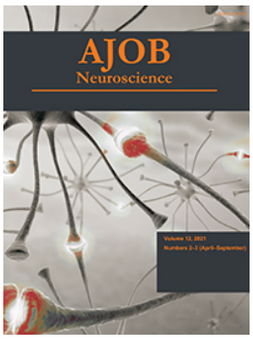


AJOB Neuroscience

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/uabn20

Mapping the Dimensions of Agency

Andreas Schönau, Ishan Dasgupta, Timothy Brown, Erika Versalovic, Eran Klein & Sara Goering

To cite this article: Andreas Schönau, Ishan Dasgupta, Timothy Brown, Erika Versalovic, Eran Klein & Sara Goering (2021) Mapping the Dimensions of Agency, AJOB Neuroscience, 12:2-3, 172-186, DOI: 10.1080/21507740.2021.1896599

To link to this article: https://doi.org/10.1080/21507740.2021.1896599

© 2021 The Author(s). Published with license by Taylor & Francis Group, LLC

Published online: 25 Mar 2021.

Submit your article to this journal

Article views: 4036

View related articles

View Crossmark data

Citing articles: 5 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=uabn20

AJOB NEUROSCIENCE
2021, VOL. 12, NOS. 2–3, 172–186
https://doi.org/10.1080/21507740.2021.1896599

Taylor & Francis
Taylor & Francis Group

TARGET ARTICLE

OPEN ACCESS Check for updates

Mapping the Dimensions of Agency

Andreas Schönau[a], Ishan Dasgupta[a], Timothy Brown[a], Erika Versalovic[a], Eran Klein[a,b], and Sara Goering[a]

[a]University of Washington; [b]Oregon Health and Science University

ABSTRACT

Neural devices have the capacity to enable users to regain abilities lost due to disease or injury – for instance, a deep brain stimulator (DBS) that allows a person with Parkinson's disease to regain the ability to fluently perform movements or a Brain Computer Interface (BCI) that enables a person with spinal cord injury to control a robotic arm. While users recognize and appreciate the technologies' capacity to maintain or restore their capabilities, the neuroethics literature is replete with examples of concerns expressed about agentive capacities: A perceived lack of control over the movement of a robotic arm might result in an altered sense of feeling responsible for that movement. Clinicians or researchers being able to record and access detailed information of a person's brain might raise privacy concerns. A disconnect between previous, current, and future understandings of the self might result in a sense of alienation. The ability to receive and interpret sensory feedback might change whether someone trusts the implanted device or themselves. Inquiries into the nature of these concerns and how to mitigate them has produced scholarship that often emphasizes one issue – responsibility, privacy, authenticity, or trust – selectively. However, we believe that examining these ethical dimensions separately fails to capture a key aspect of the experience of living with a neural device. In exploring their interrelations, we argue that their mutual significance for neuroethical research can be adequately captured if they are described under a unified heading of agency. On these grounds, we propose an "Agency Map" which brings together the diverse neuroethical dimensions and their interrelations into a comprehensive framework. With this, we offer a theoretically-grounded approach to understanding how these various dimensions are interwoven in an individual's experience of agency.

KEYWORDS

Neuroethics; brain computer interfaces (BCIs); deep brain stimulation (DBS); personal identity; privacy; responsibility

INTRODUCTION

Most end users of neural technologies are active agents who seek to express themselves—their feelings, emotions, thoughts, and desires—through goal-directed actions. Often, a neural device enables end users to regain abilities lost due to a disease or an injury. A person with Parkinson's disease, for example, may benefit from a deep brain stimulator (DBS) that alleviates tremor and rigidity, and thus restores the ability to fluently perform movements. A person living with spinal cord injury may benefit from a brain computer interface (BCI) to control a robotic arm, or even to regain a lost sensation of touch. A person with amyotrophic lateral sclerosis (ALS) may use a BCI to communicate with loved ones through the translation of thought to computer-generated speech. A depressed person may use a DBS to improve mood, in the hope of regaining a brighter, more authentic self.

Across these different types of neural devices, users actively aim at particular goals, and use devices in service of achieving them. Reporting on an interview study, Kögel, Jox, and Friedrich (2020) argue that BCI users perceive themselves as "active operators" who recognize and appreciate a neural device's capacity to maintain or restore human capabilities. Nonetheless, the neuroethics literature is replete with examples of end users expressing concerns, such as a sense of alienation or a perceived lack of control (See Table 1). These examples have motivated scholarship around ethical dimensions most commonly implicated by user experiences—responsibility, privacy, and authenticity. To what extent is a person with Parkinson's disease responsible for DBS-mediated actions that result in unintended outcomes? How can a person with ALS,

CONTACT Andreas Schönau schoenau@uw.edu Department of Philosophy, University of Washington, Savery Hall 378, Seattle, 98195-0005, USA.

© 2021 The author(s). Published with license by Taylor & Francis Group, LLC
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

Table 1. Four ethical dimensions in neurotechnology research.

Dimensions	Agency Concerns	Quotes
Responsibility	*Did I do that?* *Am I responsible?*	"I guess I didn't concentrate hard enough. But it also may be that the measuring is not optimal." (Kögel, Jox, and Friedrich 2020) "I've begun to wonder what's me and what's the depression, and what's the stimulator. I mean, for example, I can be fine, and then all of a sudden … I'll say something that is insensitive or just misread a person entirely, say something that either makes *me* look like a fool, or hurts them, or, something along that line. I can't really tell the difference. There are three things—there's me, as I was, or think I was; and there's the depression, and then there's depression, and the device and … it blurs to the point where I'm not sure, frankly, who I am" (Klein et al. 2016).
Privacy	*Who has access to information about me?* *How can I limit that access or extend it to the people I choose?*	"I wouldn't want anyone to know my feeling when I see my husband, or go to bed with my husband, … see my doctor, or any of those things" (unpublished, DBS user at MGH).
Authenticity	*Is this me?* *Is this who I want to be?*	"I did not like that at all … No, that clearly didn't fit with who I am … it was really too much; that really wasn't me, you know; I really felt as if there was someone [else] standing next to me … " (Haan 2017).
Trust	*Did that really just happen? Can I trust myself to know if it did?* *Do my sensations feel 'natural'?*	"Both subjects reported that the sensations elicited by electrical stimulation of the SI cortex felt 'unnatural' and unlike anything they had ever felt before" (Collins et al. 2017).

Each ethical dimension (responsibility, privacy, authenticity, trust) is aligned with sample core questions most strongly associated with the respective dimension and relevant quotes from individual end users.

using a BCI device to communicate, protect against the inadvertent sharing of private thoughts? Can a person using a DBS for depression experience changes in mood that feel inauthentic? Most inquiries into the nature of these concerns emphasize one of the above issues, at the expense of overlooking how these issues are interconnected. In addition, the ethical dimension of trust—how (sensory) experience mediated by devices and sent directly to the cortex can alter a person's trust in what they perceive, and thus threaten self-trust regarding what they know—has been under-explored in the neuroethics literature.

We believe that examining these ethical dimensions separately fails to capture a key part of the experience of living with a neural device: how they may influence one's overall agency (the capacity to enact one's intention on the world) and the experience of agency (the phenomenal component of exercising agency or what it is like to enact one's intention on the world). These four dimensions of agency are intricately inter-related in an individual's experience of agency. Questioning one's responsibility, for instance, may lead to feelings of confusion over authenticity ("Did I really do that? Is this me?"). Concerns about the trustworthiness of sensory experience resulting from direct stimulation of the cortex may raise worries about protecting the privacy of one's mind ("Can I trust that feeling? Could someone change my private experiences?"). A narrow analysis of device "side effects" in a particular area overlooks the ways in which devices can alter multiple parts of a user's experience, and how a treatment may shift problems from one dimension to another. Specifically, it overlooks how these various dimensions are interwoven in an individual's experience of *agency*.

There are two things worth noting here. The first is that an agent's perception of each interconnected dimension of agency can come apart from external metrics (such as how it is viewed by others or readings from devices).[1] Imagine someone who begins acting out of character after getting DBS (say, acting more impulsively [Haan et al. 2015]). This behavior may perplex loved ones, but the individual may not even recognize that a change has occurred. Or conversely, someone may feel inauthentic, but others may not find their "new" choices or behavior at all out of character. A similar sort of disconnect can occur with other dimensions as well. Whether someone feels that privacy has been violated (or not) can come apart from whether privacy has in fact been violated (e.g., an unfounded suspicion that data is being used in an unapproved manner). Something similar can be said of trust in one's sensory experience or feelings of responsibility. One may have good reasons to trust, but still find themselves unable to trust. And one may feel responsible for outcomes over which one has no causal connection, and the opposite can be true as well (Haselager 2013). In what follows, we focus on the subjective experience of these dimensions of agency, all the while recognizing the role that external sources of information play in shaping one's perceptions.

