ENGINEERING THE BRAIN

ETHICAL ISSUES AND THE INTRODUCTION OF NEURAL DEVICES

BY ERAN KLEIN, TIM BROWN, MATTHEW SAMPLE, ANJALI R. TRUITT, AND SARA GOERING

Neural devices now under development stand to interact with and alter the human brain in ways that may challenge standard notions of identity, normality, authority, responsibility, privacy, and justice.

Neural engineering technologies such as implanted deep brain stimulators and brain-computer interfaces represent exciting and potentially transformative tools for improving human health and well-being. Yet their current use and future prospects raise a variety of ethical and philosophical concerns.[1] Devices that alter brain function invite us to think deeply about a range of ethical concerns—identity, normality, authority, responsibility, privacy, and justice. If a device is stimulating my brain while I decide upon an action, am I still the author of the action? Should I be held accountable for every action in which a device is operative? Does a device make the interiority of my experience accessible to others? Will the device change the way I think of myself and others think of me? Such fundamental questions arise even when a device is designed for only a relatively circumscribed purpose, such as restoring functioning via a smart prosthetic.

We are part of a National Science Foundation–funded engineering research center tasked with investigating philosophical and social implications of neural engineering research and technologies.[2] Neural devices already in clinical use, such as deep brain stimulators for Parkinson's disease or essential tremor, have spurred healthy debate about such implications.[3] Devices currently under development—such

as the BrainGate System of implanted brain sensors coupled to robotics in persons with paralysis,[4] exoskeletons for augmented movement,[5] transcranial do-it-yourself stimulators (tDCS),[6] closed-loop brain stimulating systems,[7] or even brain-to-brain interfacing[8]—promise to extend and deepen these debates. At our center, brain-computer interfaces (BCIs) are the principal focus of work.[9] Even acknowledging that the clinical translation of neural devices and seamless integration by end users may still largely reside in the future, [10] the potential of these devices calls for careful early analysis. The launching of the BRAIN (Brain Research through Advancing Innovative Neurotechnologies) Initiative in April 2013 provides further impetus for this work.[11]

In our work alongside neural engineers, we have come to view the work of engineering the brain as incredibly complex, not simply because of

best explored together. Even careful examination of BCIs and privacy, for instance, leaves out something of critical import if it fails to attend to ways that outside access to our thoughts or intentions—even where secured by firewalls and encryption—may deeply affect our ideas about identity or responsibility.

Consider two fictional case studies, both based on developments in neural engineering: Joan is a thirty-year-old mechanical engineer and combat veteran who sustained an injury from an improvised explosive device. On her right arm, she has an above-elbow amputation, and her left arm has significant nerve damage that causes pain. After a year of aggressive rehabilitation at a military hospital, she volunteers to work with a research team developing a state-of-the-art robotic prosthetic controlled by a system that involves a brain-computer interface. A powered array of sensing

psychiatric hospitalizations. Antidepressant medications, counseling, and electroconvulsive therapy have been unsuccessful. Although he is well-liked by his coworkers and supervisors, his debilitating depressive episodes have made his deadline-oriented work as a technical writer difficult to accomplish at times. These complications threaten his job security. He volunteers for a study investigating an experimental BCI, a closed-loop deep brain stimulation (DBS) device for refractory depression as "my last hope." A set of sensing and stimulating electrodes are implanted in his brain and connected to a unit implanted in his upper chest. The unit is responsible for interpreting signals from the sensory electrodes, determining when treatment is needed, and applying a current through the stimulation electrodes. This sensor-stimulator loop detects changes in his mood and

Neural engineering devices implanted in brains and designed to interface with nervous tissue in closed-loop systems may change our notions of both social identity and body schema.

the technical feats it requires but also because of the varied ethical domains upon which the work touches. The functioning of the brain is intimately connected to an individual's and a culture's understandings of identity, human responsibility, privacy, authority, justice, and normality. In this paper, drawing on and extending work in neuroethics focused on deep brain stimulation, we explore how neural engineered devices, particularly brain-computer interface devices, challenge or may soon challenge our understanding of these six domains. We have structured this paper to focus on each of these domains individually, but as will quickly become apparent, we recognize that doing so is part artifice; these domains, particularly in their relationship to human agency, are often entangled and

electrodes and a chipset implanted in her brain records and sends signals to a robotic arm prosthesis through a wireless connection. Joan works with the engineers to customize the appearance of the prosthetic to fit her needs and personality: "strong, durable, and nothing frilly, but with just enough of a soft exterior that I can hold my newborn daughter." When a new model of prosthetic is developed for improved functionality that requires extensive training or sharing of movement data from her existing prosthetic, she declines the upgrade, preferring to "keep that part of me just as it is, just as my daughter knows me."

John, a forty-five-year old man, has struggled with depression since his late teens, including several suicide attempts and inpatient

adjusts stimulation to achieve an appropriate set point determined during consultations with his physician. After implantation, John's depressive symptoms dramatically decrease, a fact that he frequently shares with his work colleagues, noting, "I am finally back to being me."

Identity

Our identities are complicated by our relationship with technologies. Many of us rely on tools such as smart phones, laptops, and GPS devices. Indeed, we sometimes incorporate these tools into our self-understandings—"He's a Mac guy," or, "She couldn't live without her smart phone." When our tools not only aid us but also directly replace parts or functions of our bodies, as may be

the case for Joan and John, the effects of technology on identity are potentially even more significant. The experience of coming to identify with a prosthetic is rich and complex.[12] People who are blind sometimes think of their canes as part of their perceptual systems;[13] individuals with communication disorders can identify with their computer-synthesized voices.[14] When a tool functions so well that it becomes an integral part in our lives, we might say it has become part of us—it is no longer a mere tool.[15] Are neural devices likely to be taken up into our identities, and if so, to what benefit, or at what cost?

