.B. (2016). Time to expand the mind. *Nature*, 531, S9.
- Fitz, N.S., Nadler, R., Manogaran, P., Cong, E.W.J., and Reiner, P.B. (2014). Public attitudes toward cognitive enhancement. *Neuroethics*, 7(2), 173–188.
- Fogg, B.J. (2002). *Persuasive Technology: Using Computers to Change What We Think and Do*. San Francisco, CA: Morgan Kaufmann.
- Frankfurt, H.G. (1971). Freedom of the will and the concept of a person. *Journal of Philosophy*, 68(1), 5–20.
- Friedman, M. (2003). Autonomy and social relationships: rethinking the feminist critique. In: *Autonomy, Gender, Politics*, pp.81–97. Oxford: Oxford University Press.
- Fuchs, T. (2011). The brain—a mediating organ. *Journal of Consciousness Studies*, 18, 196–221.
- Gallagher, S. (2005). *How the Body Shapes the Mind*. New York: Oxford University Press.
- Glannon, W. (2009). Our brains are not us. *Bioethics*, 23(6), 321–329.
- Greely, H.T. and Illes, J. (2007). Neuroscience-based lie detection: the urgent need for regulation. *American Journal of Law & Medicine*, 33(2–3), 377–431.
- Greely, H.T., Sahakian, B., Harris, J., et al. (2008). Towards responsible use of cognitive-enhancing drugs by the healthy. *Nature*, 456(7223), 702–705.
- Greene, J. and Cohen, J. (2004). For the law, neuroscience changes nothing and everything. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359(1451), 1775–1785.
- Haslam, N. (2004). Essentialist beliefs about personality and their implications. *Personality and Social Psychology Bulletin*, 30(12), 1661–1673.

- Haynes, J.-D. and Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, 7(7), 523–534.
- Hill, K. (2012). *How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did*. Available from: <http://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/4/#583f593d34c6> [accessed March 27, 2016].
- Holzkamp, K. (1985). *Grundlegung der Psychologie*. Frankfurt: Campus-Verlag.
- Ilieva, I.P., Hook, C.J., and Farah, M.J. (2015). Prescription stimulants' effects on healthy inhibitory control, working memory, and episodic memory: a meta-analysis. *Journal of Cognitive Neuroscience*, 27(6), 1069–1089.
- Illes, J. (Ed.) (2006). *Neuroethics: Defining the Issues in Theory, Practice, and Policy*. Oxford: Oxford University Press.
- Illes, J. (2007). Empirical neuroethics. Can brain imaging visualize human thought? Why is neuroethics interested in such a possibility? *EMBO Reports*, 8, S57–S60.
- Levy, N. (2007). Rethinking neuroethics in the light of the extended mind thesis. *American Journal of Bioethics*, 7(9), 3–11.
- Lynch, M. (2013). *Brief of Michael P. Lynch as Amicus Curiae in Support of the Plaintiffs, American Civil Liberties Union v. James R. Clapper*. Available from: <https://www.aclu.org/legal-document/aclu-v-clapper-amicus-brief-michael-p-lynch-philosophy-professor-university> [accessed May 1, 2017]
- Mackenzie, C. (2010). Imagining oneself otherwise. In: Mackenzie, C. and Stoljar, N. (Eds.) *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, pp.124–150. New York: Oxford University Press.
- Menary, R. (2010). *The Extended Mind*. Cambridge, MA: MIT Press.
- Meyers, D. (2005). Decentralizing autonomy: five faces of selfhood. In: Christman, J. and Anderson, J. (Eds.) *Autonomy and the Challenges to Liberalism: New Essays*, pp.27–55. New York: Cambridge University Press.
- Mitchell, T.M., Shinkareva, S.V., Carlson, A., et al. (2008). Predicting human brain activity associated with the meanings of nouns. *Science (New York, NY)*, 320(5880), 1191–1195.
- Nagel, S.K. (2015). Thickening descriptions with views from pragmatism and anthropology. In: Metzinger, T. and Windt, J.M. (Eds.) *Open MIND*, pp. 829–839. Cambridge, MA: MIT Press.
- Nagel, S.K., Hrinco, V., and Reiner, P.B. (2016, May 13–14). *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, BC.
- Naselaris, T., Prenger, R.J., Kay, K.N., Oliver, M., and Gallant, J.L. (2009). Bayesian reconstruction of natural images from human brain activity. *Neuron*, 63(6), 902–915.
- Nedelsky, J. (1989). Reconceiving autonomy: sources, thoughts and possibilities. *Yale Journal of Law & Feminism*, 1(1), 5.
