

UC Berkeley

UC Berkeley Previously Published Works

Title

Recommendations for Responsible Development and Application of Neurotechnologies

Permalink

<https://escholarship.org/uc/item/74h49860>

Journal

Neuroethics, 14(3)

ISSN

1874-5490

Authors

Goering, Sara
Klein, Eran
Specker Sullivan, Laura
[et al.](#)

Publication Date

2021-12-01

DOI

10.1007/s12152-021-09468-6

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NoDerivatives License, available at <https://creativecommons.org/licenses/by-nd/4.0/>

Peer reviewed



Recommendations for Responsible Development and Application of Neurotechnologies

Sara Goering · Eran Klein · Laura Specker Sullivan · Anna Wexler · Blaise Agüera y Arcas · Guoqiang Bi · Jose M. Carmena · Joseph J. Fins · Phoebe Friesen · Jack Gallant · Jane E. Huggins · Philipp Kellmeyer · Adam Marblestone · Christine Mitchell · Erik Parens · Michelle Pham · Alan Rubel · Norihiro Sadato · Mina Teicher · David Wasserman · Meredith Whittaker · Jonathan Wolpaw · Rafael Yuste

Received: 11 June 2020 / Accepted: 15 April 2021 / Published online: 29 April 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract Advancements in novel neurotechnologies, such as brain computer interfaces (BCI) and neuromodulatory devices such as deep brain stimulators (DBS), will have profound implications for society and human rights. While these technologies are improving the diagnosis and treatment of mental and neurological diseases, they can also alter individual agency and estrange those using neurotechnologies from their sense of self, challenging basic notions of what it means to be human. As an international coalition of interdisciplinary scholars and practitioners, we examine these challenges and make

recommendations to mitigate negative consequences that could arise from the unregulated development or application of novel neurotechnologies. We explore potential ethical challenges in four key areas: identity and agency, privacy, bias, and enhancement. To address them, we propose (1) democratic and inclusive summits to establish globally-coordinated ethical and societal guidelines for neurotechnology development and application, (2) new measures, including “Neurorights,” for data privacy, security, and consent to empower neurotechnology users’ control over their data, (3) new methods of identifying

S. Goering (✉) · E. Klein · M. Pham
University of Washington, Seattle, WA, USA
e-mail: sgoering@uw.edu

E. Klein
Oregon Health & Science University, Portland, OR, USA

L. Specker Sullivan
Fordham University, New York, NY, USA

A. Wexler
University of Pennsylvania, Philadelphia, PA, USA

B. Agüera y Arcas · M. Whittaker
Google, Mountain View, CA, USA

G. Bi
University of Science and Technology of China, Hefei, China

G. Bi
CAS Shenzhen Institute of Advanced Technology, Shenzhen, China

J. M. Carmena · J. Gallant
University of California, Berkeley, CA, USA

J. J. Fins
Weill Cornell Medicine, New York City, NY, USA

P. Friesen
McGill University, Montreal, Canada

and preventing bias, and (4) the adoption of public guidelines for safe and equitable distribution of neurotechnological devices.

Keywords Identity · Agency · Enhancement · Privacy · Bias · Public policy

“Science without conscience is but the ruin of the soul” Rabelais

Introduction

The continued development of neurotechnologies is likely to profoundly alter the human experience. Devices such as brain computer interfaces (BCI) and deep brain stimulators (DBS) interact directly with the human brain, whether from electrodes implanted deep in the brain, electrodes on the surface of the brain, or non-invasive devices that operate through the skull. These developments are being driven by a number of large global neuroscience initiatives, including the United States-based Brain Research through Advancing Innovative Neurotechnologies (BRAIN) initiative [1], the Human Brain Project (HBP) in Europe [2], and other coordinated research projects around the world including China, Japan, South Korea, Australia and Canada

[3]. Industry is also pushing the field forward, with companies such as Neuralink, Kernel, Iota, ControlLabs, Facebook and Microsoft, among others, devoting major investments to neurotechnology development, and probably already surpassing public investments [4]. Neurotechnologies are enabling users to manipulate distant objects [5], prevent, mitigate, or prepare for disruptive neurological events [6], and monitor, influence or regulate mood, emotion, and memory [7].

Yet, these phenomenal feats of science and technology are double-edged. The neural modifications brought about by BCIs and DBS, sensory or motor augmentation devices, and other emerging technologies may not only enhance experiences of agency, but also reduce or confuse them. They have the potential to disrupt users’ narratives, estrange them from actions and emotions they should feel ownership over, and make their sense of self more precarious. They may intrude on the key domains of privacy that are important for maintaining a discrete sense of self, and they may also significantly shift the norms of human functioning within a society. Even more broadly, these technologies may alter the connection between the body and the mind and blur current boundaries between minds.

These threats may arise not only from the potential abuse of these neurotechnologies but also from the unintended consequences of their intended uses. This

J. E. Huggins
University of Michigan, Ann Arbor, MI, USA

P. Kellmeyer
University Medical Center Freiburg, Freiburg, Germany

A. Marblestone
Federation of American Scientists, Washington, DC, USA

C. Mitchell
Harvard University, Cambridge, MA, USA

E. Parens
The Hastings Center, Philipstown, Garrison, NY, USA

A. Rubel
University of Wisconsin-Madison, Madison, WI, USA

N. Sadato
National Institute for Physiological Sciences, Okazaki, Aichi, Japan

M. Teicher
Bar-Ilan University, Ramat Gan, Israel

D. Wasserman
University of Maryland, College Park, MD, USA

M. Whittaker
AI Now, New York City, NY, USA

M. Whittaker
New York University, New York City, NY, USA

J. Wolpaw
National Center for Adaptive Neurotechnologies, Albany, NY, USA

R. Yuste (✉)
Columbia University, New York City, NY, USA
e-mail: rmy5@columbia.edu

means that, while neurotechnologies can certainly aid and improve the experiences of individuals who seek to use them, they may also inadvertently threaten features of human experience that society cares about preserving. Furthermore, attention to how existing biases influence neurotechnology development, and to how norms of human functioning may be radically altered through enhancement uses of neurotechnology will be critical.

This paper is a product of a large interdisciplinary, multi-national workshop that took place in 2017, with the aim of creating recommendations for developing novel neurotechnologies. The group¹ defined neurotechnology broadly, given the rapid pace at which novel developments are occurring and the desire to think expansively about potential effects of new modes of access to the brain, whether through invasive or non-invasive means, and whether for monitoring or modulating the brain itself, controlling targets outside the brain, or both. During an intensive three-day workshop, the group deliberated over four key areas of concern: identity and agency, privacy, bias, and enhancement. These areas of concern were initially identified as potential foci by workshop organizers, and, following initial exploratory group discussions, were refined and endorsed by the larger group. Following the workshop, the group published a short commentary [8] to initiate a conversation about the importance of developing recommendations in these areas. The goal of this paper is to develop those initial ideas, and to emphasize our shared commitment – across a large group with diverse training, social positioning, and investment in neurotechnology – to taking action.

In the intervening years, other neurotechnology guideline efforts have been published [9–11]. Those welcome efforts provide valuable first steps, but they require expansion and further elaboration. They call attention to some key components of responsible innovation (e.g., safety, privacy protection, attention to agency and autonomy) but miss some broader concerns, and often lack more specific recommendations. For instance, the NIH working group's guidelines [9] focus more on individual consent and understanding of risk than broader societal risks of neural interventions, and recommend attention to the management of private neural data without more explicit guidance. The neuroethics questions provided by the Global Summit Delegates (2018) [10] identify potential cultural differences in understanding of

neuroethical issues (e.g., privacy), and warn that social or cultural bias may affect research design. Their list of questions is an important contribution to the responsible innovation effort, but they do not yet put forward options for answering them. The OECD 2019 report [11] articulates a set of guiding principles – including the need for trustworthy and agile governance structures, protection of cognitive liberty and autonomy, attention to social inequality and potential exacerbations of it through neurotechnology, and the need for a diverse workforce – that provides an overarching set of commitments to guide responsible innovation in neurotechnology. Still, the guidelines are succinct and call out for additional justification and elaboration regarding implementation. The recommendations developed in this paper go beyond these reports and are intended to signal the need for greater anticipatory regulation of a field that holds significant promise, but may also threaten key features of human life. Rather than viewing these potential perils as hyperbole or too far off to merit close attention, this paper offers several concrete recommendations and identifies governance structures that should be developed soon in order to meet this challenge.

This paper includes reflections in four focused areas: 1) identity and agency, 2) privacy, 3) bias, and 4) enhancement. Each includes a discussion of ethical and societal challenges raised by neurotechnologies. Recommendations that can help address these challenges are shared at the end of the paper. While many of the issues and recommendations apply quite broadly, in cases where our discussion focuses on the legal and regulatory terrain surrounding neurotechnologies, we limit ourselves to the context of the United States, given space limitations.

Setting the Stage: Recognizing What Is Novel about Neurotechnology and What Is Not

There is good reason to be concerned about the development and application of neurotechnology. From perception to memories, imagination, emotions, decisions and actions, all mental or cognitive states derive directly from the activity of neural circuits in the central nervous system. Technologies that provide access to these circuits, either for recording that activity (“reading”) or altering it (“writing”) have the potential to register and alter the inner workings of human mentality. Because of this, fundamental human values including biographical

¹ After briefly using the “Morningside Group” name, the group now identifies as the Neurotechnology Ethics Taskforce (NET).

identity, agency and mental privacy could in principle be decipherable and made directly susceptible to outside influence.

At the same time, it is important to recognize that the threats posed by neurotechnology to our fragile senses of human identity, agency, and privacy are not unique or exceptional. Much of the research demonstrating this fragility employs ordinary social manipulation, like verbal communication. Consider, for example the threat to psychologically-constituted personal identity posed by neurotechnological memory-transfer. Hildt sees a grave threat to people's sense of identity in future brain-to-brain interfaces (BBI; a special type of BCIs where human brains are connected together) that implant the memories of others, such that the recipient "would not be able to distinguish between his own genuine memories and the quasi-memories" being implanted [12]. This prospect, consistent with recent research demonstrating normal behavioral responses to artificially implanted perceptions and memories in laboratory animals, is disturbing, but no more so than the prospect of implanting false memories of traumatic and transformative childhood events by verbal and visual suggestion, a feat that has already been powerfully demonstrated [13]. Much of the research that revealed humans' suggestibility in claiming agency was done before the advent of neurotechnology [14, 15].

Neurotechnology, however, promises precision and effectiveness in altering the brain and with it, human agency, identity, and privacy. Fortunately, neurotechnologies may also be more readily subject to public oversight than social manipulation: it is far more difficult to regulate the conversations that can intentionally or inadvertently alter one's memories than to alter memories deliberately with BCI. Family, friends, and society cannot help but shape an individual's values and beliefs. After all, humans are social animals embedded in an environment that requires communication and other forms of social interaction for us to thrive. As neurotechnological devices are framed and marketed to extend human cognitive, motor or other mental abilities, protecting these realms from undue interference is paramount. Still, determining what counts as excessive interference, manipulation, or undue influence is notoriously difficult. Ethical boundaries may be considerably easier to draw for neurotechnological modifications, at least in the foreseeable future; they are discrete, highly salient intrusions, for which participants can demand consent, and a careful assessment of risks and benefits. The prospect of neurotechnologies that modify, enhance, and

threaten users' sense of self, agency and privacy can "concentrate the mind," forcing society to address fundamental questions in these arenas with a sense of urgency and with the involvement of various stakeholders, including the public.