Becoming "part of us" can, of course, have multiple meanings. The technology might simply become part of how I consciously think of myself and want others to see me; it's part of my social identity, a way that people recognize me for who I am. It might, however, also become part of how I understand myself even at subconscious or neural levels, as when it enters into my body schema. Both philosophical and empirical work has suggested that we readily incorporate various tools into our body schemas,[16] and these seamlessly incorporated tools are sometimes considered an "extended body" beyond the confines of our skin. Similarly, we might envision the possibility of what philosophers Andy Clark and David Chalmers call "extended mind," where a person relies on external aids, such as smart phones or notebooks, to perform cognitive functions.[17] Neural engineering devices implanted in brains and designed to interface with existing nervous tissue in closed-loop systems may complicate our thinking about identity by changing our notions of both social identity and body schema. Shifts in identity can be a positive development—as when device-based alterations put us more in line with how we see ourselves—but they may also undermine identity in certain ways and come with acceptance costs for users.[18]

Think about Joan's case. Joan, like others who have become disabled, adjusts to this fact not only in how she interacts with her environment but in how she sees herself: "I am," she thinks, "a person without a functional arm." The transition to this new identity can be difficult and hard won. The offer of a smart prosthetic—even one that promises improved functionality—can be met with resistance if it is felt to undermine this new identity. Recall Joan's concern about her daughter's ability to recognize her *through her prosthetic*. Trying a new device is not necessarily cost free. Even when this cost is deemed worth paying, further threats to identity may follow soon upon adopting it. What if a new device is too cumbersome for her to achieve seamless integration into her body schema?[19] What if the new prosthetic is only temporary, a bridge technology to yet another device? Or what if Joan has lingering concerns about whether every action of the arm is indeed hers, given the possibility of device malfunction, design flaw, or interference from a third party? Her nagging doubts may stand in the way of her incorporating the device into her identity.[20] From Joan's perspective, a device's potential identity costs may be significant.

Her identity might also be stretched if the technology took a different form. The BCI used to control Joan's attached smart prosthetic might be used to control a detached assistive device, such as a robotic arm. In this case, the arm, even as it functions, no longer needs to be permanently attached to her body. The device might be wirelessly connected to her BCI in a way that allows her easy control of the device, and if the device had sufficient range and maneuverability, she might be able to send it to the next room on a task—getting her a drink or patting her child to sleep. Would a roaming robotic "arm" of this kind seem more like a mere tool, or would Joan be able to think of such an arm as "her arm" also? In the latter case, our common understandings of identity would experience foundational shifts.[21] Related questions about who and where an individual is might have implications for moral and legal responsibility as well.[22]

A different sort of identity question—one linked to concerns about authenticity[23]—arises in the context of John's DBS system. Will John feel like himself when his mood elevates as a result of an algorithmic feedback loop rather than through his usual physiologic process? If he experiences a negative mood that is context appropriate—perhaps in response to sad news—will he find it disorienting that the device quickly causes him to feel better? Would he even notice this change, or would it be something he feels obliged to explain, ex post facto, by reevaluating his own assessment of the news, given that his quick recovery from it occurs internally and without an obvious trigger, such as taking a pill? A person learns to "read" her reactions to events as a means of understanding her own internal states ("I must not have liked him so much, given that I didn't even cry when I heard he had died"). Similarly, John may find himself rethinking his values given his steady mood in the face of loss. Even if *he* does not notice, others around him may wonder about the authenticity of his responses, noting, for instance, that "John doesn't seem like himself" if his response does not fit the situation.

Carl Elliott has explored this question of authenticity in the context of neuropharmacology. Elliott critiques the widespread acceptance and promotion of Prozac and other selective serotonin reuptake inhibitors, given their capacity to make us feel better about situations in which we might more authentically feel despair. Some ways of responding to those situations are uniquely ours; they signal to others who we are. If we alter those response patterns through "smart" implants and effectively take the individual out of the response—even in the name of attempting to treat debilitating depression—we risk undermining the authenticity not just of the mood but also of the person's capacity for self-expression. As we

develop neural technologies, we need to examine our aspirational norms and the ways such technologies may interfere with (or perhaps enhance) our capacities for self-definition and self-expression.

As a preview, if an individual's BCI is hacked, that person's movements or mood could be manipulated by another. She could easily find this more threatening—for her sense of self, for her moral and legal responsibility, and so forth—than if her computer or bank account is hacked and data stolen. Similarly, if the datasets from a person's BCI are recorded, how might such recordings be used by educational systems, courts of law, or employers? With ever more complex recording systems and algorithms for identifying and translating intentions, concerns about privacy become

challenging, and not just for technical reasons.

Neural engineers often take themselves to be employing "objective" measures of normal and abnormal function. In John's example, normal brain function can be drawn in strict physiologic terms, such as by appeal to particular regional patterns of electrical activity or neurotransmitter levels. A device capable of repeatedly sampling physiologic brain function for deviation from the norm and iteratively intervening to reestablish normal function could constitute a closed-loop neuromodulatory system—a "pacemaker" for mood. Whether there is or could be one pattern of abnormal (or normal) electrical activity or neurotransmitter levels that faithfully represents mood across individuals, or even across the

a grip, and when does a lingering grip become awkward? A firm, prolonged handshake may communicate trustworthiness during a business interaction but could indicate aggression or one-upmanship in a meeting of new acquaintances. A purely scientific rendering of normal hand-shaking behavior, one that tries to set aside the normative, may not be possible.