[1]We recognize that some might refer to external metrics, like the views of others or clinical tests, as being objective. We avoid the language of objective and subjective to highlight that external metrics can sometimes be faulty and convey a distorted reality. We simply want to distinguish between an agent's first person perspective and other external factors that may shape agency.

The second thing worth noting is that although the neuroethics literature has devoted significant attention to the related issue of autonomy (e.g., Kellmeyer et al. 2016; Gilbert 2015; Kramer 2011; Glannon 2009), we focus here on agency, in part given its foundational relationship to autonomy. Agency involves the capacity to act; autonomous agency requires action that is self-governed (i.e. not subject to undermining influences; what counts as autonomous agency is clearly widely debated, see Buss and Westlund 2018). As such, agency is a necessary condition for autonomy, but it is a lower level phenomenon, i.e., part of what makes autonomy possible (Sherwin 1998). We want to call attention to the reality that neural devices can alter our very capacity to act—to make choices and enact our intentions on the world—in ways that precede questions about whether or not those actions are autonomous.

In this paper, we offer a map of the dimensions of agency and their interrelations as they are implicated in neural technology, situating this map within the context of existing neuroethics literature and reports from users of neurotechnologies. The paper is structured as follows. In section two, we summarize how each of the four ethical dimensions have been discussed in the neuroethics literature. In section three, we organize these seemingly disparate dimensions into a coherent conceptual Agency Map. In section four, we present three case studies to illustrate how our Agency Map can be used to bring agential concerns into focus—both to highlight problems of agency due to disease or disability (that users hope device adoption will remedy or enhance) but also to anticipate and address potential threats to agency resulting from device use. We also offer a brief suggestion regarding the practical utility of our map.

FOUR KEY DIMENSIONS OF ETHICAL INTEREST IN NEUROTECHNOLOGY

Responsibility

While neural devices offer novel prospects for motor rehabilitation, the intimate connection between devices and human intentional actions can obscure whether the machine or the end-user caused the action. In some cases, a BCI may misinterpret the user and produce an action they did not intend fully. In such cases, it remains an open question to whom the performance of the action should be attributed. Is it the end user, the device, both, or a third party (e.g., the manufacturer or the software developer) that should be held responsible for the resulting outcome? In the literature, this difficulty of assigning responsibility for complex human–machine interactions has been coined the "responsibility gap" (O'Brolchain and Gordijn 2014).

Recent accounts address this issue by viewing it through the lens of intentionality. On this view, the agent is responsible for the outcome of an action if she is intentionally performing that action (Fritz 2018). This general approach allows a form of reasoning for BCI-mediated actions that is similar to the form of reasoning which is typically used for describing everyday human actions, i.e. by identifying whether the user's intention matches the observed, goal-directed outcome (Holm and Voo 2011).

One feature relevant to responsibility ascription for intentional action is the agent's capacity to exercise control over their actions (Bublitz and Merkel 2009). For instance, for BCI use, a reasonable ascription of responsibility hinges on how integrated into the system's control-loop the user is. The more an agent is able to perform intentional actions within this control-loop, the more responsible he is for the outcome of that action. Conversely, if a malfunction results in the agent falling out of this control-loop, responsibility ascription is gradually lost (Kellmeyer et al. 2016). However, some authors disagree, arguing that BCIs introduce no new ethical problems relative to those already generated in conventional therapy (Clausen 2009), or that end users are ultimately responsible for unintended outcomes of implanted devices insofar as they deliberately take on BCI-related risks (Grübler 2011).

Despite these opposing views, understanding responsibility ascription through the connection between intentional action and control marks a contemporary argumentative trend. This theoretical background offers the opportunity to make more fine-grained distinctions between different levels of control in relation to various types of intentional actions (Steinert et al. 2019). Our framework understands the issue of responsibility ascription primarily as a problem of intentional control.

Privacy

Most broadly, privacy denotes the right for individuals to establish their boundaries and dynamics with others. This protection from unwanted access is a foundational piece to building and expressing an authentic sense of self (Igo 2018; Kupfer 1987). Through exerting control over who can access what parts of our lives, we establish boundaries that are

crucial for self-individuation and expression. Privacy needs reflect individuals' intimate sense of ownership (Braun et al. 2018) over their mental and physical states and boundaries for when, and by whom, their bodily and mental states can be accessed, their information taken, and their decisions encroached upon. However, as critical as privacy is, many guiding principles for neural research programs, such as the BRAIN Initiative and the European Human Brain Project, focus privacy considerations on the protection of neural *data* at the risk of oversimplifying the many ways neural devices affect users' sense of self and social dynamics (Greely et al. 2018; Salles et al. 2019). The currency of privacy extends beyond information—it crucially includes the ability for users to make decisions and regulate the access others have to their bodies (Klein and Rubel 2018).

Informational privacy within medical research often revolves around protecting patient confidentiality, or restricting access to personal information to authorized recipients (Allen 1997). This form of privacy is particularly important within the context of neural devices as it protects access to users' neural data. The data gathered by BCIs can be associated with conversations with loved ones, emotional responses to particular people or situations, or sensitive contexts and associated behavior (e.g., using the bathroom, sexual activity, etc.)—all of which could be considered extraordinarily private, depending on the user.

Beyond protecting information, privacy protections importantly also serve to ensure patients' autonomy in making medical decisions, often referred to as *decisional privacy* (Allen 2014, Beauchamp 2000). This kind of privacy is relevant to "the most intimate and personal choices a person may make" (Planned Parenthood v. Casey, 505 U.S. 833, 1992), including choices about abortion, birth control, etc. that require protection from governmental intervention or external influence. Participating in BCI research can require participants to engage in or forgo certain kinds of activities that fit in these identified realms of privacy. For example, in her autobiography, *On My Feet Again*, Jennifer French expresses frustration in having to choose between participating in a trial of an implanted neural device and having biological children (French 2012). The potential restriction of life choices that comes with taking on a neural device may be felt as a privacy burden. Most broadly, then, decisional privacy protects people's rights to fully understand what they are consenting to—as their use of these devices may affect their ability to make other personal decisions in the future and may allow others to have more of a say in their decision-making than they imagined.

Finally, privacy encapsulates people's right to set and negotiate boundaries for who can access their bodies under what conditions (Klein and Rubel 2018). Individuals with conditions like ALS may need frequent hands-on assistance from others to enable them to move, feed themselves, and go to the bathroom—routines that may require them to renegotiate what kind of contact they are comfortable with from those around them and adjust their *physical privacy* boundaries. This process may become all the more complicated if they receive an assistive neural device that may restore communicative and/or movement abilities. Further, bodily access concerns become all the more pressing if the neural implant introduces a portal to the brain that could be maliciously hacked.

These aspects of privacy overlap. For example, certain aspects of a person's body—from intimate tattoos to cancerous tumors to genital configurations—may be considered not for public knowledge, and a person may actively work to shape how others perceive them by controlling how those others access their body. They may work to control who has access to their body as they decide which kinds of information they wish different people to have. For instance, as a person's Parkinson's disease progresses, they may feel particularly vulnerable and exposed when a colleague eyes their trembling hand (Dubiel 2009). Similarly, someone who receives a neural implant may feel uncomfortable when they catch people staring at their cranial scar. Reclaiming their privacy may take the form of working to accept these reactions or testing different ways of concealing the betraying characteristic, reestablishing their privacy boundaries in the process.

Neural implants can complicate each of these aspects of privacy. They may increase the access others have to the user's intimate information, limit the kinds of decisions available to them, and complicate the ways others access and perceive their body. Our framework understands the issue of privacy as the degree to which one can negotiate access to their information, decision-making, and physical body.

Authenticity

Authenticity was an early topic of debate in the neuroethics literature and continues to be frequently discussed in relation to newer technologies like BCIs. Because the brain is often thought to be central to a person's identity, direct modification of it (regardless

of the means) may introduce changes to the self that trouble authenticity. Commonly, to be authentic means to act in accordance with one's "true" self (for an overview, see Vargas & Guignon 2020). However, many scholars (e.g., Walker and Mackenzie 2020) now repudiate the idea of any definitive, static attributes of self. That is, in some sense we cannot help but be ourselves, even through a wide variety of changes, as long as we maintain some authorial power over ourselves (Baylis 2013). Still, some ways of thinking, deciding, and acting seem out of place with one's self image, resulting in the individual feeling alienated.