Identity and Agency

We define personal identity as the concept of self for an individual agent, whereas agency can be understood as the ability of this individual to make and communicate choices, often through action. Our argument is that while neurotechnologies may help to enable identity and agency, both these central features of human beings can also be put at risk by certain uses of neurotechnologies. Although other forms of personal intervention (e.g., education, social pressure, pharmaceuticals) may also support or manipulate these important human features, neurotechnologies present a form of intervention that aims to be more precise and effective, and that may open up greater opportunities for user manipulation without user awareness, making regulation paramount. Although identity and agency are closely related, these concerns are addressed separately.

Personal Identity

Neurotechnologies that aim to restore a person to a state that existed before the onset of an illness (e.g., DBS for Parkinson's tremors or rigidity or BCIs to restore lost sensation or movement) appear to support the preservation of identity. Yet, sometimes they create side effects that complicate the recipient's sense of identity. The price of removing Parkinsonian tremors, for instance, might be the loss of voice modulation [16], an increase in impulsive behavior [17], or confusion over one's phenomenological sense of self [18]. In such cases, the side effects may be apparent to the device user (e.g. [16]) or primarily noticed by family members or caregivers [19]. Acknowledging the capacity of neurotechnology to alter one's identity in these ways is important, even if the individual's numerical identity does not change [20] (see also [21] for discussion of issues related to potentially deeper shifts in identity). Although identity is dynamic – affected by both voluntary and involuntary changes to bodies, relationships, and social circumstances [22] – having some relatively stable characteristics is common and typically desirable.

Unwanted shifts in identity are typically perceived as harmful (although some cultural perspectives may not interpret changes in this way [23]). If neurotechnologies have the capacity to create unwanted disruptions of identity, there is reason to proceed with caution.

Neurotechnologies designed to alter psychological functioning (e.g., aimed at alleviating symptoms for depression, obsessive compulsive disorder (OCD), post-traumatic stress disorder (PTSD) or dementia) may raise concerns about identity even more directly. While the body and its functions are significant for the understanding of the individual's narrative identity, their psychological states more directly provide the interpretive frames through which their experiences are comprehended and their narratives are shaped. Changing these psychological states, then, potentially more fully transforms the narrative identity of the person.

Features sometimes ascribed to personality – impulsivity, conscientiousness, neuroticism, openness, or agreeableness – may be altered through neural interventions. DBS, for instance, may address debilitating symptoms of treatment-resistant depression, but leave individuals unsure about their own recognition of themselves (e.g., “I’ve begun to wonder what’s me and what’s the depression, and what’s the stimulator...it blurs to the point where I’m not sure, frankly, who I am.”) [24]. Brains lack proprioception and some self-monitoring of cognitive and emotive states happens at a preconscious level. People are not aware of all that their brains are doing in real-time as they are experiencing themselves thinking and acting. Consequently, they are not easily able to individuate what their brains are responsible for and what the neural device is responsible for [25]. This is especially true in cases where the device may be “smart” and receiving feedback to automatically help guide its functioning (e.g., closed loop devices). But even in the simplest implanted devices, the device may be perceived as a “third party” in the user’s head [26] that potentially competes for control rather than simply enabling control [27].

Consider neural devices designed to alter memory for the treatment of PTSD (erasing or dampening memories or perhaps implanting new ones) or dementia (retrieving or reinstating lost memory connections). While these efforts address significant problems, they simultaneously demonstrate the potential fragility of our sense of identity. Conditions such as PTSD, with symptoms sometimes tied to feelings of guilt and shame over one’s own wrongdoing (e.g. a soldier who unintentionally

kills a combatant’s civilian family), dampening the salience of such experiences could disrupt the formation of that individual’s moral convictions [28]. Developing a narrative of self requires drawing on salient personal memories and deemphasizing memories considered not adequately reflective of themselves. Although others may help to narrate their story [29], neural modulation of memories could lead to the loss of a key part of identity. Forgetting is also important to how a person navigates the world, since it allows the opportunity for both losing track of embarrassing or difficult memories, and focusing on future-oriented activity [30]. Efforts to enhance identity through memory preservation thus run the risk of inadvertently damaging a valuable, if less consciously-driven cognitive process.

Concerns about identity need not arise solely from the addition of electrical stimulation or complex patterns of electrical activity, but also from the addition of neural devices that “read” from the brain and help to control external targets (like a wheelchair, computer cursor, robotic arm, etc.). BCIs that provide the link between the brain and an external effector allow for a kind of extension of the self, beyond the boundaries of the body (the skin) and into the world [31, 32].

Although our body schemas seem quite flexible, enabling us to include everything from a fork to a cane to a car as sensory extensions (consider how it “feels” to drive on a wet road), typical experience nonetheless involves extensions that are directly connected to our bodies (sitting in the car). BCI devices, working wirelessly, will allow the capacity to directly control an action at a distance just through thinking, and perhaps also to receive sensory feedback from a distance directly to users’ brains, without the involvement of the usual sensory gateways (e.g., noses, eyes, ears, tongues, or skin). Listening in on a conversation in a distant part of the world (e.g., through radio or the internet) is already common, but neural devices would enable the capacity to send those inputs directly to the brain, perhaps making them appear as if they are much more immediate or, even in some cases, indistinguishable from local sensory impressions. This may have the effect of shifting how users think about their presence in a way that may be unsettling [33]. Where is a person when the body parts they control and receive feedback from are widely distributed? Similarly, the possibility of brain-to-brain interfaces may expand an individual’s cognitive and sensory capacities beyond their individual brain, affecting the individuality of experience (e.g., [34]). These

matters are much larger than traditional questions about informed consent and understanding of individual risk [9]; they pose a challenge to traditional conceptions of ethics and law that take a clearly bounded individual agent as their basic unit [35, 36].

Personal Agency

Most individuals take themselves to be agents, that is, individuals who *act* in the world vs. merely being acted upon (for a useful review, see [37]). Having a sense of agency involves a subjective awareness that “I am the one who is causing or generating an action” [38]. This sense of agency is a critical part of one’s ability to identify as the author of one’s actions and to take responsibility for them. A sense of agency, however, may sometimes be distinct from actual exercises of agency, in which the individual is indeed the one generating the action [39], as there could be a difference between the perception of agency and the actual reality of it.

Haselager [25] explores the implications of this potential disconnect between agency and a sense of agency for BCI devices, especially where the devices are “intelligent” and the shift to device control may be implicit (not directly user controlled). A user may think they are authoring an action and simply using the BCI to enact it (e.g., reaching with a BCI-controlled robotic arm) when in fact the BCI device is “smartly” operating on its own, based on its visual inputs and artificial intelligence that accurately predicts what the individual wants. If the agent’s intent and the device’s output can come apart (think of how the auto-correct function in texting sometimes misinterprets the user’s intent and sends problematic text messages), the user’s sense of agency may be undermined. They may find their wheelchair or body moving in ways they did not intend, and thus feel controlled, or they may question their own intent. BCI “users may be insecure about the extent to which the resulting behavior, whether successful or unsuccessful, is genuinely their own. Though they may be certain about what they wanted, they may be insecure about the extent to which they did or did not ‘do’ it” [25]. Neural devices that target motivational centers in the brain that modulate desires (e.g., devices intended to treat depression, anorexia, obesity, etc.) may also complicate a user’s sense of certainty about *what* they want.

Ironically, many neurotechnologies are explicitly designed with the goal of enabling agency. For instance, a

BCI-operated wheelchair that responds to thought-initiated commands from a user might provide a quadriplegic person with the capacity to initiate a causal sequence that results in them moving across the room, enabling them to control their mobility. Processors and machinery mediate the action, but the device restores a sense of agency to the user, given the newly introduced connection between thought and action.² Wolpaw and colleagues describe how one of their research participants used a BCI device designed to facilitate typing and communication so that he could continue running his lab, despite significant motor impairments due to ALS [40]. For practical purposes, this looks like a technological restoration of agency. In addition, future BCI devices will likely aim not simply to control external devices (e.g. robotic arms, wheelchairs, computer cursors), but also to “reanimate” paralyzed limbs [41]. In such instances, the BCI extracts neural signaling indicative of particular motor intentions and wirelessly transfers it past a spinal cord lesion directly to the peripheral nerves or neuromuscular junctions, thus even more closely approximating typical electrical signaling involved in movement.

Given that the sense of agency can be enabled but also manipulated or confused, and that neurotechnologies operate at the nexus of typical agency (intervening between intention and action), neurotechnologies seem positioned to alter human’s sense of agency.

Brain Data Privacy

Another key area of concern about neurotechnologies is privacy. Data privacy is a general problem resulting from technological access to personal data via electronic devices such as smartphones, but one which is greatly exacerbated with neurotechnology, since the data it generates and manipulates reflects the neural activity of the individual. While definitions of privacy are contested, we understand privacy as a *right that others not access one’s personal information and personal*

² Some action theorists might dispute whether this counts as an “action” given that the wheelchair is external to the body. Action theorists debate, for instance, whether the action “she moved her left arm” is applicable only to the person who has direct control over her body (e.g., raising her left arm) or also to the person who more indirectly shifts her body (e.g., by using a non-paralyzed right arm to lift her paralyzed left arm). The left arm is moved in both cases, and the agent intends them in both cases, but the pathway of control is different.

space. Three features of privacy in relation to neurotechnology are: the intimate nature of brain data, the general trend of increased intrusions into privacy via technology, and the relative inaccessibility of the brain to privacy intrusions without neurotechnologies.

First, brain data, or neural data (understood as any data recorded from the activity of brain tissue), could provide access to highly intimate information that is proximal to one's identity. Such data is sensitive because it contains information about the organ that generates the mind [42, 43]. Though not all neural data is decipherable, some of it can be "read" or interpreted. Imagined handwriting, for instance, can be decoded and translated by BCI into quick and accurate texting [44]. Similarly, BCIs that identify and convert covert speech could be used to drive a computerized voice [45] and so externalize what was previously private. Information derived from neurotechnological means has not passed through executively controlled sensory or motor systems, meaning that it potentially lacks the mechanisms by which people normally control information they convey to the world. One can often control, to some extent, what one says, one's facial expressions, and other ways in which one behaviorally presents oneself to the world. While some inadvertent action may unintentionally reveal certain information, collecting brain data may provide new avenues to circumvent even this limited control. This information can include intentions and emotions. Hiding unsanctioned emotions is a common strategy for maintaining privacy that may be threatened. Moreover, such access may reveal facts that affect how one sees oneself, for example, by revealing subconscious tendencies and biases. Furthermore, brain data may contain information on brain pathology, for example, epileptiform EEG patterns, that might be revealed without explicit consent. Intimate aspects of the individual that are personal and otherwise relatively hidden from others may become accessible through neurotechnology.