Feminist and disability critiques have challenged prevailing notions of normality, pointing out that individuals are not abnormal or defective by virtue of their disabilities; rather, individuals are disabled due to inhospitable environments that make their abilities a poor fit for individually or socially desired ends.[25] The implication for development of neural devices is that "normal" may not be what all end users want and how normal

If the sensors that trigger John's deep brain stimulator indicate that he is depressed at a time when he denies feeling depressed and fails to exhibit outward signs of depression, is he in fact depressed?

entangled with our understandings of identity and authenticity. Other closed-loop implantable devices, such as cardiac pacemakers, function relatively autonomously, but their control is less obviously linked to central features of our identity.

Normality

The concept of normality is central to the development and implementation of BCI devices. Take, as an example, John's BCI used for the regulation of mood. A neuromodulatory device that aims to change an emotional state in a particular direction—say, from depressed to happy—relies on norms of affective function. This requires that we set parameters delineating what counts as the abnormal state to be corrected and what counts as the normal state to be sought. Ascertaining or setting the standards for "normal" function is

same individual over time, is an open question. Naturalistic theories of normality, such as Christopher Boorse's biostatistical theory or Norman Daniels's theory of species-typical functioning, would seem a natural place to start for defining normal brain function, but these approaches have been criticized for failing to recognize that appeals to normality are often ineliminably value laden.[24]

Take Joan's motor-oriented BCI and the ability to shake another's hand. The ability to shake hands is, on one level, a rather mundane sensorimotor skill, but it also is implicated in important social practices, such as expressions of autonomy, trust, friendship, negotiation, and courtesy, and this linkage complicates the definition of "normal." What counts as a normal grip, not too firm and not too soft? When is shaking another's hand too vigorous or too rigid? When is it too soon to release

is best understood may deviate from simple notions of replacing lost human functions.

Neural devices can also improve functioning in ways that raise questions about neuroenhancement. Imagine if Joan's prosthetic arm affords her greater strength or endurance and allows her to perform a job that displaces several other workers. Could such a device give her a competitive advantage? Would it privilege her in some way or yield benefits that are not earned in a traditional, "authentic" way, such as through strength or endurance training? Conversely, should such a neural device be embraced, provided it is medically safe, because it offers a chance to improve on arbitrary limitations to human abilities? Analogous worries have been raised by the use of neuropharmacology to enhance cognitive function and mood and by the use of prosthetics in sport.[26] The concept

of normality has been recognized as central to understanding and making progress in debates over pharmacological enhancement.[27] The same can be said about enhancement concerns raised by neural technology.[28]

The challenge of defining normality in BCIs may be complicated still further by the intertwining of related concepts such as authority and identity. For instance, physiologic measures of normal or abnormal function can come apart from introspective experience. If the sensors that trigger John's deep brain stimulator indicate that he is depressed at a time when he denies feeling depressed and fails to exhibit outward signs of depression, is he in fact depressed? Closed-loop neuromodulation of mood may lead to a disconnect between what is felt and what an individual or others think ought to be felt. The implications for identity from experiencing such a disconnect could be significant. Conversely, what if John reports feeling depressed but objective measures do not bear this out? What or who is the ultimate arbiter of normal affective function? Where does authority reside?

Authority

The goal of BCI technology is to translate brain processes underlying thought and action into desired outcomes, like grasping of a prosthetic hand or elevation of mood. BCIs are mediated by complex algorithms that take data from carefully placed brain sensors or electrodes, mine them for a desired signal, and convert them into a mechanical or electrical activity. Normative questions arise along this translational pathway.[29] How are the relevant characteristics of a common (or, as above, "normal") input signal to be defined? What signal qualifies as a person's intention, rather than a fleeting, fragmented, or even personally abjured thought? The challenge for neural engineers is not only to design a signal processing algorithm sufficiently sensitive to allow Joan's arm to gently grasp an offered

flower, for instance, but also selective enough to prevent her hand from crushing the flower (or whatever else she is holding, like her daughter!) as a scene from *Little Shop of Horrors* momentarily passes through her consciousness. And perhaps most importantly, how much authority should we invest in the translational algorithms of BCIs and in what ways?

BCI systems offer a powerful, alternative way to access a person's mental life apart from first-person testimony. While BCI systems like Joan's are not designed to monitor intention, they might be able to record past commands. If so, we can imagine instances where BCI recordings disagree with a user's subjective reports. A BCI might indicate, for instance, that Joan "meant" to turn a steering wheel to the right and into a neighbor's fence, despite her insistence to the contrary. The alternative ontology afforded by BCIs has the potential to reshape how we understand and lend credence to claims of self-knowledge.