Being authentic means not only feeling authentic, but also expressing one's self in ways that are fitting, whether through direct action, or being able to express one's thoughts and intentions openly and honestly to others. Furthermore, a person can feel authentic in the moment, but authenticity requires a capacity to understand and make sense of oneself over time. Finally, in our view, authenticity is not merely individual, from a person's own understanding, but also intersects with how others perceive the individual. In this sense, being authentic will typically require some uptake from close relational others who know the individual well and have a role in "holding' the person in their identity (Lindemann 2014).

Following the introduction of DBS for Parkinson's disease, reports emerged of patients sometimes not feeling like themselves and manifesting behaviors that they and others found not to match with the person they knew before surgery (Schüpbach et al. 2006). These troubling reports have led to a rich debate around different questions related to authenticity and DBS: does DBS cause changes to one's personal identity? (Gilbert, Viaña, and Ineichen 2018); how do end users experience feelings of inauthenticity? (Haan 2017); can authenticity be assessed? (Ahlin 2018); what is the role of authenticity in medical decision making? (Maslen, Pugh, and Savulescu 2015); and, do closed-loop systems introduce new issues related to authenticity (Kramer 2013a, 2013b)?

Authenticity is central to human agency. When we talk about an agent, we assume that the individual has some qualities that remain somewhat consistent over time. It is hard to feel agential if one's sense of self is significantly discontinuous. People care about feeling authentic and that requires having both some set of relatively stable core components and the ability to create change through goal-directed behavior that is coherent in relation to one's core components (Parens 2014). Our framework understands authenticity as the ability to reflect on one's past (Taylor 1992), envision one's future self (Degrazia 2000), and work to align one's actions with this vision while preserving continuity (Pugh, Maslen, and Savulescu 2017).

Trust

Neural engineers have created neurally-controlled assistive devices—like robotic limbs and prosthetic limbs. One problem with these devices may carry over from prosthetics, orthotics, and assistive robots: these devices may not feel as "natural" to use as an individual's own limbs would. That is, users may not feel ownership over or feel a sense of embodiment with these devices (Tbalvandany et al. 2019; Heersmink 2011). One potential way of overcoming this problem is to give users real-time, biomimetic feedback from their assistive devices through neurostimulation (Farina and Aszmann 2014; Kramer et al. 2019; Perruchoud et al. 2016). Systems like these could induce feelings of ownership through mechanisms like the rubber hand illusion (RHI)—a sensorimotor illusion whereby a person is given the temporary feeling that a life-like rubber limb is their own limb (Botvinick and Cohen 1998; Perez-Marcos, Slater, and Sanchez-Vives 2009; Johnson et al. 2013; Cronin et al. 2016; Collins et al. 2017).

Scholars in human-computer interaction have focused on features of reliability (e.g., low likelihood of malfunction) and other aspects of performance (e.g., consistency of achieving its goals) (Ahn and Jun 2015). In the BCI literature, the feel of the sensory feedback is an additional feature of human-device interaction that goes beyond mere reliability or performance. For instance, Collins et al. (2017) note that people receiving stimulation feedback during a BCI control task "reported that the sensations elicited by electrical stimulation […] felt 'unnatural' and unlike anything they had ever felt before" (166). Users may have trouble recognizing or accepting artificial sensations as their own. Further, biomimetic stimulation may conflict with other sensations that target populations sometimes experience: phantom limb sensations in amputees, residual sensations in people with incomplete spinal cord injuries (SCI), or even just input from other sensory modalities (e.g., vision). Some amputees report that their phantom limb sensations sometimes help and sometimes hinder their use of EEG-controlled prosthetic limbs (Bouffard et al. 2012). So, beyond the notions of reliability and performance there seems to be a question of whether users *trust* a device. There is an open question, then, if BCI users will be able to trust assistive neural

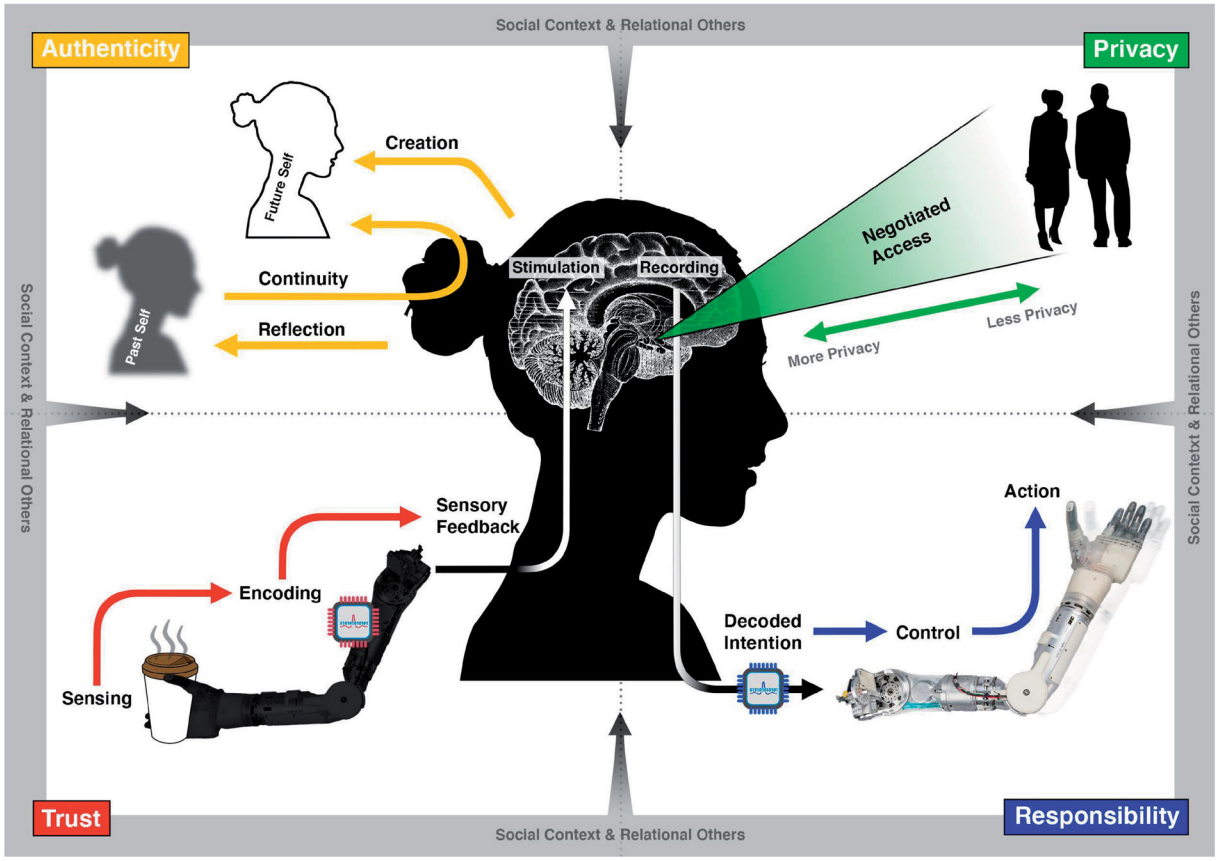


Figure 1. The Agency Map shows how different dimensions of agency are integrated in a single user of neurotechnology: (1) authenticity as reflecting upon the past self and creating a future self that has continuity over time; (2) privacy as a function of negotiating other people's access to private data, thoughts, body, and so on; (3) trust as the ability to discern and make use of sensory feedback received from or through the device; and (4) responsibility as the capacity to exercise control over an intentional action.

devices enough so that they become "transparent" (i.e., they disappear from the user's attention) (Ihde 1990; Heersmink 2011), even with sensory feedback. Richard Heersmink (2011) argues that rudimentary BCIs are not capable of this level of transparency, because a tool must reliably perform its function to become transparent.

But in what sense can a BCI system be trustworthy? In the philosophical literature, trust is often distinguished from mere reliance—where trust has a normative dimension that reliance does not (e.g., Hawley 2014). When we rely on others, we predict the other's future behavior and act in ways that presume they will act as they did before. When we trust others, however, we count on them to act in ways that stem from their good will toward us (Baier 1986; Jones 1996) or in ways that uphold our values (McLeod 2002; O'Neill 2002); breakdowns in trust result in feelings of betrayal. Assistive devices are not agents—at least prima facie—and so are incapable of either acting out of good will, or upholding our values. On this view, the devices are not candidates for trust. However, if biomimetic feedback makes BCI-enabled devices begin to feel embodied to the user, they may serve to help users learn to trust *themselves* as they act using the devices. Trust in neural devices then becomes as much about self-trust as about trust in the engineered system. However, such devices may also make it difficult for users to trust themselves or their ability to evaluate their own senses. That is, if the device sends phantom sensory feedback (e.g., of holding a coffee cup when no such object is being held), or if it fails to send feedback with appropriate intensity (e.g., to indicate that a coffee cup is scalding hot), the user may end up less able to gauge their surroundings and, as a result, how to act within them. The user may feel like they cannot trust their ability to evaluate their surroundings—or that they cannot trust themselves with their BCI. Our model, then, is meant to capture and elucidate how biomimetic sensory feedback provided by BCIs could make it more or less difficult for users to trust themselves.