Second, pressure on individual privacy has increased dramatically in recent years, with a vast expansion of government surveillance, not only in the United States (U.S.), but also internationally [46]. Such tradeoffs between security and civil liberties – relevant now in the context of COVID-19 pandemic control via digital tracking – can conform to the model of 'securitization', i.e. reframing social issues as security issues, often with the aim to solidify state power [47]. For-profit enterprises have long sought information about individuals

for a variety of purposes, often pertaining to marketing and commerce. Today, the commercial sector's access to individuals' information is at unprecedented levels. Social media companies have the capability to distinguish their users' social, political, religious, and consumer affiliations, and use or exchange that information to enable selectively targeted information dissemination [48], while many are creating medical data from social media content without users' awareness [49]. Hence, questions about brain data privacy arise against a background of diminished privacy in other contexts. With so many aspects of people's private life accessible through data, brain data poses the final frontier to directly access still more intimate data, which could profoundly deepen already robust personal data profiles.

Third, brain data may be one of the few remaining domains in which the most substantial invasions of privacy have not yet been realized. It may be too late to restrict the acquisition of location data/video surveillance, commercial preferences, and behavioral data, but devices to permit ubiquitous brain recording do not yet exist. This is subject to change however, particularly because of the large-scale commercial capital investment pouring into the development of consumer neurotechnologies. Over 2019 alone, Microsoft invested 1 billion into OpenAI, a company cofounded by Elon Musk to build artificial general intelligence [50], while Musk announced progress in another company he founded, Neuralink, to combine AI with invasive BCIs to augment human brains. Meanwhile, Facebook purchased CTRL-Labs, a company specializing in non-invasive BCIs, for between 500 million and 1 billion dollars [51]. Consequently, brain data privacy is important not only for the reasons we have noted above, but because the establishment of regimes to protect brain data may be one of the few remaining bulwarks against fully compromising privacy in modern life. The window for implementing such measures proactively, as opposed to reactively, is likely to shrink as neurotechnology investment grows [52].

Combining these three foundational issues with the ability to link brain data with other types of personal data, makes brain data especially powerful. Several specific brain data privacy concerns are addressed below.

Unauthorized Access

Brain data could be stolen or released accidentally, and thereby made accessible to unauthorized parties. The

concept of BCI “App Stores” has been implemented by some neurotechnology companies including Emotiv and NeuroSky to expand BCI applications. Most of the applications included are granted unrestricted access to users’ raw electroencephalogram (EEG) signals [53]. In 2012, Martinovic presented “brain spyware,” which can extract confidential information about an individual via a BCI-enabled malicious application [54]. Preventing or deterring efforts to gain unauthorized access to the information contained in brain data and building effective safeguards against accidental data release should be a priority [55–57]. Effective policies will require a diverse set of approaches tailored to the circumstances and method of collection, as well as the type of data and format of storage.

Mind “Reading” and Consent to Share Brain Data

Another series of concerns involves individual consent to collection and use of brain data. Personal information is generally regulated on an individual consent model. If a person agrees to share their information (even when the agreement is based on language buried deep within an end-user license agreement or EULA, or when a person has no realistic opportunity to understand how their data will be used), then the party to the agreement can generally collect, combine, use, and transfer that information per the terms of that consent. As such, consent is a vehicle for determining which data an individual authorizes to be collected and which are off limits. Consent agreements also determine what happens after data collection. Individuals may unknowingly consent to share data that will grant companies further, and perhaps unwanted, insight into their personal customer profiles. The recent field of “Neuromarketing” has revealed significant new data about human preferences and emotional responses by measuring the brain activation of customers using magnetoencephalography (MEG) and wearable EEG. This data could be used to predict future consumers’ choices and therefore could have high resale value [58]. If it had been authorized by a consent agreement, such personal insight could then be made available to any entity willing to pay for it. In other instances, users might feel comfortable authorizing companies to collect their brain data for certain purposes (i.e., product improvement), but not authorize that same company to sell their data to another company for targeted marketing or use the data for a purpose to

which they object (e.g., a company that produces both commercial and military products).

Obstacles to meaningful and specific consent are a problem in the collection and sharing of brain data. Brain data is a complicated concept which is difficult to communicate to a broad population. The same is true regarding the vast network of commercially available personal data that can be used to make further inferences about an individual’s brain data. Given the relatively early stage of neurotechnological development, it is difficult to predict the future uses and risks of collecting brain data. One reasonable criticism is that consent will be perfunctory and fail to be meaningful because it is impossible for most subjects to understand enough about possible risks to be adequately informed. Consider how rarely people read EULAs before buying or using products. One study [59] found that over 80% of their participants either reported “not reading the EULA at all” or “not really reading anything.” Of the remaining 20%, 16.5% described their reading behavior as “skimming.” While the Common Rule requires research overseen by an Institutional Review Board (IRB) to outline key information in an accessible manner in informed consent documents, similar requirements are not as forthcoming from commercial entities. These problems are intractable if the individual consent model remains as is, with companies able to make commercial use of information with sparse, difficult-to-understand consent procedures, or the employment of opt-out paradigms.

“Writing” and Opt-in Consent for Brain Data

Privacy is important not just because information can be gleaned from brain data but because neurotechnologies allow for new ways of “writing” information into the brain. While “writing” is a metaphorical term to describe the many ways that electrical activity can be precisely delivered to the brain to specify particular outcomes, neuroscientists are increasingly able to stimulate the brains of animals to create, for instance, behavioral responses that suggest a specific visual experience was delivered even in the absence of actual visual content, i.e., a hallucination of sorts can be “written” into the brain [60]. Importantly for our discussion, the behavior of animals is identical whether they are optogenetically implanted with a visual image or they see the image with their eyes. This indicates that neurotechnological manipulations in humans may be interpreted as part of the self. Consider two brain processes central to human

experience and human identity: processing of fear and formation of memory. What a person fears, when they fear, and how they respond to fear all shape identity. Something similar can be said for memory. Neurotechnologies offer the prospect of making changes to these and other brain processes by encoding new information into the brain. A carefully placed electrode may induce a feeling of fear or a memory (or feeling of *déjà vu*) [61] that is disconnected from the typical ways in which the subject experiences changes to their mental states (e.g., watching a scary movie, seeing an old photograph). Technologically-mediated changes to mental life can be psychologically disruptive or alienating. As such, robust consent to undergo such changes ought to be a prerequisite.

Individual Consent and Collective Action

As noted, personal information is generally regulated on an individual consent model, which itself is often not implemented well. However, one does not need to share much information about oneself for others to be able to make important inferences. Kosinski et al. [62] used easily accessible digital behaviors (Facebook “likes”) to accurately predict a “range of highly sensitive personal attributes” such as sexual orientation, religious and political views, and personality traits, while others have demonstrated the ease with which social media data can be used to reveal which users are likely to be diagnosed with ADHD or depression [63]. Thus, one could infer the intimate from the available data. Furthermore, by analyzing the data that *others* share, companies can make inferences based on a few similarities. In other words, relying on information from people who consent to information gathering allows for fine-grained inferences *even about non-consenters*. Consider, for example, consumer EEG devices used for gaming or biofeedback. If, in the process of providing services to consumers, companies collect neural data from these devices, this data may be useful for making inferences about non-users of their devices (e.g., emotional reactions to in-game purchase options among a particular socioeconomic demographic). This is not dissimilar to the way that media companies (Netflix, Amazon) can take a small amount of information about a person (a few movies or shows that one likes) and infer much about one’s preferences based on the tastes of others. Likewise, social media companies (Facebook in particular) can narrowly tailor and push information to suit

individuals based on their activities within the platform (i.e., what one posts, ‘likes,’ comments on, and so forth). While the line from non-neural sources of data to inference may be more obvious than that for neural data, this may make the individual consent model all that more difficult to apply.

Power Disparities

A related set of issues concerns the substantial power disparities between individuals, whose brain data may be obtained, and the larger entities that may wish to collect, analyze, use, and share that information. This issue concerns whether – even in the context of clear consent procedures – individuals are empowered to exercise their options to refuse.

Many employees sign agreements to substantial surveillance within (and even beyond) their work environments. Employers regularly surveil communications, health habits, and other behaviors. It may well be justifiable for institutions to collect some of this information, but it is not always clear how easy or viable it is to refuse consent. When sensitive or intimate data is commonly recorded, a person contemplating refusal may reasonably worry that doing so could cost them their job. In the context of collecting brain data from consumer technologies, imagine a wearable EEG system for closed-loop interaction with software content on social media (as pursued in Facebook’s “brain typing” project, see [64]. Vulnerable individuals such as teens or others may come under pressure for fear of social alienation or exclusion from peer groups when not participating in these “services.” Some research suggests that for users of social media, the actual (or perceived) psychological rewards for using the services often seem to outweigh the possible threats to privacy [65]. Still, in a study comparing levels of concern over privacy issues online in 2002 and then in 2008, expressed concerns rose substantially [57]. This may have been due to increases in rates of fraud and identity theft, and breaking news on large scale data breaches.

Bias

Broadly, bias can occur when scientific or technological decisions are based on limited data, methodologies, values, or concepts. Bias is inherent to most human endeavors – each person has a particular perspective

from which they understand the world. It can, however, become problematic if we are not vigilant of its effects. Without careful consideration, bias can exclude, oppress, or denigrate alternative perspectives, usually those of minority or vulnerable populations. Bias can have significant impacts on what we know about the world. For instance, bias can influence which populations are included in research and are therefore likely to benefit from it (e.g. receiving effective treatments), and what sort of research questions are studied and therefore what “truths” are found [66, 67]. Similarly, existing social biases can be reinscribed through the design of technologies that fail to identify problematic assumptions. As one critic puts it, “research practices can reconstitute fixed understandings of difference. Therefore, researchers must excavate how ‘optimally working’ technological practices insidiously encode normative ideas about racial worth without need for a specific racist intent.” [68].

More specific examples of bias relate to projected views of what is considered “normal” brain function and what is not. Feedback loops between commercial entities, funding opportunities, and research trends can perpetuate assumptions about how a “normal brain” should function, and what count as desirable and acceptable behaviors. Assumptions regarding the importance of particular research targets or outcome measures may wrongly be viewed as universally shared [see [10]]. One example that has come strongly to the fore is the rejection of research that aims to “cure” autism by neurodiversity advocates [69]. Similarly, many have questioned research that aims to describe differences between “male” and “female” brains [70]. These types of studies remind us what assumptions drive the scope and content of research programs. They lead us to ask how efforts to reveal how the brain works can proceed with a sensitivity to such value-laden assumptions [71].

Below, we illustrate how biases can have an impact on neurotechnologies at various stages of their development and distribution: research goals and questions, participant/data set selection, dissemination, and assessment and feedback.

Examples of Biases within Neurotechnologies

Research Goals and Questions Research goals are often shaped by trends within a field, funding mechanisms [72], pressures related to job security [73], shared community or cultural norms, publication biases [74, 75],

and conflicts of interest [76]. Additionally, mainstream biases – such as ableism – can influence the direction of research. Given the broad social acceptance of interventions aimed at assisting those with a medical condition, for instance, study aims may medicalize conditions that their bearers do not view as detrimental.