- Noë, A. (2009). *Out of Our Heads: Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness*. New York: Hill and Wang.
- Pardo, M.S. and Patterson, D. (2010). Philosophical foundations of law and neuroscience. *University of Illinois Law Review*, 5, 1211–1250.
- Parens, E. (1998). *Enhancing Human Traits: Ethical and Social Implications*. Washington, DC: Georgetown University Press.
- Parens, E. and Johnston, J. (2014). Neuroimaging: beginning to appreciate its complexities. *Hastings Center Report*, 44(s2), S2–S7.

- Pew Research Center (2016). *Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies*. Available from: <http://www.pewglobal.org/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/> [accessed January 31, 2017].
- Preston, J.L., Ritter, R.S., and Hepler, J. (2013). Neuroscience and the soul: competing explanations for the human experience. *Cognition*, 127(1), 31–37.
- Reiner, P.B. (2011). The rise of neuroessentialism. In Illes, J. and Sahakian, B.J. (Eds.) *Oxford Handbook of Neuroethics*, pp.161–175. Oxford: Oxford University Press.
- Reiner, P.B. (2013). The biopolitics of cognitive enhancement. In: Hildt, E. and Franke, A. (Eds.) *Cognitive Enhancement: An Interdisciplinary Perspective*, pp.189–200. New York: Springer Science+Business Media.
- Rissman, J., Greely, H.T., and Wagner, A.D. (2010). Detecting individual memories through the neural decoding of memory states and past experience. *Proceedings of the National Academy of Sciences of the United States of America*, 107(21), 9849–9854.
- Rose, S.P.R. (2005). Human agency in the neurocentric age. *EMBO Reports*, 6(11), 1001–1005.
- Roskies, A.L. (2007). The illusion of personhood. *American Journal of Bioethics*, 7(1), 55–57.
- Rupert, R.D. (2004). Challenges to the hypothesis of extended cognition. *Journal of Philosophy*, 101(8), 389–428.
- Rupert, R.D. (2013). Distributed cognition and extended mind theory. In: Kaldis, B. (Ed.) *Encyclopedia of Philosophy and the Social Sciences*, pp.209–213. Thousand Oaks, CA: Sage Publications Inc.
- Scialabba, G. (2011). *The Modern Predicament*. Boston, MA: Pressed Wafer.
- Shen, F.X. (2013). Mind, body, and the criminal law. *Minnesota Law Review*, 97, 2036–2175.
- Stanovich, K.E. (2009). Distinguishing the reflective, algorithmic, and autonomous minds: is it time for a tri-process theory. In Evans, J. and Frankish, K. (Eds.) *In Two Minds: Dual Processes and Beyond*, pp.55–88. Oxford: Oxford University Press.
- Stutzman, F., Gross, R., and Acquisti, A. (2012). Silent listeners: the evolution of privacy and disclosure on Facebook. *Journal of Privacy and Confidentiality*, 2, 7–41.
- Tokmetzis, D. (2014). *How Your Innocent Smartphone Passes on Almost Your Entire Life to the Secret Service*. Available from: <https://www.bof.nl/2014/07/30/how-your-innocent-smartphone-passes-on-almost-your-entire-life-to-the-secret-service/> [accessed March 27, 2016].
- Varela, F.J., Thompson, E., and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
- Verbeek, P.-P. (2006). Persuasive technology and moral responsibility toward an ethical framework for persuasive technologies. *Persuasive*, 6, 1–15.
- Verbeek, P.-P. (2009). Ambient intelligence and persuasive technology: the blurring boundaries between human and technology. *NanoEthics*, 3(3), 231–242.
- Warren, S.D. and Brandeis, L.D. (1890). The right to privacy. *Harvard Law Review*, 4(5), 193–220.
- Wittes, B. and Chong, J. (2014). *Our Cyborg Future: Law and Policy Implications*. Center for Technology Innovation at Brookings. Available from: <https://www.brookings.edu/research/our-cyborg-future-law-and-policy-implications/> [accessed January 31, 2017].
- Wittes, B. and Liu, J.C. (2015). *The Privacy Paradox: The Privacy Benefits of Privacy Threats*. Center for Technology Innovation at Brookings. Available from: https://www.brookings.edu/wp-content/uploads/2016/06/Wittes-and-Liu_Privacy-paradox_v10.pdf [accessed January 31, 2017].
- Wolpe, P., Foster, K., and Langleben, D. (2005). Emerging neurotechnologies for lie-detection: promises and perils. *American Journal of Bioethics*, 5(2), 39–49.