AGENCY MAP BASIC FIGURE

In this section, we present our Agency Map (c.f. fig. 1) that brings together the previously discussed dimensions of agency in a comprehensive and integrated manner. This framework captures the user from a paradigmatic action-theoretical perspective by situating the agent in the center of four different quadrants that represent the agential dimensions we introduced in section 2, tying each to a particular power (i.e., an agentive competency). The agent actively navigates through a domain by exercising the respective agentive competency (i.e., maintaining *authenticity* through achieving *continuity* via the integration of previous, future, and current states of the self; defining individual *privacy* realms by *negotiating* with others over their access to the individual; *trusting* oneself to *interpret* sensory feedback regarding one's positioning, and feeling *responsible* for exercising intentional *control* over a goal-directed action). We begin by explaining the novel features of the agency map and then describe how it is based on an extended version of existing action-theoretical concepts.

With this Agency Map, we aim to draw upon contemporary findings in action theory from a relational perspective (a connection that we will flesh out in the second part) in order to tie together several of the most pressing neuroethical issues under the heading of agency. This unifying view allows us to depict these neuroethics issues together with the agential competencies we identified in section 2.

1. The dimension of responsibility is linked to the agential competency of exercising control. Here, we want to emphasize the ability of an agent to intentionally exercise control over a BCI-mediated movement. The agent may either find himself moving in ways that were unintended, or be uncertain about his level of control over successfully enacted intentions. The neuroethical question consists in asking to what extent an agent might be differently praise- or blameworthy if a perceived lack of control results in the inability to intentionally perform an action (as has been further explored by Schönau 2021).

2. The dimension of privacy is linked to the agential competency of negotiating access. Here, we want to highlight that the way a user understands privacy is always negotiated in a relational context with others. First, though, for an individual even to have a sense of self that is separate from others requires some degree of privacy, of recognizing individual thoughts and control over a body that others do not share (Kupfer 1987). Privacy is integral to developing agency, and a precondition to being able to function as an agent (Ramsay 2010, 293). Individuals participate in the process that determines who accesses their personal information, for which purposes, how they want to make decisions that guide their care and life after implantation, and what kinds of physical boundaries are most important to them. Through these negotiations, they actively shape their relationships with the world around them, enable chosen intimacies, and assert who they are in the process. Without adequate privacy an agent is incapable of making decisions/performing actions that are characteristic of agentive behavior.

3. The dimension of authenticity is tied to the agential competency of achieving continuity through integrating aspects of the self. Here, we want to emphasize the continuous agentive process of shaping an understanding of the self by comparing the past with the current self and the challenge of creating a possible future self that allows the individual to feel continuous with, even if changed from, these previously lived experiences. The neuroethical issue consists in alienation or inauthenticity that might be experienced if this comparison reveals inconsistencies.

4. The dimension of trust is tied to the agential competency of fostering self-trust. Here, we want to spotlight how an agent processes (artificially generated) sensory feedback. Usually, agents use their senses to ascertain if their actions were performed successfully or not. However, neural technology users will need to rely on sensory feedback that may feel unfamiliar, and therefore may need to develop other ways of validating their sense perceptions or rely on others' perceptions more than usual. Our aim is to ask: to what extent can a user evaluate the feedback they receive from their neural devices in a way to supports self-trust, and what role will this feedback play in their experience of being an intentional agent?

Overall, these four neuroethics dimensions help to comprise the agent, who is at the core of our agentive framework. The agent is the one who experiences phenomenologically what it is like to be in control, receive sensory information, share private information with others and shape an understanding of herself. By creating this map from this agent-centered perspective, we aim to show that these four dimensions are not independent, but rather inherently connected and highly influential on each other. A change in one of those domains – whether from natural causes like disease or through (artificially) induced alterations from neurotechnology use – might very well affect any other domain. In section 4, we will further elaborate

upon this connection with concrete examples that illustrate their mutual influence.

Our Agency Map draws a paradigmatic picture of agentive behavior that is built on contemporary findings in action theory, but it extends that view by incorporating relational influences that are usually outside the scope of action-theoretical endeavors.

Contemporary action theory acknowledges intentions as indispensable and irreducible mental states that prompt an agent to act (Brand 1984; Bratman 1987; Mele 1992; Pacherie 2008). Across these theories, the main aim is to understand how an agent generates, initiates, and performs an intentional action.[2] As such, the explanatory power of action theory has been fruitful in generating a novel perspective for the ethical issues of responsibility and culpability, by carefully utilizing formal language to distinguish acts of doing from acts of not-doing in the respect to moral and legal implications (among others, Chappell 2002; Clarke 2014; Wollard and Howard-Snyder 2016; Henne, Pinillos, and De Brigard 2017; Nelkin and Rickles 2017).

While we acknowledge the rich applicability of this approach, a major limitation is the narrow scope of action theory on the individual. Although a new trend under the heading of joined action (Pacherie 2012, 2013) and shared agency (Gilbert 1990; Bratman 2014) aims to understand how a group of individuals can generate and act upon mutual goals (e.g., taking a walk together), the relational influence of others on one's individual goals and capacity to act is still underexplored. We propose to broaden action theory by situating the intentional agent in a relational environment, highlighting individuals as relational beings (Doris 2015; Bierria 2014; Mackenzie and Stoljar 2000). This combination of contemporary action theory with a relational perspective allows us to tie the most commonly discussed neuroethical issues under the common heading of agency, with a novel understanding of their mutual connection, respective impact, and relational influence to and from other agents.

Viewing the agent as an active participant in the world allows us to expand the current action theoretical focus beyond the scope of individual action performance by incorporating several agential competencies that mutually shape the agentive experience and correspond to the depicted neuroethical issues. In what follows, we present several case examples that show how these different neuroethical dimensions are interconnected throughout the agency map.

ENACTING THE AGENCY MAP

Despite being organized under the broad heading of neurotechnology, devices that interface with the brain are varied in their design and function. Accordingly, the types of agential concerns they introduce will depend on the nature of the condition they are treating and the type of interaction with the brain. We have chosen three types of BCIs[3] that are actively being researched for clinical use. This focuses our discussion on the near-term future rather than more speculative technologies that may be developed. In each case, we highlight problems of agency for which users may turn to BCI devices for help and then consider problems of agency that may result in turn from subsequent use of such devices.

BCI for Spinal Cord Injury

A number of research projects aim to create assistive robotic devices for people with sensory and motor impairments. For example, the BRAINGATE device—an investigational system that drives a robotic arm using neural recordings—gave study participant Jan Scheuermann the ability to eat a chocolate bar on her own, despite the spinocerebellar degeneration that left her body paralyzed (Upson 2014). Earlier, we noted adjacent efforts aim to equip devices like these with sensory feedback (Cronin et al. 2016; Lee et al. 2018) for the sake of inducing feelings of ownership over them. Consider, then, the following fictional case based on these developments:

> Rachel is a 35 year old woman with an incomplete C5 spinal cord injury with loss of most motor function in her arms and legs and patchy sensory loss below her neck. She relies on a caregiver for activities of daily living, like bathing, dressing, and toileting. With residual function in her upper arms, she is able to navigate a power wheelchair, and she works part-time in software development from home using assistive inputs on her computer. She desires greater independence and misses the feel of physical contact with her friends and family. She decides to enroll in a pilot research study which, using electrodes implanted

[2]Elisabeth Pacherie's (2008: 189) depiction of the "intentional cascade" that introduces three types of intentions (i.e., future-directed intention, present-directed intention, motor intention) is a good example of what this threefold distinction could look like.

[3]We recognize debates in the literature regarding what counts as a BCI (Nijboer et al. 2013), potentially excluding open-loop DBS devices. Nevertheless, all three examples we discuss contain established BCI technologies (see Wolpaw and Wolpaw 2012) as part of the overall system.

> in her brain, will allow her to control a robotic limb mounted on her wheelchair by using only her "thoughts." Electrodes will be implanted in separate cortical areas to both record neural activity associated with her intentions, allowing her to control the robotic arm (e.g., to grasp items, like a cup of coffee), and stimulate the sensory cortex of her brain by channeling feedback from sensors on the arm, allowing the device to feel more real.