Cochlear implants were developed to restore hearing, but with the problematic assumption that being deaf is understood as a biological deficit to be fixed [77]. The limitation of this bias is evident in responses from many members of the Deaf community, who do not view deafness as a deficit, and are not interested in eliminating deafness [78]. Rather, they communicate through sign language and view being deaf as a valuable embodied experience [79]. Although not all deaf individuals view deafness in this way, scientists aiming to “help” the deaf community should be aware of how deaf individuals may understand and value their condition, and work in ways sensitive to those values. More generally, including the perspectives of likely end-users of neurotechnology throughout the research and development process, including in the early stages of setting research goals and questions, would help to promote a just and well-targeted product [80]. In a variety of areas of clinical research, engagement with study participants has been shown to introduce perspectives that have not been considered [81] and better translation of research outcomes [82].

Participant/ Data Set Selection In participant or data selection, it is often difficult to capture a sample that is sufficiently diverse and inclusive. Bias introduced in this phase of a research project can have a significant impact on the generalizability of the conclusions drawn from the data [83]. For example, clinical trials for medical devices have historically enrolled predominantly white test subjects [84]. A study on novel medical devices performed from 2014 to 2017 found that despite laws being passed in the United States to increase test group diversity in 1993 [85] and 2012 [86], diversity and racial/ethnic subgroup testing has remained low [87]. As a result, novel medical devices, which include some BCIs, may be less well tested for safety among demographics with subtly varying medical characteristics.

Dissemination Medical and consumer devices offer people ways to “improve themselves” – to be happier, smarter, more agile, and more alert. This reinforces the idea that a person living well is a person who is happy,

smart, agile, and alert, increasing social pressure to strive to exhibit these traits. The dissemination of research results and marketing of products can exacerbate these social pressures, regardless of the accessibility or effectiveness of these devices. For example, Transcranial Direct Current Stimulation (tDCS) has been marketed by consumer companies as a way to modulate mood and make users variously calm, alert, or energized [88]. While it is not clear that these devices are effective for these purposes, the chosen target moods or mental states reflect value assumptions about which moods are preferable. Such a value assumption subliminally or explicitly prompts users to modulate their energy and temperament levels so that they might adhere to the social standard marketed by the seller (i.e., agreeability, engagement), rather than users honoring their authentic subjective experience and the circumstances that may have induced them. People should be encouraged to acknowledge that feeling frustrated or anxious is often a normal and healthy reaction in response to specific circumstances and one should not always attempt to artificially exchange that mode for a more positive one.

Assessment and Feedback Bias can also arise in the assessment phase of neurotechnology development. Researchers and industry groups often use assessment tools to determine who should have access to a new technology, when it should be made available, and how to measure successful implementation. In such situations, a medical device may be described as efficacious based on the company's assessment tool, yet this tool may not adequately capture patient concerns about the device. Conflicts of interest can play a significant role here as well. As Eaton and Illes point out [89], "In the context of combined assessment and post-assessment treatment services, companies face an inherent self-interest that can affect their business decisions where the frequency of diagnosis directly promotes growth in the treatment or service arm of their business." Companies also routinely use market research and feedback from particular users to guide their design of neurotechnology devices. Whether specific instances of feedback are then reflected in device design depends on how valuable the feedback is perceived to be – an assessment that itself depends on the priorities and biases of the company.

Even though countless strategies have been developed to help minimize or counteract inappropriate biases within research (e.g. conflict of interest policies,

double-blinding, data safety monitoring boards), the examples above illustrate how biases can shape the development of neurotechnologies in stages that are not covered by existing strategies or regulations [90]. The responsibility for recognizing and responding to biases in the field of neurotechnology falls on all who are involved in their development and use: scientists, clinicians, industry, funders, regulatory bodies, journals, consumers, and the media.

Enhancement

Enhancement (or augmentation) interventions are those that "improve human form or functioning beyond what is necessary to sustain or restore good health" [91]. By contrast, treatment interventions generally refer to those that restore an individual to a "healthy" state. However, clear lines between enhancement and treatment are difficult to elucidate [92], as what is considered "healthy" or "normal" varies across social, cultural, and temporal contexts. A given intervention may be interpreted differently depending on the circumstances of use. A consideration of the literature on cognitive enhancement may help to elucidate themes of relevance for a wider array of neurotechnological enhancements (including, e.g., emotional, prosocial, or physical enhancements).

Contemporary discussions of the ethics of cognitive enhancement began with the advent of pharmacological drugs—primarily stimulants such as methylphenidate and amphetamine derivatives—previously used for the treatment of depression and now commonly used for the treatment of indications such as ADHD [93]. Although these drugs now require a prescription, in practice they are often obtained illicitly by those looking to improve learning, memory and concentration [94]. Scholars have raised numerous ethical concerns, relating to authenticity [95–97], fairness [98, 99], and disruptions to personhood [100, 101], among others [102].

In recent years, the discussion of the ethics of cognitive enhancement has shifted from pharmacological enhancers to neurotechnological techniques, such as neurostimulation, that can modulate brain function. For example, DBS can have indirect effects on identity and personality [103, 104]. Other external (or "noninvasive") techniques, such as transcranial direct current stimulation (tDCS), have been shown to have both therapeutic and cognitive enhancement effects [105–107]. In addition, advances in robotics, AI, and

BCI interface technology have led to speculation about the ethical issues that may arise in a future world of unprecedented human intelligence and cognitive capabilities, as human brains connect more directly with the impressive power of machine learning and vast data available via the internet. The literature on cognitive enhancement via neurotechnological modification has largely paralleled the earlier literature on pharmacological enhancement, with debates arising about whether there are meaningful differences between various methods of enhancement (neurotechnological, pharmacological, and others). Empirical work has examined attitudes towards enhancement with neurostimulation [108, 109], how such technologies are used recreationally [110, 111], and the ethical and regulatory issues raised by direct-to-consumer marketing of enhancement products [112–114].

Neurotechnologies may alter not only cognitive processing but also emotional regulation, social skills, and even physical capacities. A neuroprosthetic device could potentially dampen or raise mood, provide access to facial recognition or name recall, or allow a person to exert superhuman strength via a thought-controlled robotic arm. There is already anecdotal evidence of clinical cases where human abilities are significantly altered by neuroprosthetic devices and where patients feel that the device is part of their bodies (M. Nicolelis, pers. comm.). Given the influx of industry funding targeting consumer uses of the new devices, the reality is that neurotechnologies are already and increasingly will be designed explicitly for enhancement purposes. Salient ethical concerns relate to safety, commercial responsibility, social coercion, distributive justice, and unintended/dual uses.

Safety (Short-Term and Long-Term Effects)

Safety is often defined in relation to the probability of an adverse event—a short-term, quantifiable, health-related effect – and in the neurotechnology space, encompasses everything from mild symptoms (e.g., skin tingling) to medically significant events (e.g., seizures) and death [115]. In most countries, government regulation requires that medical devices and drugs marketed for medical purposes demonstrate a minimum level of safety. However, products marketed solely for enhancement may not be required to comply with drug and medical device regulations. Thus, it is unclear which government agencies, if any, will maintain oversight

over the safety of enhancement products [111, 116]. Additionally, even beyond typical health-related adverse events, neuroscientists have speculated that brain enhancement may be a “zero sum” endeavor—that is, enhancement of one cognitive ability may come at the cost of others [117, 118].

To some degree, the characterization of safety as a measure of the probability of near-term adverse events has obscured attention to the potential long-term risks of enhancement interventions. As at least one study has shown that a subset of recreational users of neurostimulation devices marketed for enhancement utilize the device much more frequently than in scientific protocols [88], the issue of safety with regard to chronic use may be of particular concern. Furthermore, there are questions regarding the long-term effects of implantable neurological devices. While current research focuses on mitigating foreign body responses and neuroinflammatory reactions in the short-term [119], such devices may interact with brain tissue in additional unknown ways or cause unforeseen health problems many years after implantation. Regulation focused on the long-term effects of emerging technologies, such as neurotechnologies for brain stimulation, often requires decision-making under considerable uncertainty, given the lack of longitudinal observations and data. In this context, a precautionary approach might be most appropriate [120].

Commercial Responsibility

As noted above, some neurotechnologies may not fit into traditional medical regulatory frameworks; similarly, research conducted on them may also fall outside the scope of federal research regulations. This is a particularly acute problem for neurotechnologies built for human enhancement, given the already great commercial interest that exists in the consumer space (e.g., Neuralink or Kernel). In the case of the U.S., while the Belmont report and subsequent U.S. federal regulations (known as the “Common Rule”) set out principles for the protection of human subjects in research contexts, such regulations apply only to research that is being conducted with federal funding [121]. Companies conducting neurotechnology research on human subjects using private funds are not required to comply with the Common Rule. Particularly in cases where neurotechnologies are noninvasive, no Food and Drug Administration (FDA) or clinical approval would be required for use, resulting in high degrees of freedom

within the direct-to-consumer noninvasive neurotechnology market [122]. This raises concerns, as companies may have competing interests—such as the financial interests of their investors—that conflict with the goal of ensuring the safety of participants or the public good.

Social Coercion

Scholars have raised concerns about the potential for coercion in the use of neurotechnologies. For example, if a neurotechnology becomes widespread in educational, military or occupational contexts, individuals may feel compelled to adopt such technologies, either explicitly, via regulations and policies, or implicitly, via social pressure. In the domain of sports, many athletes have felt peer pressure to use performance-enhancing drugs, despite their illegality, in order to remain competitive [123]. Other scholars have pointed out that coercion may only be a practical concern if a given neurotechnology is both effective and has demonstrated a propensity for widespread social uptake [124]. While to date, no contemporary pharmacological or neurotechnological enhancement intervention has achieved widespread social uptake, coercion may be a potential future concern given the rapid development of neurotechnological tools.

Distributive Justice

If neurotechnologies are disproportionately available to those in higher socioeconomic classes, they may exacerbate current gaps in inequality. Indeed, there is evidence that those who purchase direct-to-consumer neurostimulation devices are in much higher income brackets relative to the general U.S. population [88]. Other avenues for inequitable distributions of neurotechnologies could come from the workplace (i.e., some companies, but not others, might be motivated to provide enhancement technologies for their employees). While issues of inequality with regard to neurotechnologies may not differ in principle from non-neurotechnological enhancements (e.g., better educations are available to those with greater means), equitable and fair distribution of these techniques still represents a potential concern. Different neurotechnologies will likely be developed and disseminated at varying rates; implanted technologies that require surgical placement of high-tech devices may begin as expensive and

relatively inaccessible, whereas non-implanted technologies that use relatively simple hardware may be inexpensive and readily available, and therefore more widely accessible to a greater number of individuals [125].

Unintended/Dual Uses

Though many neurotechnologies are intended for positive purposes—such as assistive BCIs for people with disabilities—such interventions could be used by malicious actors for harmful purposes [126]. For example, criminals could hack into individuals' BCIs, or prisoners of war could be subjected to unwanted neural recording or neurostimulation. The capacity to act directly on the brains of one's enemies sets up a potentially problematic kind of control. Though nefarious uses are in principle possible with any new technology, caution is particularly warranted with relation to interventions that can modify brain function [126–128], given the variety of concerning reasons articulated in the earlier sections of the paper. Given the concerns outlined above, we suggest a number of guiding principles to ensure the safe, appropriate and fair development of neurotechnologies. These recommendations echo those made in the identity section.