Such discussions of authority have some precursors in the literature on the neuroimaging of pain. Functional MRI has been explored as a possible tool in the diagnosis of chronic pain.[30] Whereas diagnosis and treatment of pain have traditionally relied predominantly on first-person subjective reports (such as "I am in pain now" and "My pain is a 4 out of 10"), neuroimaging may offer an alternative way to assess the level, kind, and even the presence of pain. The move toward neuroimaging evidence of pain raises concern. The limitations of neuroimaging of pain can be underestimated, as can the potential harm to individuals when imaging is misread or overinterpreted; if "objective" measures are placed on a par with personal testimony, false negative results may deny or diminish legitimate claims of pain-related suffering in both legal and medical contexts.[31] Consequently, medical practitioners should use a precautionary principle in cases of subjective-objective disagreement, according to which the deference to subjective

reports reflects the seriousness of error.[32] Responsible use of the technology is not provided by the technology itself, and "objective" measures, such as colorful images of pain, may provide a kind of "illusory accuracy."[33]

A precautionary principle in the setting of subjective-objective disagreement is a lesson that could carry over into BCI systems. The complicated intersectional nature of agency, however, prevents the precautionary principle from applying straightforwardly. If Joan has come to include a BCI as part of her identity, then a device malfunction or a subjective-objective disagreement might be an experience of self-alienation that a precautionary principle on its own will not ameliorate. Joan may adamantly deny intending to swerve her car, but the mere fact of the event may be enough to engender feelings of guilt and self-doubt.[34] Legal and moral responsibility leads to similar complications; unlike cases of pain assessment, which are more clearly about limited reports of individual needs, legal and moral judgments will be made about *any* of the actions Joan makes with her prosthesis. Recall Joan's denial of an intention to run her car into her neighbor's fence. If a court of law introduces BCI recordings as evidence, and the data set clearly shows that there was such a motor intention immediately preceding the action, what should be done? Here, a precautionary principle is not a clear solution to conflicts between the objective and the subjective, though it may offer an appropriate starting point.

Moral and Legal Responsibility

We typically hold people responsible for actions over which they can exercise control. Taking responsibility for our own actions and holding others to account is a fundamental feature of living in moral community with others.[35] Neural devices, as we have seen in the cases of John and Joan, can complicate our notions of responsibility. Neural

devices can provide a new source of information with which we can judge responsibility, as in John's closed-loop DBS recording of electrical activity related to his mood, even if the extent to which authority should be vested in this information is not self-evident. Neural devices can also be themselves *involved in* actions for which responsibility is at issue, such as in Joan's car accident. Insofar as they influence actions, thoughts, or feelings, they affect responsibility, both moral and legal.[36]

An individual with a neural device can be held responsible in different ways. Recall Joan's prosthetic arm and the car accident that causes destruction of her neighbor's fence. She claims she did not intentionally turn the car. Still, she might be

difficulty of her task. Fairly distributing responsibility can be a complicated affair.

Individuals with neural devices may be owed special moral consideration by others. People with disabilities can come to rely on such devices to secure social and other goods. For instance, Joan's device may allow her to gain employment, care for her child without assistance, or maintain a healthy sense of self-confidence. Given the extent of her reliance on the device, she may be owed special consideration, such as affordable replacement in the case of device failure. This is in part a consideration of distributive justice, but it may be more than this. It may also be a responsibility of everyone who encounters her individually to acknowledge

Responsibility for actions involving neural devices is made more complicated by the intersection of identity and responsibility. Individuals with neural devices may not only become functionally dependent on devices but, as we have seen, may also incorporate these devices into their sense of self and body schema. If the connection to identity is taken seriously, the implications for responsibility might be striking. For instance, if Joan's prosthetic arm becomes an extension of her body and an integral part of her identity, its destruction could be traumatic. From a moral standpoint, its willful destruction could be taken as a significant moral transgression. From a legal standpoint, its destruction might be appropriately classified as battery, rather than mere property

A brain-computer interface might indicate that Joan "meant" to turn a steering wheel to the right and into a neighbor's fence, despite her insistence to the contrary.

held responsible for choosing to have and use a prosthetic arm in the first place; after all, had she not had it, her neighbor's fence might be safe. More commonly we ascribe responsibility for actions more proximate to the event of interest. Along these lines, we might hold Joan responsible for failing to train her arm adequately and in so doing increasing the risk of a resultant harm. If lost control due to inexperience were a foreseeable possibility, as intoxicated driving is a foreseeable consequence of alcohol ingestion, some attribution of responsibility might be appropriate.[37] Or we might hold Joan responsible for a momentary lapse of focus immediately preceding the event. However, we might absolve Joan of moral or legal responsibility if the device were poorly constructed or made use of faulty software and, in turn, malfunctioned. We might also wonder if the level of concentration required by Joan to control her prosthetic should temper the assignment of responsibility for mistakes, given the relative

the device's value in their interactions with her. A sudden malfunction of the prosthetic would not be like losing a favorite hat; it would be more akin to a blind person's losing her cane on a busy street.

Neural devices also affect responsibility insofar as they leave an auditable information trail. Whereas assigning responsibility for an action typically entails piecing together its history out of available (and sometimes unreliable) elements—reported memories, states of mind, environmental circumstances—neural devices offer the possibility of a detailed history of brain states leading up to an action. Even setting aside debates about how such brain states are causally related to actions, the mere presence of this trove of information is significant. It offers a detailed and available—and hence likely quite attractive—source of information about an action of interest. Courts, insurance companies, and others will be interested in that data.

destruction, even if it is not being worn at the time.[38]

Privacy

As the technologies to monitor and intervene in complex neurological systems become more robust and useful to end users, platforms like Joan's prosthetic arm are likely to be integrated with wireless technologies, many of which are already implemented in consumer products. Wireless standards like near-field communication—exemplified by the ubiquitous radio-frequency identification systems in cellular phones, debit cards, security passes, and so on—are vulnerable to tampering, misuse, and attack. BCIs using such standards will be made vulnerable to the same security exploits that affect all devices using those standards. Again, Joan's prosthetic device might store her previous motor commands such that they could be retrieved later by either a medical professional or, potentially, a malevolent agent. Simi-

larly, John's closed-loop deep brain stimulator might record the level and frequency of his treatment such that its wireless diagnostic interface could be used to extract details about his medical condition.