Rachel's interest in a BCI-robotic arm derives in part from constraints she experiences on her agency. Her injury interferes with her ability to initiate goal-directed movements using her body. Many of the actions that she could perform prior to her injury are now outside her autonomous control. Though she is agential in her decision making, for at least some purposes she relies on caregivers to follow her instructions, and accomplish her ends. She wishes to extend her agency with the use of an assistive device.

A BCI-controlled robotic limb may extend aspects of her agency, even as it raises other problems with agency. Imagine, for instance, that Rachel's artificial limb behaves in ways she doesn't expect. Perhaps the system is difficult to control, or isn't as responsive as Rachel would like it to be, or it malfunctions. In such cases, she may not be sure whether she or the device is in control (e.g., when she pounds the table with her robotic hand). And, hence, her sense of responsibility for what the robotic arm does may become muddled, or even disconcerting. Or consider the way that her agency might be affected by her interpretation of sensory feedback from her device. Imagine that sensory feedback from her robotic arm is at odds with other sensory feedback (e.g., vision, smell, hearing). What if Rachel uses the robotic arm to shake a visitor's hand and the visitor cries out in pain, as if from a crushing grip and yet the sensory feedback from the robotic arm registers only a gentle grasp? Should she trust her ability to use the robotic arm safely and interpret its feedback appropriately, or her ability to detect an instance of pain caused to another (or possibly a stranger's prank of feigning harm)? Issues of responsibility and self-trust may in turn lead to problems of authenticity (e.g., if Rachel's use of the robotic arm results in actions that feel out of character) and privacy (e.g., if records kept on-device can be audited to assign responsibility). The technology intended to enhance agency can help in one dimension (providing control over a prosthetic) even as it alters the others, creating new issues for the agent.

BCI for ALS

Implantable BCI devices are also being developed to enable verbal expression for people with impaired communicative capabilities due to conditions like amyotrophic lateral sclerosis (ALS). One such project enabled a partially locked-in woman with ALS to learn how to select letters on a computer screen by imaging moving her hand (Vansteensel et al. 2016). In addition, researchers are exploring ways that individuals can use "covert" or imagined speech to create classifiable neural signals (Martin et al. 2014) which could then produce synthetic speech (Anumanchipalli, Chartier, and Chang 2019). We consider the implications these next generation BCI devices have on users' agency in this hypothetical case:

> Armando is a 56-year-old who slowly lost his ability to verbally communicate due to amyotrophic lateral sclerosis (ALS). He used various assistive devices, such as a communication board, until he could no longer use them reliably due to loss of motor control, and was not able to use an eye-tracking device for communication due to longstanding stroke-related eye jerking (nystagmus). The inability to express his needs and share his thoughts with his wife has made him feel increasingly isolated. He enrolls in a study of an implantable BCI-based communication device, and eventually becomes skilled at communicating several dozen concepts or actions (e.g., "thirsty", "rest", "funny", "help", "TV").

A BCI-controlled communication device promises to extend agency by allowing a new mode of communication. For those who are limited or unable to communicate in other ways, like Armando, such a device provides a means of actualizing his intentions to express himself. The ability to communicate is central to enacting control over one's environment, building and nurturing relationships (and maintaining their boundaries), and sustaining a sense of self. Yet, a BCI-controlled communication device also raises some issues about agency.

One concern relates to privacy. Imagine, for instance, that Armando is paid a surprise visit by a friend, Sam, even though Armando is in no mood for company. Armando entertains the idea of choosing the concepts "friend" and "leave" but recognizes that this would hurt his friend's feelings, and changes his mind to communicate instead an invitation to "TV" and "sit". Yet, despite his change of mind, the device nevertheless greets Sam with "leave" and "friend". Even though Armando follows this with "error" and "TV" and "sit", a slight sting lingers for Sam and shame for Armando. One can see how such a device may make it more difficult to protect mental privacy

and shape social dynamics. This may be particularly true if devices foreground features of communication that are normally hidden (e.g., inner monologue or sub-vocalizations) or merely allow for the possibility of interrogation (e.g., what neural activity was present when you said X?). This issue can be potentially more severe because a BCI may pick up brain signals that have not yet been endorsed for execution by higher cognitive processing, but become appropriated by the agent as intended action after having been executed via the BCI (Krol et al. 2020). Thus the use of a BCI device can lead to a kind of *agential confabulation*.

BCI-based communication invites consideration of other kinds of agency issues as well. Communication is a form of action for which we hold people responsible. But is Armando responsible for hurting Sam's feelings? Should he feel remorse? Is the device to blame or is Armando to blame for failing to use it effectively? A BCI-communication device also raises questions about authenticity. The words one uses, what one chooses to say out loud (versus keep to oneself), and the tone one takes are all intimately connected to one's identity. In so far as a device becomes the medium for communication, it becomes a vehicle for being (or failing to be) authentic. Is Armando the kind of person who gives priority to speed of communication over accuracy? Or would he rather slow down and ensure greater accuracy? Is he the kind of person who would say hurtful words out loud? Using the device, can he trust himself not to have an inner dialogue that might get exported inadvertently?

Closed-Loop DBS for Major Depressive Disorder

BCI-based DBS or closed-loop DBS has been proposed as a possible treatment for psychiatric disorders such as depression, OCD, and addiction (Widge, Malone, and Dougherty 2018). Proponents argue that the ability to read neural signals directly from the brain and utilize machine learning algorithms to analyze the data will allow the system to provide dynamic stimulation that is automatically adjusted based on symptoms (Widge, Malone, and Dougherty 2018). Our anticipatory qualitative work indicates that closed-loop DBS for treatment of depression or OCD may introduce concerns around control over the device, the capability to express an authentic self, and relationship effects (Klein et al. 2016). The case below explores how a closed-loop DBS may impact user agency:

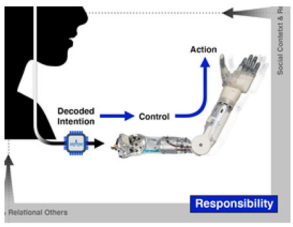
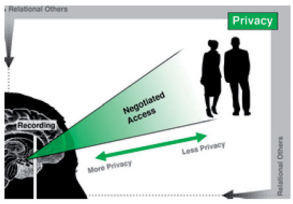
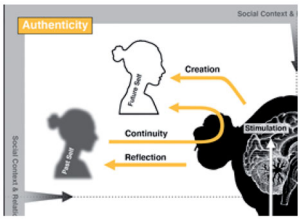
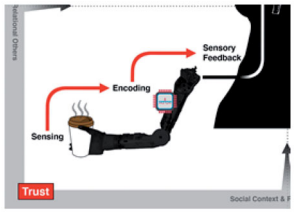
> Cora is a 28 year old woman with severe major depressive disorder (MDD). Over the last 5 years, she has tried multiple types of counseling, medications, and electroconvulsive therapy (ECT), all with minimal effect. She feels like depression is robbing her of who she used to be—outgoing, inquisitive, and carefree. Even dancing, which used to be a passion she shared with friends, has lost its appeal. She has become increasingly isolated. On the urging of friends and family, she enrolls in a trial of closed-loop DBS for depression. This device will both record her brain activity and stimulate areas of the brain thought to be important for symptoms of depression. She, and they, are optimistic that the device will allow her to get back to the Cora they remember.

BCI-based DBS has appeal as therapy in part owing to its potential to address problems of agency. Psychiatric conditions often interfere with the ability of people to act as they want or think they ought to act. By relieving symptoms of depression, BCI-DBS promises to restore the agency that depression (or other mental illness) has dampened. For some, the hope is to return to a more authentic self.

The BCI-based DBS opens up concerns about agency as well. First, in so far as stimulation triggers new beliefs or intentions, ones that are different from the pre-depression self, we may simply be replacing one feeling of inauthenticity with a different one. Imagine, for instance, that Cora is faced with a situation which would normally engender feelings of sadness (e.g., a funeral), but because her device detects a certain neural activity pattern it provides extra stimulation so that she does not feel sad. If she fails to feel sad during the funeral, will she attribute this to the working of the device or to her feelings (or lack of them) for the deceased? Will this experience cause her to reexamine the history of her relationship with this person?

Authenticity is not the only relevant concern about agency here. Such devices have implications for responsibility in so far as they affect one's ability to be motivated to initiate goal-directed action, to get up and do what one feels one should be doing. A closed-loop device, which aims to automate much of the traditional tinkering required in psychiatric treatment or open-loop DBS, may take the user "out of the loop" and lead to people feeling as if they are no longer fully in control of their own actions (Goering et al. 2017). Issues of trust are implicated here in so far as devices may change the way people experience feedback from the outside world. Imagine that Cora becomes too reliant on her own device, feeling that because the DBS system is adaptive and can "automatically" treat her MDD, she no longer needs to take an active role in dealing with her condition.