Recommendations

Taken together, these four areas of concern related to neurotechnologies demand attention and action. The ten recommendations offered below articulate and briefly explain precautionary measures for the responsible development and application of neurotechnologies.

- **Recommendation 1: Building on existing human rights frameworks, establish “Neurorights” (e.g., mental liberty, mental privacy and mental integrity)** Together with others [129], we recognize that people may soon require explicitly stated rights to keep their internal mental space free of unwanted recording and manipulation. Such rights would not mean that someone could act with impunity in defining their identity or exercising their agency but they would protect individuals from unwanted intrusion. Similarly, to preserve individual privacy and individual power to control access to their own intimate mental spaces, we recommend that all entities engaged in collection, analysis, use, and sharing of

brain data recognize several baseline rights that individuals have with respect to their data. First, people have a strict right to not be compelled to have brain data or code written into them. Second, people have a strict right to not be compelled to give up brain data. Third, people have a right to the restriction of the commercial transfer and use of their brain data, such that commercial reading and writing of brain data is prohibited (regardless of consent status), depending on what information is contained in (or could be inferred from) that data. These rights would be a conceptual re-thinking [32] of already recognized rights (freedom of thought, bodily integrity) in response to emerging technological opportunities to directly record and manipulate the brain.

- **Recommendation 2: Improve informed consent for neurotechnology.** We recommend that users of neurotechnology, including research participants, must be fully informed about potential psychosocial side effects in advance of device adoption, with attention to ensuring that individuals comprehend short and long-term risks. Given the current limited understanding of what the long-term risks are, this will require funding and completing research to study long-term effects of neurotechnology use. In addition, we recommend that consent procedure tools be improved, using plain, simple and comprehensible wording with complementary aids such as visualization where possible (see for example [130]). Still, the consent process should be specific. What information will be recorded from or “written” into the brain, who will do this, for how long, and for what purpose? What are the relevant risks? When and how can an individual revoke initial consent or stop the “reading” or “writing” process and ensure that access is secured? In relation to neurotechnology, revisiting questions and revising consent over time should become the norm, not the exception. In the commercial sector, transparent end-user licensing agreements (EULAs) are helpful but not sufficient in themselves. To ensure better comprehension, widespread efforts to increase public understanding of machine learning and big data, including what insights they provide as well as their limitations, must be undertaken. Although data literacy efforts have already begun in many parts of the world [52], future efforts should ensure that brain data is included as an emerging class of sensitive personal data [39].
- **Recommendation 3: Create defaults that require an active opt-in to share brain data.** We recommend the default stance toward any collection of brain data mandate explicit “opt-in” authorization. That is, brain data should not be collected passively or rely on individuals to “opt-out” if they do not wish their data to be collected. Rather, the default should require data collectors to obtain specific consent for not just data collection, but for how data will be used, for what purpose, and for how long. Greater granularity in consent options gives the individual a broader axis of control, even if it creates a greater burden on participation. A higher level of protection should help to signal the potential salience of these data [131, 132]. There may be instances where imminent public safety concerns supersede this default, but such instances would require explicit attention and would need to follow established legal proceedings.
- **Recommendation 4: Encrypt brain data along its full arc, from brain recording site to output device.** We recommend full encryption [133] to help protect privacy. For example by using homomorphic encryption, data can undergo analysis while remaining encrypted [134]. It should be collected and stored in open data formats using open-source code but with objective and verifiable block chain tracking (or equivalent). This will provide both deterrence and a mechanism for assigning responsibility when unauthorized access does occur. Priority should be given to brain data processing that uses encapsulated modules located in close proximity to the brain recording site (for example, on the local recording device). The collection, storage, and use of brain data should occur on verifiable hardware. Within any research organization or company that collects, stores and/or processes brain data, access to the data should be strictly limited for pre-specified purposes, each instance of access by individuals (whether researchers or employees) should be logged, and sensible guidelines on the duration of data storage should be developed. Finally, brain data should be governed by a principle of succinctness—filtering-out and thereby transmitting only the minimum data needed at each stage along the data arc.
- **Recommendation 5: Restrict sharing of brain data (given re-identification risks) and concerns about rise of commercial markets.** We recommend

substantial restrictions on commercial sharing or sale of brain data. Similar restrictions exist in other contexts. For example, the commercial use of human tissue and organs is tightly regulated in the U.S. and elsewhere, with the sale of organs strictly prohibited (though with admitted ambiguity in permissible processing costs [135]). These regulations serve to protect bodily integrity and avoid exploitation by eliminating monetary incentives. Similar protections may be needed to protect mental privacy and integrity and avoid exploitation of brain data for commercial purposes. Similarly, health information is subject to important limitations on sharing for Health Insurance Portability and Accountability Act (HIPAA)-covered entities, through these limitations are not without their gaps [136]. This option maybe made more viable if brain data is considered a form of medical data, as was found in a recent survey of neuroscientists [137]. Medical data are protected to prevent discrimination, maintain confidentiality and trust, and ensure that the individual exerts control over what parts of the intimate are shared; brain data – even when not recorded in a medical setting – may require similar protections. The temptation to commercially exploit uncontrolled and potentially powerful brain data is too great to leave it to a consent regime that does little to protect individual information.

- **Recommendation 6: Recognize Bias.** The assumptions, values, and limitations underlying research, whether they are appropriate or inappropriate, intentional or unintentional, should be acknowledged throughout the process of research and development [138]. Discussion should take place regularly regarding the role of biases related to the sample selected, the conceptions of well-being and quality of life being relied upon, the pressures from funders or industry, and so on. It is crucial not only to recognize the role of bias within a research team or company, but also to communicate to others how these biases affect an intervention or product. This communication might be done within direct-to-consumer advertising (DTCA), in peer-reviewed publications, or in the popular press. Journals can play an important role by requiring that research submissions both recognize and respond to the biases that have shaped their findings. Finally, regulators can help to fill gaps related to DTCA in neuroproducts [139]. For instance, in 2019 New York insurance regulators explicitly cited UnitedHealth Group for a racist algorithm that was directing black patients away from higher quality health care in favor of white patients. In this case, regulators not only identified and raised awareness about bias, but took action to ensure the biased tool was corrected or abandoned.
- **Recommendation 7: Actively Counteract Bias.** It is important to engage with communities that may be affected by research and obtain constructive feedback [80]. Researchers and industry groups could, for instance, seek end user feedback through focus groups and/or surveys that ask questions about bias regarding the device they are designing. Bias can also be counteracted through the selection of research participants who are diverse with respect to income, gender, and race. Research by social scientists can offer insight about potential end users that could help minimize the effect of problematic biases in research. To this end, psychological research on effective structural and individual strategies for debiasing should also be considered [140]. Academic researchers and companies may also want to consider implementing a “bias checklist” or “unconscious bias training” in the process of device development. Finally, an important part of actively counteracting bias is diversifying research teams. Research and consulting teams comprised of members with diverse sets of backgrounds, disciplines, social identities, and training will be more likely to identify alternative ways to approach a shared problem, and to recognize new issues that might otherwise be overlooked [141]. Inclusion of women, people of color, disabled people, etc. in neuroscience research and neurotechnology development teams will help to surface implicit assumptions about neural differences and their evaluation and significance and will help to ensure that training sets and their supervision strategies are inclusive. We support ongoing efforts to recognize and address diversity deficits in science (e.g. Gordon Research Conference “Power Hour” [142]).
- **Recommendation 8: Encourage commercial responsibility in the development of neurotechnologies** Scientists and ethicists should work alongside companies to ensure that neurotechnologies are

developed with appropriate ethical foresight. Where possible, companies that comply with ethically responsible standards should be recognized for their efforts. IBM for example is working to create a crowdsourced, iterative framework for ethical AI called “Everyday Ethics for AI” which they are committed to upholding within their work. Francesca Rossi, IBM’s AI Ethics Global Leader, participated in the creation of the European Union’s ethics guidelines for AI [143].

- Recommendation 9: Promote equitable access to neurotechnologies. While we recognize that many forms of human cognitive augmentation are still in the distant future, early efforts should be made by manufacturers and insurance companies to ensure broad access to effective enhancement technologies. Companies should consider innovative pricing arrangements that allow for effective products to be subsidized and made available to people with lower incomes who desire them. Given the complexity and range of possible enhancements, this recommendation serves as a reminder to keep matters of equity and access centered as development continues.
- Recommendation 10: Create a broad international commission designed to meet regularly and assess neurotechnology developments with the aim of providing ethical guidance and shared commitments to responsible innovation. We recommend the establishment of a transparent international commission to examine how neurotechnology research should be structured, regulated, and facilitated. While that the development of neurotechnology has enormous medical and scientific benefits, this development should be done with an eye to broader social ramifications, not simply individual consent. Similar efforts have been pursued in relation to the ethical and societal implications of human gene editing [144], with the recommendation for an ongoing international forum with widespread representation from science and technology, ethics, law, community leaders, health care providers, funders, among others. to ensure continued discussion, shared responsibility and proactive governance. We recognize that this model has limitations: many such meetings only include a narrow set of relatively privileged perspectives, often from people already invested in the development of the technology in

question [145]. It is critical that these meetings include voices from those who are directly affected by the application of the technology, including consumers, patients, caregivers, and the public. Intentional democratic engagement of the public is necessary in a scientific field that has the potential to impact everyone [146]. We stand on the verge of a transformative shift in how humans experience the world, and we would do well to collectively explore the likely implications of neurotechnology before we make the leap. Developing international, democratic, inclusive efforts to assess transformative technologies is likely to be “imperfect, slow, difficult and expensive” [147]. However, it is often better than the alternative. To be clear, this recommendation is not intended to replace but rather to complement and expand a broader system of responsible research and innovation (RRI) [148], such as that adopted by the European Human Brain Project [2] or the more principle-based approach recommended by the NIH BRAIN Initiative ethics working group [9]. On the RRI model, science is understood to be done “for and with” society rather than simply “in” society, and it treats ethics and values as integral to science rather than as a constraint on science. The international effort envisioned here would aspire to provide a greater democratic accountability [33] and would aim to foster global efforts to achieve moral consensus, again, akin both to human rights and to international research agreements.

Conclusion: A Call to Action

In summary, neurotechnologies have the potential to significantly alter elements of the human experience. BCIs, particularly those utilizing artificial intelligence, can expand or disrupt senses of identity and agency for users. We recommend that users be given access to education about the potential psychosocial impacts of BCI use and that the collective international public, scientific, political, medical and corporate communities participate in an inclusive conversation about the elements of the human experience that should be preserved within this domain.

Neurotechnology also offers unique access to some of the most intimate data we have to offer: brain data.

What happens inside people's heads is the last bastion of privacy remaining. Maximum privacy and security efforts must be employed to protect this data from being accessed illicitly. The regulation of sharing such data should go beyond the individual consent model that is currently employed – even if not always as intended – for much personal data. Where consent models are still used, we should seek to improve the digestibility of consent agreement documents while improving data literacy.