These two possibilities illustrate how neural engineering involves privacy at multiple levels. Illicit access to John's neurostimulator seems to recall issues of "brain privacy" raised in previous neuroethics literature that considers how neuroimaging technologies might reveal an individual's psychological traits or mental states ("brainotyping"), attitudes toward other people, and truthfulness.[39] Even without imaging, a stranger who gains access to John's closed-loop DBS recordings might be able to infer analogous details about John's emotional state or psychological disposition from the stolen data. While brain privacy doesn't quite capture the threat posed by stolen recordings from Joan's motor prosthetic, she might experience a threat to privacy if her data is combined with other non-private information about her—from social media, tracked web activity, or public record.[40] Further, BCIs are vulnerable to "brain spyware": malicious programs that can extract private information from the right kinds of neurological data. An attacker might, for example, present Joan with specially designed visual stimuli and derive private data from her neurophysiologic responses.[41]

One response to both of these threats to privacy might be deflationary; we need just take the relevant security precautions to prevent hacking in the first place. Appropriate encryption and design constraints could eliminate nearly all security risks. That might be a reasonable response in the context of computer privacy. In the case of BCIs, however, it seems more is at stake. If an attacker compromises a user's personal computer, the attacker will be able to access that user's data, perform tasks using his computer or disable that computer altogether. If an attacker were to compromise John's or Joan's device, the

attacker might access stored neurological data, control his or her devices (against the victim's will), or disable the devices entirely. The task of securing a neurological device is a grave one: "instead of protecting the software on someone's computer, we are protecting a human's ability to think and enjoy good health."[42] To this end, Tamara Denning and colleagues coin the term "neurosecurity" as "the protection of the confidentiality, integrity, and availability of neural devices from malicious parties with the goal of preserving the safety of a person's neural mechanisms, neural computation, and free will" (p. 2).

Identity complicates privacy concerns: we have yet to anticipate how a user's self-image will change when her personal neurological data can be accessed the same way we might access a file. At the very least, it seems that something that was once inexorably private, something that often comprises our sense of self, has become potentially public: available for access, interference, or inquisition. Perhaps there is something potentially disquieting about this shift itself. We can call this—picking up on Denning's nomenclature—the potential accessibility problem, or figuring out where to draw the boundary between public and private given the very existence of devices that collect and analyze neurological data.[43]

Whatever safeguards engineers implement to prevent breaches of security and privacy in BCIs and other neural devices, issues of fairness or justice might still arise or even be exacerbated. In the near term, BCIs will most likely be the end user's last resort treatment option for both motor control and neurostimulation. As such technologies advance, users may feel pressure to accept a technology in order to address their concerns. Will consenting to a BCI with wireless technology mean consenting to diminished privacy or increased potential accessibility? Could such devices be built differently to address end user needs without compromising the privacy of collected data or

the security of the device's functions? Traditionally wired systems, while less convenient, would avert at least some of the privacy concerns. What kinds of compromises is it reasonable to expect users to accept?

Justice

Like other health-related technologies, neural devices raise issues of justice with regard to the distribution of harms and benefits and to the inclusion of perspectives from people likely to be affected by the technology. Although distribution concerns tend to attract the most attention in debates about justice—and have relevance here, particularly given the resource expense required to develop technologies that may never translate into widely available consumer goods—we should not underestimate the significance of the concept of *justice as recognition*.[44] How do we ensure that groups who are often differently socially positioned—such as the disabled people who are the intended beneficiaries of such technologies—have their perspectives on the meaning and significance of their bodily differences and these technologies heard and respected? Even once funding decisions have been made, how might we ensure that end users like Joan and John have their perspectives and concerns integrated into the development process?[45] If the technologies in question are to be designed so that end users will adopt them, attention must be paid to their particular needs, concerns, and experiences well before final products are determined. Justice as recognition demands explicit and meaningful engagement of likely downstream end users at major decision points in the design of neural technologies.[46]

To understand why, consider the potential individual trade-offs of adopting some of these technological devices. Unlike a traditional prosthetic, a BCI like Joan's does not allow the flexibility to completely abandon the device whenever she wants. An implanted BCI would require medical

intervention for complete removal, and it may provide limited ability to turn on and off its recording functions.[47] As a consequence, her macro level of control and privacy may be more limited with the BCI compared to a traditional prosthetic.[48] For some, these potential threats may not be important, especially given the sense of freedom and independence that the BCI may provide. Still, if Joan perceives her device as acting unreliably and cannot distinguish whether this is a problematic design issue or a malicious hacking event, we may consider it at least reasonable that she would prefer the capacity to simply

a risk to his health or that of others; he may neurostimulate in such a way that produces a particularly enhanced mood, perhaps with side effects linked to mania.[50] Who should determine the appropriate settings in such cases? Sharing of information about neural stimulation or decision-making with a health care provider, for instance, may seem justifiable in John's case. Some might also consider it permissible for a health care provider to intervene if John was experiencing a depressive episode, even if he chose not to neurostimulate, much as psychiatric patients are sometimes committed to involuntary holds

for some marginalized communities, use of the device may be an attempt personally, socially, or institutionally to address broader social situations that deserve attention but often do not receive adequate social support and funding.