Table 2. Qualitative Agentive Competency Tool (Q-ACT).

Dimension	Agential Competency	Agency Inquiry Prompts
	Exercising Control	1. Do you feel that the (device mediated) movements you perform are under your intentional control? 2. Do you sometimes feel that you share control with the AI/device? 3. To what extent do you feel responsible for those movements?
	Negotiating Access	1. Would you say that information about your behaviors, thoughts, or attitudes is becoming more or less private? Do you feel that you are active in determining who has access? 2. Do you feel that having the device alters your privacy with respect to decision making in your life? 3. Do you feel that others know more about the state of your physical body due to the presence of the device? Are you comfortable with that access?
	Integrating Self	1. Has living with the device changed your perception of yourself? Have others noticed changes in you? 2. Do you feel comfortable with these changes? 3. Do you feel able to adequately shape the person you are now and want to be?
	Fostering Self-Trust	1. Do you think the device works reliably and records/stimulates correctly? 2. Does the feedback from the device feel natural/trustworthy? 3. Do you trust yourself when you are using the device? 4. Do you think that others trust you when you are using the device?

Each dimension (responsibility, privacy, authenticity, trust) is aligned with a competency agents make use of and a series of prompts researchers and clinicians can use to ascertain how neurotechnologies impact these competencies.

Listening to End-Users: Toward a Qualitative Assessment Tool

How can our agency map help researchers and clinicians to assess these agential issues more effectively as BCI technology continues to develop? We suggest that the kinds of questions presented in the illustrative examples - related to intertwined issues of responsibility, privacy, authenticity, and trust - can form the basis of a qualitative assessment tool for researchers and clinicians. Qualitative studies—in particular, semi-structured interviews paired with thematic analyses—to ascertain what impact neurotechnologies have on users have grown all the more common (Haan 2017; Kögel, Jox, and Friedrich 2020). While qualitative approaches are ideal for capturing user experiences, many of these interview projects do not focus on agency specifically (Gilbert et al. 2019), and may fail to appreciate how these dimensions of agency intersect in the user's experience. To encourage more studies on the agency of neurotechnology users, we offer the outlines of a Qualitative Agentive Competency Tool (Q-ACT) that identifies some of the competencies agents employ in each dimension, and suggests a series of key questions researchers or clinicians could ask to more fully assess the impact of neurotechnology on an individual's agency (Table 2).

A Q-ACT tool will need to be tailored to the particular features of neurotechnologies and populations targeted by these technologies. Reflecting on the cases discussed above, one can imagine asking: Rachel, did you feel responsible for the movements of the robotic limb (e.g., pounding the table) and able to trust the feedback from the device [exercising control; fostering self-trust]? Armando, did you feel your attitude toward your friend was made less private than you wanted [negotiating access]? Cora, has your device affected your ability to make sense of yourself over time and across important life events (e.g., death of a loved one) [integrating self]? Moreover, the Q-ACT

can provide a way of structuring these discussions, by bringing together questions across dimensions of agency. For instance: Armando, beyond concerns about privacy, did you feel responsible for hurting your friend? Did it cause you to question that you know (or trust) your true feelings toward your friend? Does the device allow you to feel authentic with your friends? Do other people react differently to you, now that you have the device? Rather than functioning as a checklist of agency-related concerns, these questions provide a starting point for explorations of agency and the too often neglected interconnections between these dimensions. Work will need to be done to tailor the Q-ACT to particular technologies (e.g., DBS vs BCI) and populations (e.g., ALS vs SCI vs depression). We suggest that this work is worth undertaking in order to better understand how neural devices that appear to meet their immediate aims may nonetheless have significant agential effects on users.

CONCLUSION

In sum, we have identified four dimensions of ethical inquiry—responsibility, privacy, authenticity, and trust—as dimensions of agency that may be impacted by the use of neurotechnologies. In order to describe the role of each dimension, and how these dimensions are interrelated, we presented an Agency Map: a depiction of how these dimensions of agency are integrated into a person's experience as they use neurotechnology. The Agency Map is meant to be a guide for recognizing how neurotechnologies may change users' experience of agency—either through enhancing it or addressing a problem of agency in one dimension only to create new agential problems in others.

To further illustrate the agential concerns neurotechnology users may face, and to demonstrate how our agency map can help identify or track those concerns in the context of how people use specific technologies, we presented three hypothetical case studies built on existing technologies. Each case study raises concerns that might seem to fall squarely within one region of the map—where Armando wants to negotiate access to his private life, Rachel may just want more explicit control of her artificial hand. Each case study, however, demonstrates how the features of agency overlap—Cora may not know if her actions are authentically her own, and she may not fully trust herself as a result. A strength of our agency map is in how it considers these dimensions fit together as interwoven parts of one experience, instead of as distinct topics to discuss separately. Our map can also be used as the basis for a tool (suggested by our Q-ACT) to assist clinicians, researchers, and caretakers in assessing changes to agency after treatment with a neural device. We believe this approach, of considering the possible interactions between the competencies and dimensions of agency, will help the field anticipate issues with neurotechnologies we might otherwise overlook.

FUNDING

This work was supported by Nation Institutes of Health (MH117800-01) and National Science Foundation (EEC-1028725).

REFERENCES

Ahlin, J. 2018. The Impossibility of Reliably Determining the Authenticity of Desires: Implications for Informed Consent. *Medicine, Health Care, and Philosophy* 21 (1): 43–50. doi:10.1007/s11019-017-9783-0.

Ahn, M., and S. C. Jun. 2015. Performance Variation in Motor Imagery brain-computer interface: a brief review. *Journal of Neuroscience Methods* 243:103–10. doi:10.1016/j.jneumeth.2015.01.033.

Widge, A. S., D. A. Malone, and D. D. Dougherty. 2018. Closing the Loop on Deep Brain Stimulation for Treatment-Resistant Depression. *FOCUS* 16 (3):305–13. doi:10.1176/appi.focus.16302.

Allen, A. 1997. Genetic Privacy: Emerging Concepts and Values. In *Genetic Secrets: Protecting Privacy and Confidentiality in the Genetic Era*, edited by Mark Rothstein, 31–59. New Haven: Yale University Press.

Allen, A. 2014. Privacy in Health Care. In *Encyclopedia of Bioethics*, edited by Bruce Jennings, 4th ed. New York: MacMillan Reference Books.

Anumanchipalli, G. K., J. Chartier, and E. F. Chang. 2019. Speech Synthesis from Neural Decoding of Spoken Sentences. *Nature* 568 (7753):493–8. doi:10.1038/s41586-019-1119-1.

Baier, A. 1986. Trust and Antitrust. *Ethics* 96 (2):231–60. doi:10.1086/292745.

Baylis, F. 2013. "I Am Who I Am": On the Perceived Threats to Personal Identity from Deep Brain Stimulation. *Neuroethics* 6 (3):513–26. doi:10.1007/s12152-011-9137-1.

Beauchamp, T. L. 2000. The Right to Privacy and the Right to Die. *Social Philosophy and Policy* 17 (2):276–92. doi:10.1017/S0265052500002193.

Bierria, A. 2014. Missing in Action: Violence, Power, and Discerning Agency. *Hypatia* 29 (1):129–45. doi:10.1111/hypa.12074.

Botvinick, M., and J. Cohen. 1998. Rubber Hands 'feel' touch that eyes see. *Nature* 391 (6669):756. doi:10.1038/35784.

Bouffard, J., C. Vincent, É. Boulianne, S. Lajoie, and C. Mercier. 2012. Interactions Between the Phantom Limb Sensations, Prosthesis Use, and Rehabilitation as Seen by

Amputees and Health Professionals. *JPO Journal of Prosthetics and Orthotics* 24 (1):25–33. doi:10.1097/JPO.0b013e318240d171.

Brand, M. 1984. *Intending and Acting: Toward a Naturalized Action Theory*. Cambridge: MIT Press.

Bratman, M. 1987. *Intention, Plans, and Practical Reason*. Cambridge: Harvard University Press.

Bratman, M. 2014. *Shared Agency: A Planning Theory of Acting Together*. New York, NY: Oxford University Press.

Braun, N., S. Debener, N. Spychala, E. Bongartz, P. Sörös, H. H. O. Müller, and A. Philipsen. 2018. The Senses of Agency and Ownership: A Review. *Front Psychol* 9: 535(April 16). doi:10.3389/fpsyg.2018.00535.