Bias, often implicit within the research, development and application of BCI devices, is inevitable and must be exposed, acknowledged, and mitigated wherever possible. Preventative measures should be exhaustively pursued to prevent problematic biases from influencing neurotechnology development in ways that re-inscribe and exacerbate existing inequalities.

Finally, neurotechnology-enabled cognitive enhancement must be scrutinized from the perspectives of short and long-term safety and distributive justice. Exploration into adverse and unexpected impacts should be thoroughly explored before introducing such potentially disruptive technologies, particularly within commercial markets.

We have offered recommendations related to each of these issues, detailing particular concerns that lead to the need for ethical sensitivity and guidance. Within each of these realms, regulators, researchers, and companies should prioritize working with and for society, taking on the responsibility to ensure transparency and responsible leadership within the development and application of such technologies. We urge that citizens be empowered to take advantage of available information to learn about novel neurotechnologies, critically consider the potential societal impacts of such technologies, and demand from their political representatives a clear and public stance on this issue.

Acknowledgements The authors would like to thank Adrian Carter and Judy Illes for their helpful contributions to earlier drafts of the paper. Special thanks to Belle Saar Lively and Emily Kate Abrams Einhorn for assistance with the manuscript preparation. Supported by NSF DBI 1644405 (“Coordinating Global Brain Project,”; R. Yuste and C. Bargmann, PIs.), the IBM- Columbia University Data Science Institute grant (“Noninvasive Brain Computer Interface Data: Ethical and Privacy Challenges”; R. Yuste and Ken Shepard, PIs.), and the Precision Medicine & Society Program, from Columbia University College of Physicians & Surgeons (“Genomic Data Regulation: A Legal Framework for NeuroData Privacy Protection”; R. Yuste, and G. Hripesak, PIs.).

References

1. Jorgenson, Lyric A., William T. Newsome, David J. Anderson, Cornelia I. Bargmann, Emery N. Brown, Karl Deisseroth, John P. Donoghue, Kathy L. Hudson, Geoffrey S.F. Ling, and Peter R. MacLeish. 2015. The BRAIN initiative: Developing technology to catalyse neuroscience discovery. *Philosophical Transactions of the Royal Society B: Biological Sciences* 370 (1668): 20140164.
2. N.A. 2017. Welcome to the Human Brain Project. <https://www.humanbrainproject.eu/en/>. Accessed May 6, 2020.
3. N.A. 2020. Transforming the future through brain science. <https://canadianbrain.ca>. Accessed May 6, 2020.
4. New York Times. 2020. The Brain Implants That Could Change Humanity.
5. Muelling, Katharina, Arun Venkatraman, Jean-Sebastien Valois, John E. Downey, Jeffrey Weiss, Shervin Javdani, Martial Hebert, Andrew B. Schwartz, Jennifer L. Collinger, and J. Andrew Bagnell. 2017. Autonomy infused teleoperation with application to brain computer interface controlled manipulation. *Autonomous Robots* 41 (6): 1401–1422. <https://doi.org/10.1007/s10514-017-9622-4>.
6. Cook, Mark J., Terence J. O'Brien, Samuel F. Berkovic, Michael Murphy, Andrew Morokoff, Gavin Fabinyi, Wendyl D'Souza, Raju Yerra, John Archer, Lucas Litewka, Sean Hosking, Paul Lightfoot, Vanessa Ruedebusch, W. Douglas Sheffield, David Snyder, Kent Leyde, and David Himes. 2013. Prediction of seizure likelihood with a long-term, implanted advisory system in patients with drug-resistant epilepsy: A first-in-man study. *The Lancet Neurology* 12 (6): 563–571. [https://doi.org/10.1016/S1474-4422\(13\)70075-9](https://doi.org/10.1016/S1474-4422(13)70075-9).
7. Steinert, Steffen, and Orsolya Friedrich. 2020. Wired emotions: Ethical issues of affective brain–computer interfaces. *Science and Engineering Ethics* 26 (1): 351–367. <https://doi.org/10.1007/s11948-019-00087-2>.
8. Yuste, Rafael, Goering, Sara, and et al. 2017. Four ethical priorities for neuro technologies and AI. *Nature* 551: 159–163.
9. Greely, Henry T., Christine Grady, Khara M. Ramos, Winston Chiong, James Eberwine, Nita A. Farahany, L. Syd M. Johnson, Bradley T. Hyman, Steven E. Hyman, and Karen S. Rommelfanger. 2018. Neuroethics guiding principles for the NIH BRAIN initiative. *Journal of Neuroscience* 38 (50): 10586–10588.
10. Rommelfanger, Karen, Sung-Jin Jeong, Arisa Ema, Tamama Fukushi, Kiyoto Kasai, Khara Ramos, Arleen Salles, and Iliina Singh. 2018. Neuroethics questions to guide ethical research in the international brain initiatives. *Neuron* 100 (1): 19–36.
11. OECD Recommendation on Responsible Innovation in Neurotechnology. 2019. ed. OECD. oecd.org: The Organisation for Economic Co-operation and Development.
12. Hildt, Elisabeth. 2015. What will this do to me and my brain? Ethical issues in brain-to-brain interfacing. *Frontiers in Systems Neuroscience* 9 (17). <https://doi.org/10.3389/fnsys.2015.00017>.

13. Loftus, Elizabeth F. 2005. Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory* 12 (4): 361–366.
14. Wegner, Daniel M., Betsy Sparrow, and Lea Winerman. 2004. Vicarious agency: Experiencing control over the movements of others. *Journal of Personality and Social Psychology* 86 (6): 838–848.
15. Sato, Atsushi, and Asako Yasuda. 2005. Illusion of sense of self-agency: Discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of self-ownership. *Cognition* 94 (3): 241–255.
16. Dubiel, Helmut. 2009. *Deep in the brain: Living with Parkinson's disease*. New York: Europa Editions.
17. Pham, Uyen, Anne-Kristin Solbakk, Inger-Marie Skogseid, Mathias Toft, Are Hugo Pripp, Ane Eidahl Konglund, Stein Andersson, Ira Ronit Haraldsen, Dag Aarstrand, and Espen Dietrichs. 2015. Personality changes after deep brain stimulation in Parkinson's disease. *Parkinson's Disease* 2015.
18. Agid, Yves, Michael Schüpbach, M Gargiulo, L Mallet, JL Houeto, C Behar, D Maltête, V Mesnage, and ML Welter. 2006. Neurosurgery in Parkinson's disease: The doctor is happy, the patient less so? In *Parkinson's disease and related disorders*, 409–414. Springer.
19. Haan, De, Erik Rietveld Sanneke, Martin Stokhof, and Damiaan Denys. 2017. Becoming more oneself? Changes in personality following DBS treatment for psychiatric disorders: Experiences of OCD patients and general considerations. *PLoS One* 12 (4).
20. Schechtman, Marya. 2009. Getting our stories straight: Self-narrative and personal identity.
21. Pugh, Jonathan, Hannah Maslen, and Julian Savulescu. 2017. Deep brain stimulation, authenticity and value. *Cambridge Quarterly of Healthcare Ethics* 26 (4): 640–657.
22. Baylis, Françoise. 2013. "I am who I am": On the perceived threats to personal identity from deep brain stimulation. *Neuroethics* 6 (3): 513–526.
23. Sullivan, Specker, and Laura. 2019. Insight and the no-self in deep brain stimulation. *Bioethics* 33 (4): 487–494.
24. Klein, Eran, Sara Goering, Josh Gagne, Conor V. Shea, Rachel Franklin, Samuel Zorowitz, Darin D. Dougherty, and Alik S. Widge. 2016. Brain-computer interface-based control of closed-loop brain stimulation: Attitudes and ethical considerations. *Brain-Computer Interfaces* 3 (3): 140–148. <https://doi.org/10.1080/2326263X.2016.1207497>.
25. Haselager, Pim. 2013. Did I do that? Brain-computer interfacing and the sense of agency. *Minds and Machines* 23 (3): 405–418. <https://doi.org/10.1007/s11023-012-9298-7>.
26. Lipsman, N.I.R., and Walter Glannon. 2013. Brain, mind and machine: What are the implications of deep brain stimulation for perceptions of personal identity, agency and free will? *Bioethics* 27 (9): 465–470.
27. Kellmeyer, Philipp, Thomas Cochrane, Oliver Müller, Christine Mitchell, Tonio Ball, Joseph J. Fins, and Nikola Biller-Andorno. 2016. The effects of closed-loop medical devices on the autonomy and accountability of persons and systems. *Cambridge Quarterly of Healthcare Ethics* 25 (4): 623–633.
28. Bassil, Katherine C., Bart P.F. Rutten, and Dorothee Horstkötter. 2019. Biomarkers for PTSD susceptibility and resilience, ethical issues. *AJOB Neuroscience* 10 (3): 122–124.
29. Lindemann, Hilde. 2014. *Holding and letting go: The social practice of personal identities*. 1st ed. Oxford: Oxford University Press.
30. Wimber, Maria, Arjen Alink, Ian Charest, Nikolaus Kriegeskorte, and C. Anderson. 2015. Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression. *Nature Neuroscience* 18: 582–589. <https://doi.org/10.1038/nn.3973>.
31. Aas, Sean, and David Wasserman. 2016. Brain-computer interfaces and disability: Extending embodiment, reducing stigma? *Journal of Medical Ethics* 42 (1): 37–40.
32. Reiner, Peter Bart, and Saskia K. Nagel. 2017. Technologies of the Extended Mind: Defining the issues. In *Neuroethics: Anticipating the future*, ed. J. Illes, 108–122. Oxford: Oxford University Press.
33. Dennett, Daniel C. 2017. Who am I? In *Brainstorms: Philosophical essays on mind and psychology*, Daniel C. Dennett, 333–346. MIT Press.
34. Hildt, Elisabeth. 2019. Multi-person brain-to-brain interfaces: Ethical issues. *Frontiers in Neurosciences* 13: 1177.
35. Heersmink, Richard. 2017. Distributed cognition and distributed morality: Agency, artifacts and systems. *Science and Engineering Ethics* 23 (2): 431–448.
36. Bublitz, Christoph, Andreas Wolkenstein, Ralf J. Jox, and Orsolya Friedrich. 2019. Legal liabilities of BCI-users: Responsibility gaps at the intersection of mind and machine? *International Journal of Law and Psychiatry* 65: 101399–101399.
37. Wilson, George, and Samuel Shpall. 2016. Action. *Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/win2016/entries/action/> accessed 22 December 2019.
38. Gallagher, Shaun. 2000. Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Sciences* 4 (1): 14–21.
39. Wegner, Daniel. 2002. *The illusion of conscious will*. MIT Press.
40. Sellers, Eric W., Theresa M. Vaughan, and Jonathan R. Wolpaw. 2010. A brain-computer interface for long-term independent home use. *Amyotrophic Lateral Sclerosis* 11 (5): 449–455.
41. Ajiboye, A. Bolu, Francis R. Willett, Daniel R. Young, William D. Mernberg, Brian A. Murphy, Jonathan P. Miller, Benjamin L. Walter, Jennifer A. Sweet, Harry A. Hoyen, and Michael W. Keith. 2017. Restoration of reaching and grasping movements through brain-controlled muscle stimulation in a person with tetraplegia: A proof-of-concept demonstration. *The Lancet* 389 (10081): 1821–1830.
42. Christen, Markus, Josep Domingo-Ferrer, Bogdan Draganski, Tade Spranger, and Henrik Walter. 2016. On the compatibility of big data driven research and informed consent: The example of the human brain project. In *The ethics of biomedical big data*, 199–218. Springer.
43. Kellmeyer, Philipp. 2018. Big brain data: On the responsible use of brain data from clinical and consumer-directed neurotechnological devices. *Neuroethics*:1–16.