Ultimately, attending to concerns of justice as recognition will require that end users have input about design decisions and have eventual access to multiple options and the flexibility to change their minds as life warrants. To ensure this, researchers must begin to recognize the trade-offs that individuals face when choosing and using a device. Because of the complexity of

We have yet to anticipate how a user's self-image will change when her personal neurological data can be accessed the same way we might access a file.

turn it off. In the interest of justice, researchers might be compelled to direct resources toward development of noninvasive BCIs, like caps or bands using electroencephalography, which may provide more user flexibility, although signal resolution still poses a technical hurdle. Balancing moral considerations with technical ones is a notable challenge but central to equitable decision-making. Consider the cochlear implants debates and the importance of allowing individuals to choose multiple pathways for enhanced communication (cochlear implants, speech therapy, sign language interpreters, or some combination).[49] In neural engineering, similar concerns arise if dominant social forces and arguably narrow notions about "normal" functioning pressure individuals to choose one modality over the other or to give up local control over the capacity to "turn off" a prosthetic.

In John's case, our moral as well as legal understanding of responsibility may complicate the concern about local control. Assuming that John has freedom to self-regulate his deep brain stimulator, he may pose

based on perceived threats to self or others. While the health care provider may be acting in what she perceives as the best interest of her patient, John might disagree, or alternatively, he might understand his depression as part of who he is, a significant factor in his identity. This kind of discrepancy between a health care provider and psychiatric patients has been well-documented by the Mad Pride, psychiatric survivors, and neurodiversity movements.[51]

As in the case with Joan, some might argue that John could have chosen an alternative to a BCI and that these potential consequences were foreseeable and avoidable. In choosing the BCI, he accepted these risks. Yet even if John and his health care provider understand the therapeutic benefits of a BCI (for example, better adherence, reliable access to medication, and less need for transportation to get to therapy) as significant, perhaps they would still recognize John's need to retain some form of control, and they might even prefer to err on the side of giving more control to the individual user. In terms of justice, we may also consider that, particularly

end users' lives, neural engineers must engage with end users to design and develop devices that address community and individual needs.

The Beginning of a Discussion

As public funding for neurotechnology expands and interest in neural engineering increases, the ethical issues raised by these technologies must be carefully explored and analyzed.[52] Important steps have been made in this direction,[53] but more needs to be done. The Presidential Commission for the Study of Bioethical Issues, for instance, notes that advances in neuroscience raise complex issues related to cognitive enhancement, consent capacity, and legal responsibility and decision-making. Focusing attention on these areas of ethical, social, and legal concern is an important step for neuroscience and society, but understanding and addressing such concerns (and others) will require further empirical and normative work—such as what we hope to have begun to offer here.

The six core areas of ethical concern that we have identified—identity,

normality, authority, responsibility, privacy, and justice—by no means form an exhaustive list; other areas of ethical concern, such as stigma and autonomy, could be added. Our list derives from both our ongoing discussions with neural engineers and from the bioethics and neuroethics literature. We believe that these six areas cover a substantial swath of the conceptual ground relevant to neural engineering and provide a starting point for discussion inside and outside neural engineering. Our hope is that these core areas can be a useful scaffolding for scholars and others as they work through challenges ushered in by neural engineering.

Acknowledgments

Disclosure

Notes

1. M. Farah and P. Wolpe, "Monitoring and Manipulating Brain Function," *Hastings Center Report* 34, no. 3 (2004): 35-45.

2. Center for Sensorimotor Neural Engineering, http://csne-erc.org/.

3. M. Schermer, "Ethical Issues in Deep Brain Stimulation," *Frontiers in Integrative Neuroscience* 5 (2011): 1-5.

4. H. F. M. Van der Loos, "Design and Engineering Ethics Considerations for Neurotechnologies" *Cambridge Quarterly of Healthcare Ethics* 16 (2007): 305-9.

5. J. Sadowski, "Exoskeletons in a Disabilities Context: The Need for Social and Ethical Research," *Journal of Responsible Innovation* 1, no. 2 (2014): 214-19.

6. N. S. Fitz and P. B. Reiner, "The Challenge of Crafting Policy for Do-It-Yourself Brain Stimulation," *Journal of Medical Ethics* 41 (2015): 410-12.

7. E. Klein, "Models of Patient-Machine-Clinician Relationship in Closed-Loop Machine Neuromodulation," in *Machine Medical Ethics*, ed. S. P. Van Rysewyk and M. Pontier (New York: Springer, 2014), 273-90.

8. J. B. Trimper, P. R. Wolpe, and K. S. Rommelfanger, "When 'I' Becomes 'We': Ethical Implications of Emerging Brain-to-Brain Interfacing Technologies," *Frontiers in Neuroengineering* 7, no. 4 (2014): 1-4.

9. In generic terms, a BCI is a device for capturing and using brain-derived information to facilitate human control and communication. The extension of the term "BCI" is itself contested, and competing terms such as "brain-machine interface" and "neuroprosthesis" populate the literature. While these terminological debates are important, given our interest in the philosophical and ethical implications of BCI as a type of technology, we will largely put them to the side. For more on these debates, see F. Nijboer et al., "The Asilomar Survey: Stakeholders' Opinions on Ethical Issues Related to Brain-Computer Interfacing," *Neuroethics* 6, no. 3 (2013): 541-78.

10. R. Heersmink, "Embodied Tools, Cognitive Tools, and Brain-Computer Interfaces," *Neuroethics* 6, no. 1 (2013): 207-19.

11. BRAIN Initiative, accessed April 9, 2015, http://www.whitehouse.gov/share/brain-initiative.

12. See, for instance, P. Gallagher, D. Desmond, and M. MacLachlan, eds., *Psychoprosthetics* (London: Springer, 2007); F. B. Mills, "A Phenomenological Approach to Psychoprosthetics," *Disability and Rehabilitation* 35, no. 9 (2013): 785-91.