Bublitz, J. C., and R. Merkel. 2009. Autonomy and Authenticity of Enhanced Personality Traits. *Bioethics* 23 (6):360–374. doi:10.1111/j.1467-8519.2009.01725.x.

Buss, S., and A. Westlund. 2018. Personal Autonomy. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2018. Stanford, CA: Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2018/entries/personal-autonomy/.

Chappell, T. 2002. Two Distinctions That Do Make a Difference: The Action/Omission Distinction and the Principle of Double Effect. *Philosophy* 77 (2):211–233. doi:10.1017/S0031819102000256.

Clarke, R. 2014. *Omissions: Agency, Metaphysics, and Responsibility*. Oxford, New York: Oxford University Press.

Clausen, J. 2009. Man, Machine and in Between. *Nature* 457 (7233):1080–1081. doi:10.1038/4571080a.

Collins, K. L., A. Guterstam, J. Cronin, J. D. Olson, H. Henrik Ehrsson, and J. G. Ojemann. 2017. Ownership of an Artificial Limb Induced by Electrical Brain Stimulation. *Proceedings of the National Academy of Sciences of the United States of America* 114 (1):166–71. doi:10.1073/pnas.1616305114.

Cronin, J. A., J. Wu, K. L. Collins, D. Sarma, R. P. N. Rao, J. G. Ojemann, and J. D. Olson. 2016. Task-Specific Somatosensory Feedback via Cortical Stimulation in Humans. *IEEE Trans Haptics* 9 (4):515–22. doi:10.1109/TOH.2016.2591952.

Degrazia, D. 2000. Prozac, Enhancement, and Self-Creation. *The Hastings Center Report* 30 (2):34–40. doi:10.2307/3528313.

Doris, J. M. 2015. *Talking to Our Selves: Reflection, Ignorance, and Agency*. 1. ed. Oxford: Oxford Univ. Press.

Dubiel, H. 2009. *Deep Within the Brain: Living with Parkinson's Disease*. New York: Europa Editions UK.

Farina, D., and O. Aszmann. 2014. Bionic Limbs: Clinical Reality and Academic Promises. *Science Translational Medicine* 6 (257):257ps12. doi:10.1126/scitranslmed.3010453.

French, J. 2012. *On My Feet Again: My Journey Out of the Wheelchair Using Neurotechnology*. San Francisco, CA: Neurotech Press.

Fritz, K. G. 2018. Moral Responsibility, Voluntary Control, and Intentional Action. *Philosophia* 46 (4):831–55. doi:10.1007/s11406-018-9968-7.

Gilbert, F. 2015. A Threat to Autonomy? The Intrusion of Predictive Brain Implants. *AJOB Neurosci* 6 (4):4–11. doi:10.1080/21507740.2015.1076087.

Gilbert, M. 1990. Walking Together: A Paradigmatic Social Phenomenon. *Midwest Studies in Philosophy* 15 (1):1–14. doi:10.1111/j.1475-4975.1990.tb00202.x.

Gilbert, F., T. Brown, I. Dasgupta, H. Martens, E. Klein, and S. Goering. 2019. An Instrument to Capture the Phenomenology of Implantable Brain Device Use. *Neuroethics*. https://link.springer.com/article/10.1007%2Fs12152-019-09422-7#citeas.

Gilbert, F., J. N. M. Viaña, and C. Ineichen. 2018. Deflating the 'DBS Causes Personality Changes' Bubble. *Neuroethics*. https://link.springer.com/article/10.1007/s12152-018-9373-8#citeas. doi:10.1007/s12152-018-9373-8.

Glannon, W. 2009. Stimulating Brains, Altering Minds. *Journal of Medical Ethics* 35 (5):289–292. doi:10.1136/jme.2008.027789.

Goering, S., E. Klein, D. D. Dougherty, and A. S. Widge. 2017. Staying in the Loop: Relational Agency and Identity in Next-Generation DBS for Psychiatry. *AJOB Neuroscience* 8 (2):59–70. doi:10.1080/21507740.2017.1320320.

Greely, H. T., C. Grady, K. M. Ramos, W. Chiong, J. Eberwine, N. A. Farahany, L. S. M. Johnson, B. T. Hyman, S. E. Hyman, K. S. Rommelfanger, et al. 2018. Neuroethics Guiding Principles for the NIH BRAIN Initiative. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 38 (50):10586–88. doi:10.1523/JNEUROSCI.2077-18.2018.

Grübler, G. 2011. Beyond the Responsibility Gap. Discussion Note on Responsibility and Liability in the Use of Brain-Computer Interfaces. *AI & SOCIETY* 26 (4):377–82. doi:10.1007/s00146-011-0321-y.

Haan, S. D. 2017. Becoming More Oneself? Changes in Personality Following DBS Treatment for Psychiatric Disorders: Experiences of OCD Patients and General Considerations. *PLOS One*. 12 (4):e0175748. Edited by Christian Schmahl. doi:10.1371/journal.pone.0175748..

Haan, S. d., E. Rietveld, M. Stokhof, and D. Denys. 2015. Effects of Deep Brain Stimulation on the Lived Experience of Obsessive-Compulsive Disorder Patients: In-Depth Interviews with 18 Patients. *Plos ONE* 10 (8): e0135524. doi:10.1371/journal.pone.0135524.

Haselager, P. 2013. Did I Do That? Brain–Computer Interfacing and the Sense of Agency. *Minds and Machines* 23 (3):405–18. doi:10.1007/s11023-012-9298-7.

Hawley, K. 2014. Trust, Distrust and Commitment. *Noûs* 48 (1):1–20. doi:10.1111/nous.12000.

Heersmink, R. 2011. "Epistemological and Phenomenological Issues in the Use of Brain-Computer Interfaces." In *Proceedings IACAP 2011: First International Conference of IACAP: The Computational Turn: Past, Presents, Futures?, 4 – 6 July, 2011, Aarhus University*, 87–90. MV-Wissenschaft. https://researchers.mq.edu.au/en/publications/epistemological-and-phenomenological-issues-in-the-use-of-brain-c.

Henne, P., Á. Pinillos, and F. De Brigard. 2017. Cause by Omission and Norm: Not Watering Plants. *Australasian Journal of Philosophy* 95 (2):270–83. doi:10.1080/00048402.2016.1182567.

Holm, S., and T. C. Voo. 2011. Brain-Machine Interfaces and Personal Responsibility for Action-Maybe Not as Complicated after All. *Studies in Ethics, Law, and Technology* 4 (3). doi:10.2202/1941-6008.1153.

Igo, S. E. 2018. *The Known Citizen: A History of Privacy in Modern America*. Cambridge, MA: Harvard University Press.

Ihde, D. 1990. *Technology and the Lifeworld: From Garden to Earth. Indiana Series in the Philosophy of Technology*. Bloomington: Indiana University Press.

Johnson, L. A., J. D. Wander, D. Sarma, D. K. Su, E. E. Fetz, and J. G. Ojemann. 2013. Direct Electrical Stimulation of the Somatosensory Cortex in Humans Using Electrocorticography Electrodes: A Qualitative and Quantitative Report. *Journal of Neural Engineering* 10 (3): 036021. doi:10.1088/1741-2560/10/3/036021.

Jones, K. 1996. Trust as an Affective Attitude. *Ethics* 107 (1):4–25. doi:10.1086/233694.

Kellmeyer, P., T. Cochrane, O. Müller, C. Mitchell, T. Ball, J. J. Fins, and N. Biller-Andorno. 2016. The Effects of Closed-Loop Medical Devices on the Autonomy and Accountability of Persons and Systems. *Camb Q Healthc Ethics* 25 (4):623–33. doi:10.1017/S0963180116000359.

Klein, E., S. Goering, J. Gagne, C. V. Shea, R. Franklin, S. Zorowitz, D. D. Dougherty, and A. S. Widge. 2016. Brain-Computer Interface-Based Control of Closed-Loop Brain Stimulation: Attitudes and Ethical Considerations. *Brain-Computer Interfaces* 3 (3):140–48. doi:10.1080/2326263X.2016.1207497.

Klein, E., and A. Rubel. 2018. *Privacy and Ethics in Brain–Computer Interface Research*. In: Nam, C. S., A, Nijholt, and F. Lotte, eds. *Brain–computer interfaces handbook*, 653–668. Boca Raton: Taylor & Francis, CRC Press.

Kögel, J., R. J. Jox, and O. Friedrich. 2020. What Is It like to Use a BCI? - insights from an interview study with brain-computer interface users. *BMC Medical Ethics* 21 (1):2. doi:10.1186/s12910-019-0442-2.