44. Willett, Francis R, Donald T Avansino, Leigh R Hochberg, Jaimie M Henderson, and Krishna V Shenoy. 2020. High-performance brain-to-text communication via imagined handwriting. *bioRxiv*.
45. Pawar, Dipti, and Sudhir Dhage. 2020. Multiclass covert speech classification using extreme learning machine. *Biomedical Engineering Letters*: 1–10.
46. Graves, Lisa. 2010. Burdens of war: The consequences of the U.S. military response to 9/11 the costs to civil liberties and the rule of law in the U.S. *Watson Institute of International and Public Affairs*.
47. Floyd, Rita. 2019. The morality of security: A theory of just securitization. Cambridge University Press.
48. Steel, Emily, and Geoffrey Fowler. 2010. Facebook in privacy breach. *The Wall Street Journal* 18 (1).
49. Wongkoblap, Akkapon, Miguel A. Vadillo, and Vasa Curcin. 2017. Researching mental health disorders in the era of social media: Systematic review. *Journal of Medical Internet Research* 19 (6): e228.
50. Metz, C. 2019. With \$1 billion from Microsoft an AI lab wants to mimic the brain. *New York Times*.
51. Wagner, Kurt. 2019. Facebook to buy startup for controlling computers with your mind. *Bloomberg*.
52. Ravindra, Vikram, and Ananth Grama. 2019. De-anonymization attacks on neuroimaging datasets. *arXiv preprint arXiv:1908.03260*.
53. Takabi, Hassan, Anuj Bhalotiya, and Manar Alohaly. 2016. Brain computer interface (BCI) applications: Privacy threats and countermeasures. In *2016 IEEE 2nd International Conference on Collaboration and Internet Computing (CIC)*: IEEE.
54. Martinovic, Ivan, Doug Davies, Mario Frank, Daniele Perito, Tomas Ros, and Dawn Song. 2012. On the feasibility of side-channel attacks with brain-computer interfaces. In *Presented as part of the 21st {USENIX} security symposium ({USENIX} Security 12)*.
55. Coates McCall, I, and J Illes. 2018. Owning ethical innovation: Claims about commercial brain wearable technologies. In *International Neuroethics Society Annual Meeting*: Neuron.
56. Minielly, Nicole, Viorica Hrinco, and Judy Illes. 2020. A view on incidental findings and adverse events associated with neurowearables in the consumer marketplace. *Ethical Dimensions of Commercial and DIY Neurotechnologies* 3: 267.
57. Minielly, Nicole, Viorica Hrinco, and Judy Illes. 2020. Privacy challenges to the democratization of brain data. *Isience*:101134.
58. Plassmann, Hilke, Thomas Zoëga Ramsøy, and Milica Milosavljevic. 2012. Branding the brain: A critical review and outlook. *Journal of Consumer Psychology* 22 (1): 18–36.
59. Plaut, Victoria C., P. Robert, and I.I.I. Bartlett. 2012. Blind consent? A social psychological investigation of non-readership of click-through agreements. *Law and Human Behavior* 36 (4): 293–311.
60. Carrillo-Reid, Luis, Shuting Han, Weijian Yang, Alejandro Akrouh, and Rafael Yuste. 2019. Controlling visually guided behavior by holographic recalling of cortical ensembles. *Cell* (2): 178, 447–457.
61. Deeb, Wissam, Bryan Salvato, Leonardo Almeida, Kelly D. Foote, Robert Amaral, Jurgen Germann, Paul B. Rosenberg, et al. 2019. Fornix-region deep brain stimulation–induced memory flashbacks in Alzheimer’s disease. *New England Journal of Medicine* 381 (8): 783–785. <https://doi.org/10.1056/NEJMc1905240>.
62. Kosinski, Michal, David Stillwell, and Thore Graepel. 2013. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences* 110 (15): 5802–5805.
63. Yin, Z., L.M. Sulieman, and B.A. Malin. 2019. A systematic literature review of machine learning in online personal health data. *Journal of the American Medical Informatics Association* 26 (6): 561–576. <https://doi.org/10.1093/jamia/ocz009>.
64. Strickland, Eliza. 2017. Facebook announces “typing-by-brain” project. *IEEE Spectrum* 20.
65. Debatin, Bernhard, Jennette P. Lovejoy, Ann-Kathrin Horn, and Brittany N. Hughes. 2009. Facebook and online privacy: Attitudes, behaviors, and unintended consequences. *Journal of Computer-Mediated Communication* 15 (1): 83–108.
66. Epstein, Steven. 2008. Inclusion: The politics of difference in medical research. University of Chicago Press.
67. Perez, Caroline Criado. 2019. Invisible women: Exposing data bias in a world designed for men. *Random House*.
68. Rollins, O. 2021. Towards an antiracist (neuro)science. *Nature Human Behaviour*. <https://doi.org/10.1038/s41562-021-01075-y>.
69. Goering, Sara. 2017. Thinking differently: Neurodiversity and neural engineering. In *The Routledge handbook of neuroethics*, 37–50. Routledge.
70. Fine, Cordelia. 2014. His brain, her brain? *Science* 346 (6212): 915–916.
71. Cipolla, Cyd, and Kristina Gupta. 2017. Neurogenderings and neuroethics. In *The Routledge Handbook of Neuroethics*, 381–393. Routledge.
72. Laudel, Grit. 2006. The art of getting funded: How scientists adapt to their funding conditions. *Science and Public Policy* 33 (7): 489–504.
73. Dalton, Rex. 2003. Berkeley accused of biotech bias as ecologist is denied tenure: Nature publishing group.
74. Ferguson, Christopher J., and Michael T. Brannick. 2012. Publication bias in psychological science: Prevalence, methods for identifying and controlling, and implications for the use of meta-analyses. *Psychological Methods* 17 (1): 120–128.
75. Ioannidis, John P.A., Marcus R. Munafò, Paolo Fusar-Poli, Brian A. Nosek, and Sean P. David. 2014. Publication and other reporting biases in cognitive sciences: Detection, prevalence, and prevention. *Trends in Cognitive Sciences* 18 (5): 235–241.
76. Krinsky, Sheldon, and Leslie Steven Rothenberg. 2001. Conflict of interest policies in science and medical journals: Editorial practices and author disclosures. *Science and Engineering Ethics* 7 (2): 205–218.
77. Sparrow, Robert. 2005. Defending deaf culture: The case of cochlear implants. *Journal of Political Philosophy* 13 (2): 135–152.

78. Crouch, Robert A. 1997. Letting the deaf be deaf: Reconsidering the use of cochlear implants in prelingually deaf children. *Hastings Center Report* 27 (4): 14–21.
79. Bauman, H-Dirksen L, and Joseph J Murray. 2014. Deaf gain: Raising the stakes for human diversity. U of Minnesota Press.
80. Goering, Sara, and Eran Klein. 2019. Neurotechnologies and justice by, with, and for disabled people. In *The Oxford Handbook of Philosophy and Disability*.
81. Allard, J., F. Ballesteros, S.J. Anthony, V. Dumez, D. Hartell, G. Knoll, L. Wright, and M.C. Fortin. 2018. What does patient engagement mean for Canadian National Transplant Research Program Researchers? *Res Involv Engagem* 4: 13. <https://doi.org/10.1186/s40900-018-0096-0>.
82. Jagosh, J., P.L. Bush, J. Salsberg, A.C. Macaulay, T. Greenhalgh, G. Wong, M. Cargo, L.W. Green, C.P. Herbert, and P. Pluye. 2015. A realist evaluation of community-based participatory research: Partnership synergy, trust building and related ripple effects. *BMC Public Health* 15: 725. <https://doi.org/10.1186/s12889-015-1949-1>.
83. Antman, Karen, David Amato, Wood William, J. Carson, Herman Suit, Karl Proppe, Carey Robert, J. Greenberger, R. Wilson, and E. Frei 3rd. 1985. Selection bias in clinical trials. *Journal of Clinical Oncology* 3 (8): 1142–1147.
84. Dresser, Rebecca. 1992. Wanted single, white male for medical research. *The Hastings Center Report* 22 (1): 24–29.
85. Health, National Institutes of. 1993. *National Institutes of Health revitalization act of 1993*. Bethesda: NIH.
86. Food, US. 2012. Drug administration safety and innovation act. *Public law*: 112–144.
87. Fox-rawlings, Stephanie R., Laura B. Gottschalk, Lauren A. Doamekpor, and Diana M. Zuckerman. 2018. Diversity in medical device clinical trials: Do we know what works for which patients? *The Milbank Quarterly* 96 (3): 499–529.
88. Wexler, Anna. 2018. Who uses direct-to-consumer brain stimulation products, and why? A study of home users of tDCS devices. *Journal of Cognitive Enhancement* 2 (1): 114–134.
89. Eaton, Margaret L., and Judy Illes. 2007. Commercializing cognitive neurotechnology—The ethical terrain. *Nature Biotechnology* 25 (4): 393–397.
90. Bird, Stephanie J. 2012. Potential for Bias in the context of Neuroethics. *Science and Engineering Ethics* 18 (3): 593–600.
91. Juengst, Eric. 1998. What does “enhancement” mean? In *Enhancing human traits: Ethical and social implications*, ed. E. Parens. Washington, DC: Georgetown University Press.
92. Parens, Erik. 2000. *Enhancing human traits: Ethical and social implications*. Georgetown University Press.
93. Rasmussen, Nicolas. 2008. *On speed: The many lives of amphetamine*. NYU Press.
94. Franke, Andreas G., Caroline Bonertz, Michaela Christmann, Michael Huss, Andreas Fellgiebel, Elisabeth Hildt, and Klaus Lieb. 2011. Non-medical use of prescription stimulants and illicit use of stimulants for cognitive enhancement in pupils and students in Germany. *Pharmacopsychiatry* 44 (02): 60–66.
95. Parens, Erik. 2005. Authenticity and ambivalence: Toward understanding the enhancement debate. *Hastings Center Report* 35 (3): 34–41.
96. Bolt, L.L.E. 2007. True to oneself? Broad and narrow ideas on authenticity in the enhancement debate. *Theoretical Medicine and Bioethics* 28 (4): 285–300.
97. Bublitz, Jan Christoph, and Reinhard Merkel. 2009. Autonomy and authenticity of enhanced personality traits. *Bioethics* 23 (6): 360–374.
98. Bostrom, Nick, and Anders Sandberg. 2009. Cognitive enhancement: Methods, ethics, regulatory challenges. *Science and Engineering Ethics* 15 (3): 311–341.
99. Elisabeth, Hildt. 2013. Cognitive enhancement: An interdisciplinary perspective.
100. Farah, Martha J., and Andrea S. Heberlein. 2007. Personhood and neuroscience: Naturalizing or nihilizing? *The American Journal of Bioethics* 7 (1): 37–48.
101. Bolt, Ineke, and Maartje Schermer. 2009. Psychopharmaceutical enhancers: Enhancing identity? *Neuroethics* 2 (2): 103–111.
102. Farah, Martha J., Caroline Haimm, Geena Sankoorikal, and Anjan Chatterjee. 2009. When we enhance cognition with Adderall, do we sacrifice creativity? A preliminary study. *Psychopharmacology* 202 (1–3): 541–547.
103. Goering, Sara, Eran Klein, Darin D. Dougherty, and Alik S. Widge. 2017. Staying in the loop: Relational agency and identity in next-generation DBS for psychiatry. *AJOB Neuroscience* 8 (2): 59–70.
104. Schermer, Maartje. 2013. Health, happiness and human enhancement—Dealing with unexpected effects of deep brain stimulation. *Neuroethics* 6 (3): 435–445.
105. Buch, Ethan R., Emiliano Santarnecchi, Andrea Antal, Jan Born, Pablo A. Celnik, Joseph Classen, Christian Gerloff, Mark Hallett, Friedhelm C. Hummel, and Michael A. Nitsche. 2017. Effects of tDCS on motor learning and memory formation: A consensus and critical position paper. *Clinical Neurophysiology* 128 (4): 589–603.
106. Coffman, Brian A., Vincent P. Clark, and Raja Parasuraman. 2014. Battery powered thought: Enhancement of attention, learning, and memory in healthy adults using transcranial direct current stimulation. *Neuroimage* 85: 895–908.
107. Lefaucheur, Jean-Pascal, Andrea Antal, Samar S. Ayache, David H. Benninger, Jérôme Brunelin, Filippo Cogiamanian, Maria Cotelli, Dirk De Ridder, Roberta Ferrucci, and Berthold Langguth. 2017. Evidence-based guidelines on the therapeutic use of transcranial direct current stimulation (tDCS). *Clinical Neurophysiology* 128 (1): 56–92.
108. Cabrera, Laura Y., and Peter B. Reiner. 2015. Understanding public (mis) understanding of tDCS for enhancement. *Frontiers in Integrative Neuroscience* 9: 30.
109. Riggall, Kate, Cynthia Forlini, Adrian Carter, Wayne Hall, Megan Weier, Brad Partridge, and Marcus Meinzer. 2015. Researchers’ perspectives on scientific and ethical issues with transcranial direct current stimulation: An international survey. *Scientific Reports* 5 (1): 1–10.
110. Jwa, Anita. 2015. Early adopters of the magical thinking cap: A study on do-it-yourself (DIY) transcranial direct