13. M. Auvray and E. Myin, "Perception with Compensatory Devices: From Sensory Substitution to Sensorimotor Extension," *Cognitive Science*, 33 (2009): 1036-58.

14. C. Elliott, "The Perfect Voice," chapter 1 in *Better than Well: American Medicine Meets the American Dream* (New York: W. W. Norton & Company, 2004): 1-27.

15. C. D. Murray, "An Interpretive Phenomenological Analysis of the Embodiment of Artificial Limbs," *Disability and Rehabilitation* 26 (2004): 963-72; talking about bone-anchored prosthetic limbs, Mari Lundberg et al. note three ways that prosthetic users conceive of their prosthetics: as a practical tool, an artificial body part, and as "part of me" (with increasing identification with the prosthetic); see M. Lundberg, K. Hagberg, and J. Bullington, "My Prosthesis As a Part of Me: A Qualitative Analysis of Living with an Osseointegrated Prosthetic Limb," *Prosthetics and Orthotics International* 35 (2011): 207-14.

16. A. Clark, "Reinventing Ourselves: The Plasticity of Embodiment, Sensing and Mind," *Journal of Medicine and Philosophy* 32, no. 3 (2007): 263-82; J. M. Carmena et al., "Learning to Control a Brain-Machine Interface for Reaching and Grasping by Primates," *PLOS Biology* 1, no. 2 (2003): e42, doi:10.1371/journal.pbio.0000042; A. Sengul et al., "Extending the Body to Virtual Tools Using a Robotic Surgical Interface: Evidence from the Crossmodal Congruency Task," *PLoS ONE* 7, no. 12 (2012): e49473, doi:10.1371/journal.pone.0049473.

17. A. Clark and D. Chalmers, "The Extended Mind," *Analysis* 58 (1998): 7-19; A. Clark, *Supersizing the Mind: Embodiment, Action and Cognitive Extension* (New York: Oxford University Press, 2008).

18. M. Hilhorst, "'Prosthetic Fit': On Personal Identity and the Value of Bodily Difference," *Medicine, Health Care, and Philosophy: A European Journal* 7, no. 3 (2004): 303-10.

19. Heersmink observes that the current BCI technology is too cumbersome—requiring extreme concentration and limited range of movement—to be transparent to the user; see Heersmink, "Embodied Tools, Cognitive Tools, and Brain-Computer Interfaces."

20. Schupbach reported that nineteen of twenty-nine patients who received deep brain stimulators for Parkinson's symptoms had issues recognizing themselves after surgery, and six of the twenty-nine experienced the change as deeply problematic, noting, for instance, "I feel like a robot," and, "I don't feel like myself anymore"; see M. Schupbach et al., "Neurosurgery in Parkinson's Disease: A Distressed Mind in a Repaired Body?" *Neurology* 66, no. 12 (2006): 1811-16.

21. It is also worth considering literature on identity and remote-controlled virtual avatars, to which serious gamers often become deeply attached; see, for example, S. Turkle, "Parallel Lives: Working on Identity in Virtual Space," in *Constructing the Self in a Mediated World*, ed. D. Grodin and T. R. Lindlof (London: Sage, 1996), 156-75.

22. J. Wolfendale, "My Avatar, My Self: Virtual Harm and Attachment," *Ethics and Information Technology* 9 (2007): 111-19.

23. See V. Johansson et al., "Thinking Ahead about Deep Brain Stimulation," *AJOB Neuroscience* 5, no. 1 (2014): 24-33. The authors acknowledge that authenticity worries may be inherent in the technology, particularly for bidirectional DBS devices, but downplay the concerns by suggesting that they may be more tied to the technology's novelty than to well-founded moral worries.

24. M. Synofzik, "Ethically Justified, Clinically Applicable Criteria for Physician Decision-Making in Psychopharmacological Enhancement," *Neuroethics* 2 (2009): 89-102.

25. A. Silvers, "A Fatal Attraction to Normalizing: Treating Disabilities as Deviations

from 'Species-Typical' Functioning," in *Enhancing Human Traits*, ed. E. Parens (Washington, D.C.: Georgetown University Press, 1998), 95-123; R. Amundson and S. Tresky, "Bioethics and Disability Rights: Conflicting Values and Perspectives," *Journal of Bioethical Inquiry* 5, no. 2-3 (2008): 111-23.

26. H. Greely et al., "Toward Responsible Use of Cognitive-Enhancing Drugs by the Healthy," *Nature* 456 (2008): 702-5; G. Wolbring, "Paralympians Outperforming Olympians: An Increasing Challenge for Olympism and the Paralympic and Olympic Movement," *Sport, Ethics and Philosophy* 6, no. 2 (2012): 251-66.

27. A. Roskies, "Neuroethics for the New Millennium," *Neuron* 35, no. 1 (2002): 21-23; D. Buchman and J. Illes, "Imaging Genetics for Our Neurogenetic Future," *Minnesota Journal of Law, Science, & Technology* 11, no. 1 (2010): 79-97.

28. M. Schermer, "Health, Happiness and Human Enhancement: Dealing with Unexpected Effects of Deep Brain Stimulation," *Neuroethics* 6, no. 3 (2013): 435-45.

29. I. de Melo Martin, "Defending Human Enhancement Technologies: Unveiling Normativity," *Journal of Medical Ethics* 36, no. 8 (2010): 483-87.

30. T. D. Wager et al., "An fMRI-Based Neurologic Signature of Physical Pain," *New England Journal of Medicine* 368 (2013): 1388-97.

31. K. Davis, E. Racine, and B. Collett, "Neuroethical Issues Related to the Use of Brain Imaging," *Pain* 153, no. 8 (2012): 1555-59.