Kraemer, F. (2011). Authenticity anyone? The enhancement of emotions via neuro-psychopharmacology. *Neuroethics* 4 (1):51–64.

Kramer, F. 2013a. Authenticity or Autonomy? When Deep Brain Stimulation Causes a Dilemma. *Journal of Medical Ethics* 39 (12):757–60. doi:10.1136/medethics-2011-100427.

Kramer, F. 2013b. Me, Myself and My Brain Implant: Deep Brain Stimulation Raises Questions of Personal Authenticity and Alienation. *Neuroethics* 6 (3):483–97. doi:10.1007/s12152-011-9115-7.

Kramer, D. R., S. Kellis, M. Barbaro, M. A. Salas, G. Nune, C. Y. Liu, R. A. Andersen, and B. Lee. 2019. Technical Considerations for Generating Somatosensation via Cortical Stimulation in a Closed-Loop Sensory/Motor Brain-Computer Interface System in Humans. *Journal of Clinical Neuroscience* 63:116–21. doi:10.1016/j.jocn.2019.01.027.

Krol, L. R., P. Haselager, and T. O. Zander. 2020. Cognitive and Affective Probing: A Tutorial and Review of Active Learning for Neuroadaptive Technology. *Journal of Neural Engineering* 17 (1):012001. doi:10.1088/1741-2552/ab5bb5.

Kupfer, J. 1987. Privacy, Autonomy, and Self-Concept. *American Philosophical Quarterly* 24 (1):81–89.

Lee, B., D. Kramer, M. Armenta Salas, S. Kellis, D. Brown, T. Dobreva, C. Klaes, C. Heck, C. Liu, and R. A. Andersen. 2018. Engineering Artificial Somatosensation Through Cortical Stimulation in Humans. *Front Syst Neurosci* 12:24. doi:10.3389/fnsys.2018.00024.

Lindemann, H. 2014. *Holding and Letting Go: The Social Practice of Personal Identities. Holding and Letting Go*. Oxford University Press.

Mackenzie, Catriona, and Natalie Stoljar, eds. 2000. *Relational Autonomy: Feminist Perspectives on Automony, Agency, and the Social Self*. New York: Oxford University Press.

Martin, S., P. Brunner, C. Holdgraf, H.-J. Heinze, N. E. Crone, J. Rieger, G. Schalk, R. T. Knight, and B. N. Pasley. 2014. Decoding Spectrotemporal Features of Overt and Covert Speech from the Human Cortex. *Frontiers in Neuroengineering* 7:14. doi:10.3389/fneng.2014.00014.

Maslen, H., J. Pugh, and J. Savulescu. 2015. The Ethics of Deep Brain Stimulation for the Treatment of Anorexia Nervosa. *Neuroethics* 8 (3):215–30. doi:10.1007/s12152-015-9240-9.

McLeod, C. 2002. *Self-Trust and Reproductive Autonomy*. Cambridge, MA: MIT Press.

Mele, A. R. 1992. *Springs of Action: Understanding Intentional Behavior*. Oxford, UK: Oxford University Press.

Nelkin, Dana, and Samuel Rickles, eds. 2017. *The Ethics and Law of Omissions*. Oxford, UK: Oxford University Press.

Nijboer, F., J. Clausen, B. Z. Allison, and P. Haselager. 2013. The Asilomar Survey: Stakeholders' Opinions on Ethical Issues Related to Brain-Computer Interfacing. *Neuroethics* 6 (3):541–78. doi:10.1007/s12152-011-9132-6.

O'Brolchain, F., 2014. and, and B. Gordijn. Brain–Computer Interfaces and User Responsibility. In *Brain-Computer-Interfaces in Their Ethical, Social and Cultural Contexts*, edited by Gerd Grübler and Elisabeth Hildt, 12:163–182. Dordrecht: Springer Netherlands.

O'Neill, O. 2002. "AUTONOMY AND TRUST IN BIOETHICS." 241300. 2002. https://repository.library.georgetown.edu/handle/10822/547087.

Pacherie, E. 2008. The Phenomenology of Action: A Conceptual Framework. *Cognition "Cognition"* 107 (1): 179–217. doi:10.1016/j.cognition.2007.09.003.

Pacherie, E. 2012. The Phenomenology of Joint Action: Self Agency vs. Joint-Agency. In *Joint Attention: New Developments*, edited by Axel Seemann. Cambridge: MIT Press.

Pacherie, E. 2013. Intentional Joint Agency: Shared Intention Lite. *Synthese* 190 (10):1817–39. doi:10.1007/s11229-013-0263-7.

Parens, E. 2014. *Shaping Our Selves: On Technology, Flourishing, and a Habit of Thinking. Shaping Our Selves*. Oxford University Press.

Perez-Marcos, D., M. Slater, and M. V. Sanchez-Vives. 2009. Inducing a Virtual Hand Ownership Illusion through a Brain-Computer interface. *Neuroreport* 20 (6): 589–94. doi:10.1097/wnr.0b013e32832a0a2a.

Perruchoud, D., I. Pisotta, S. Carda, S. MM. Murray, and S. Ionta. 2016. Biomimetic rehabilitation engineering: The importance of somatosensory feedback for brain–machine interfaces. *Journal of Neural Engineering* 13, 041001.

Pugh, J., H. Maslen, and J. Savulescu. 2017. Deep Brain Stimulation, Authenticity and Value. *Cambridge*

Quarterly of Healthcare Ethics: CQ: The International Journal of Healthcare Ethics Committees 26 (4):640–57. doi:10.1017/S0963180117000147.

Salles, A., J. G. Bjaalie, K. Evers, M. Farisco, B. T. Fothergill, M. Guerrero, H. Maslen, J. Muller, T. Prescott, B. C. Stahl, et al. 2019. The Human Brain Project: Responsible Brain Research for the Benefit of Society. *Neuron* 101 (3):380–84. doi:10.1016/j.neuron.2019.01.005.

Schönau, A. 2021. The spectrum of responsibility ascription for end users of neurotechnologies. Neuroethics. https://doi.org/10.1007/s12152-021-09460-0

Schüpbach, M., M. Gargiulo, M. L. Welter, L. Mallet, C. Behar, J. L. Houeto, D. Maltete, V. Mesnage, and Y. Agid. 2006. Neurosurgery in Parkinson disease: a distressed mind in a repaired body? *Neurology* 66 (12): 1811–1816. doi:10.1212/01.wnl.0000234880.51322.16.

Sherwin, S. 1998. *The Politics of Women's Health: Exploring Agency and Autonomy*. Philadelphia, PA: Temple University Press.

Steinert, S.,. C. Bublitz, R. Jox, and O. Friedrich. 2019. Doing Things with Thoughts: Brain-Computer Interfaces and Disembodied Agency. *Philosophy & Technology* 32 (3):457–82. doi:10.1007/s13347-018-0308-4.

Taylor, C. 1992. *The Ethics of Authenticity*. Cambridge, MA: Harvard University Press.

Tbalvandany, S. S., B. S. Harhangi, A. W. Prins, and M. H. N. Schermer. 2019. Embodiment in Neuro-Engineering Endeavors: Phenomenological Considerations and Practical Implications. *Neuroethics* 12 (3):231–42. doi:10.1007/s12152-018-9383-6.

Upson, S. 2014. "What Is It Like to Control a Robotic Arm with a Brain Implant?" *Scientific American*. http://www.scientificamerican.com/article/what-is-it-like-to-control-a-robotic-arm-with-a-brain-implant/.

Vansteensel, M. J., E. G. M. Pels, M. G. Bleichner, M. P. Branco, T. Denison, Z. V. Freudenburg, P. Gosselaar, S. Leinders, T. H. Ottens, M. A. Van Den Boom, et al. 2016. Fully Implanted Brain-Computer Interface in a Locked-In Patient with ALS . *The New England Journal of Medicine* 375 (21):2060–66. doi:10.1056/NEJMoa1608085.

Varga, S., and C. Guignon. 2020. "Authenticity." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2020. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2020/entries/authenticity/.

Walker, M. J., and C. Mackenzie. 2020. Neurotechnologies, Relational Autonomy, and Authenticity. *IJFAB: International Journal of Feminist Approaches to Bioethics* 13 (1):98–119. doi:10.3138/ijfab.13.1.06.

Wollard, F., and F. Howard-Snyder. 2016. Doing vs. Allowing Harm. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Stanford, CA: Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/entries/doing-allowing/.

Wolpaw, J. R., and E. W. Wolpaw. 2012. *Brain-Computer Interfaces: Principles and Practice*. Oxford; New York: Oxford University Press.