- current stimulation (tDCS) user community. *Journal of Law and the Biosciences* 2 (2): 292–335.
111. Wexler, Anna. 2016. A pragmatic analysis of the regulation of consumer transcranial direct current stimulation (TDCS) devices in the United States. *Journal of Law and the Biosciences* 2 (3): 669–696.
 112. Fitz, Nicholas S., and Peter B. Reiner. 2015. The challenge of crafting policy for do-it-yourself brain stimulation. *Journal of Medical Ethics* 41 (5): 410–412.
 113. Maslen, Hannah, Tom Douglas, Roi Cohen Kadosh, Neil Levy, and Julian Savulescu. 2015. Do-it-yourself brain stimulation: A regulatory model. *Journal of Medical Ethics* 41 (5): 413–414.
 114. Wexler, Anna. 2016. The practices of do-it-yourself brain stimulation: Implications for ethical considerations and regulatory proposals. *Journal of Medical Ethics* 42 (4): 211–215.
 115. Antal, Andrea, Alekseichuk Ivan, M. Bikson, J. Brockmüller, André R. Brunoni, Robert Chen, L.G. Cohen, G. Dowthwaite, Jens Ellrich, and A. Flöel. 2017. Low intensity transcranial electric stimulation: Safety, ethical, legal regulatory and application guidelines. *Clinical Neurophysiology* 128 (9): 1774–1809.
 116. Maslen, Hannah, Thomas Douglas, Roi Cohen Kadosh, Neil Levy, and Julian Savulescu. 2014. Mind machines: The regulation of cognitive enhancement devices. Oxford Martin School at the University of Oxford.
 117. Brem, Anna-Katharine, Peter J. Fried, Jared C. Horvath, Edwin M. Robertson, and Alvaro Pascual-Leone. 2014. Is neuroenhancement by noninvasive brain stimulation a net zero-sum proposition? *Neuroimage* 85: 1058–1068.
 118. Sarkar, Amar, Ann Dowker, and Roi Cohen Kadosh. 2014. Cognitive enhancement or cognitive cost: Trait-specific outcomes of brain stimulation in the case of mathematics anxiety. *Journal of Neuroscience* 34 (50): 16605–16610.
 119. Gulino, Maurizio, Kim Donghoon, Salvador Pané, Sofia Duque Santos, and Ana Paula Pêgo. 2019. Tissue response to neural implants: The use of model systems towards new design solutions of implantable microelectrodes. *Frontiers in Neuroscience* 13: 689.
 120. Kaebnick, Gregory E., Elizabeth Heitman, James P. Collins, Jason A. Delborne, Wayne G. Landis, Keegan Sawyer, Lisa A. Taneyhill, and David E. Winickoff. 2016. Precaution and governance of emerging technologies. *Science* 354 (6313): 710–711.
 121. Regulations, Code of Federal. 2009. Protection of human subjects. National Institutes of Health Office for Protection from Research Risks. Title 45.
 122. Kreitmair, Karola V. 2019. Dimensions of ethical direct-to-consumer neurotechnologies. *AJOB Neuroscience* 10 (4): 152–166.
 123. Kirby, Kate, Aidan Moran, and Suzanne Guerin. 2011. A qualitative analysis of the experiences of elite athletes who have admitted to doping for performance enhancement. *International journal of sport policy and politics* 3 (2): 205–224.
 124. Arnold, Jason F., and Robert M. Sade. 2017. Wearable technologies in collegiate sports: The ethics of collecting biometric data from student-athletes. *The American Journal of Bioethics* 17 (1): 67–70.
 125. Ray, Keisha Shantel. 2016. Not just “study drugs” for the rich: Stimulants as moral tools for creating opportunities for socially disadvantaged students. *The American Journal of Bioethics* 16 (6): 29–38.
 126. Ienca, Marcello, Fabrice Jotterand, and Bernice S. Elger. 2018. From healthcare to warfare and reverse: How should we regulate dual-use neurotechnology? *Neuron* 97 (2): 269–274.
 127. Tennison, Michael N., and Jonathan D. Moreno. 2012. Neuroscience, ethics, and national security: The state of the art. *PLoS Biology* 10 (3): e1001289.
 128. Aicardi, Christine, Lise Bitsch, Nicklas Bang Bådum, Saheli Datta, Kathinka Evers, Michele Farisco, Tyr Fothergill et al.. 2018. Opinion on ‘responsible dual use’. *Ethics and Society*.
 129. Ienca, Marcello, and Roberto Andorno. 2017. Towards new human rights in the age of neuroscience and neurotechnology. *Life Sciences, Society and Policy* 13 (1): 5.
 130. Code of Federal Regulations Annual Edition 2020. In *45: Public Welfare*. Electronic Code of Federal Regulations: U.S. Government Publishing Office.
 131. Donnelly, Mary, and Maeve McDonagh. 2019. Health research, consent and the GDPR exemption. *European Journal of Health Law* 26 (2): 97–119.
 132. Marelli, Luca, and Giuseppe Testa. 2018. Scrutinizing the EU general data protection regulation. *Science* 360 (6388): 496–498.
 133. Hassan, Muneeb Ul, Mubashir Husain Rehmani, and Jinjun Chen. 2019. Differential privacy techniques for cyber physical systems: A survey. *IEEE Communications Surveys & Tutorials*.
 134. Khedr, Alhassan, and Glenn Gulak. 2017. Securedmed: Secure medical computation using gpu-accelerated homomorphic encryption scheme. *IEEE Journal of Biomedical and Health Informatics* 22 (2): 597–606.
 135. Pirnay, Jean-Paul, Etienne Baudoux, Olivier Cornu, Alain Delforge, Christian Delloye, Johan Guns, Ernst Heinen, Etienne Van den Abbeel, Alain Vanderkelen, and Caroline Van Geyt. 2015. Access to human tissues for research and product development. *EMBO Reports* 16 (5): 557–562.
 136. Tanner, Adam. 2016. *How data brokers make money off your medical records*. New York: Scientific American.
 137. Naufel, Stephanie, and Eran Klein. 2020. Brain–computer interface (BCI) researcher perspectives on neural data ownership and privacy. *Journal of Neural Engineering* 17 (1): 016039.
 138. Sullivan, Laura Specker, Eran Klein, Tim Brown, Matthew Sample, Michelle Pham, Paul Tubig, Raney Folland, Anjali Truitt, and Sara Goering. 2018. Keeping disability in mind: A case study in implantable brain–computer Interface research. *Science and Engineering Ethics* 24 (2): 479–504.
 139. Racine, Eric, H.Z. Adriaan van Der Loos, and Judy Illes. 2007. Internet marketing of neuroproducts: New practices and healthcare policy challenges. *Cambridge Quarterly of Healthcare Ethics* 16 (2): 181–194.
 140. Ludolph, Ramona, and Peter J. Schulz. 2018. Debiasing health-related judgments and decision making: A systematic review. *Medical Decision Making* 38 (1): 3–13.

141. Page, Scott E. 2019. *The diversity bonus: How great teams pay off in the knowledge economy*. Princeton University Press.
142. Conference, Gordon Research. 2020. *The GRC Power Hour*. <https://www.grc.org/the-power-hour/>. Accessed 2020.
143. Cutler, Adam, Milena Pribić, and Lawrence Humphrey. 2019. *Everyday ethics for artificial intelligence*. PDF, IBM Corporation.
144. . 2016. In *International Summit on Human Gene Editing: A Global Discussion*, ed. S. Olson. The National Academies Collection: Reports funded by National Institutes of Health. Washington (DC).
145. Jasanoff, Sheila, J. Benjamin Hurlbut, and Krishanu Saha. 2015. CRISPR democracy: Gene editing and the need for inclusive deliberation. *Issues in Science and Technology* 32 (1): 37.
146. Baylis, Françoise. 2019. *Altered inheritance: CRISPR and the ethics of human genome editing*. Harvard University Press.
147. Archard, David, Jonathan Montgomery, Simon Caney, Tara Clancy, Jeanette Edwards, Ann Gallagher, Andy Greenfield, Erica Haimes, Julian Hughes, and Roland Jackson. 2016. Nuffield Council on Bioethics.
148. Owen, Richard, Phil Macnaghten, and Jack Stilgoe. 2012. Responsible research and innovation: From science in society to science for society, with society. *Science and Public Policy* 39: 751–760. <https://doi.org/10.1093/scipol/scs093>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.