32. J. Giordano, "The Neuroscience of Pain, and a Neuroethics of Pain Care," *Neuroethics* 3 (2010): 89-94.

33. Farah and Wolpe, "Monitoring and Manipulating," 40.

34. Bernard Williams argues that such agent regret is morally expected, even when there was nothing else an agent could have done to avoid an accident; this case is somewhat different in that Joan may actually question whether she was in control of her action. See B. Williams, "Moral Luck," in *Moral Luck*, ed. D. Statman (Albany, NY: SUNY Press, 1993), 35-55.

35. D. Shoemaker, "Responsibility and Disability," *Metaphilosophy* 40, no. 3-4 (2009): 438-61.

36. N. Lipsman and W. Glannon, "Brain, Mind and Machine: What Are the Implications of Deep Brain Stimulation for Perceptions of Personal Identity, Agency and Free Will?," *Bioethics* 27, no. 9 (2013): 465-70.

37. L. Klaming and P. Haselager, "Did My Brain Implant Make Me Do It?," *Neuroethics* 6 (2010): 527-39.

38. G. Ramachandran, "Assault and Battery on Property," *Loyola Law Review* 44 (2010): 253-76.

39. Farah and Wolpe, "Monitoring and Manipulating."

40. L. Austin, "Privacy and the Question of Technology," *Law and Philosophy* 22, no. 2 (2003): 119-66.

41. T. Bonaci et al., "Securing the Exocortex: A Twenty-First Century Cybernetics Challenge," paper presented at the 2014 IEEE Conference on Norbert Wiener in the 21st Century, Boston, MA, June 2014.

42. T. Denning, Y. Matsuoka, and T. Kohno, "Neurosecurity: Security and Privacy for Neural Devices," *Journal of Neurosurgical Focus* 27, no. 1 (2009): E7, 1-4.

43. Martha Farah recognizes this complexity with respect to privacy and identity in M. J. Farah, "Neuroethics: The Practical and the Philosophical," *Trends in Cognitive Sciences* 9, no. 1 (2008): 34-40.

44. N. Fraser and A. Honneth, *Redistribution or Recognition? A Political-Philosophical Exchange* (New York: Verso, 2003).

45. In practice, some researchers understand "end user" to mean industry partners or other research scientists, because any clinical application is seemingly far downstream. Consequently, the distinct needs of underserved groups—like people with disabilities—may be overlooked during the development process.

46. A. Silvers, "Better than New! Ethics for Assistive Technologists," in *Design and Use of Assistive Technology: Social, Technical, Ethical, and Economic Challenges*, ed. M. M. K. Oishi, I. M. Mitchell, and H. F. M. Van der Loos (New York: Springer, 2010).

47. On the relative irreversibility of BCI systems, see M. Synofzik and T. Schlaepfer, "Electrodes in the Brain: Ethical Criteria for Research and Treatment with Deep Brain Stimulation for Neuropsychiatric Disorders," *Brain Stimulation* 4 (2011): 7-16. In a different paper by Matthis Synofzik et al., it is suggested that "[t]he alleged reversibility of DBS—which is still stated by most authors as one of the main ethical 'pro' arguments of DBS . . . —might only apply to the technique, but not to the person"; see M. Synofzik, T. E. Schlaepfer, and J. J. Fins, "How Happy Is Too Happy? Euphoria, Neuroethics, and Deep Brain Stimulation of the Nucleus Accumbens," *AJOB Neuroscience* 3, no. 1 (2012): 30-36, at 34-35.

48. This contrast may be yet too simplistic, as refinements in implantable devices, such as DBS systems, may increasingly allow for some patient control over stimulation. See E. Klein, "Models of the Patient-Machine-Clinician Relationship in Closed-Loop Machine Neuromodulation," in *Machine Medical Ethics*, ed. S. P. van Rysewyk and M. Pontier (New York: Springer, 2014), 273-90.

49. J. B. Christiansen and I. Leigh, *Cochlear Implants in Children: Ethics and Choices* (Washington, D.C.: Gallaudet University Press, 2002); R. Sparrow, "Implants and Ethnocide: Learning from the Cochlear Implant Controversy," *Disability & Society* 25, no. 4 (2010): 455-66.

50. For instance, Synofzik et al. describe a patient with a DBS system for general anxiety and obsessive-compulsive disorder. After calibration sessions in which the patient reported feeling "'unrealistically good' and . . . 'overwhelmed' by the sensations of happiness and ease" and asked to have the stimulations levels reduced, the patient returns and requests the higher stimulation again, because he would "like to feel 'a bit happier' during the next few weeks"; see Synofzik, Schlaepfer, and Fins, "How Happy Is Too Happy?"

51. B. Lewis, "A Mad Fight: Psychiatry and Disability Activism," in *The Disability Studies Reader* 4th ed., ed. L. J. Davis (New York: Routledge, 2013), 115-31; S. M. Robertson, "Neurodiversity, Quality of Life, and Autistic Adults: Shifting Research and Professional Focuses onto Real-Life Challenges," *Disability Studies Quarterly* 30, no. 1 (2009), accessed April 9, 2015 http://dsq-sds.org/article/view/1069.

52. R. M. Green, "The Need for a Neuroscience ELSI Program," *Hastings Center Report* 44, no. 4 (2014): inside back cover.

53. Presidential Commission for the Study of Bioethical Issues, *Gray Matters: Integrative Approaches for Neuroscience, Ethics, and Society* (PCSBI: Washington, D.C., 2014), at http://bioethics.gov/node/3543.