On the Necessity of Ethical Guidelines for Novel Neurotechnologies

Sara Goering[1,*] and Rafael Yuste[2,*]
[1]Department of Philosophy, and Center for Sensorimotor Neural Engineering, University of Washington, Seattle, WA 98195, USA
[2]Neurotechnology Center, Kavli Institute of Brain Science, Department of Biological Sciences and Neuroscience, Columbia University, New York, NY 10027, USA
*Correspondence: sgoering@uw.edu (S.G.), rmy5@columbia.edu (R.Y.)
http://dx.doi.org/10.1016/j.cell.2016.10.029

Because novel neurotechnologies may alter human identity and society in profound ways, we advocate for the early integration of ethics into neurotechnology. We recommend developing and adopting a set of guidelines, like the Belmont Report on human subject research, as a framework for development and use of brain-related technologies.

The Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative, as well as other such large-scale projects around the world, is poised to revolutionize our capacity for recording and manipulating large-scale neuronal activity. These methods could spur a new era in the scientific understanding of neural circuits and also enable powerful novel therapeutic approaches to mental and neurological diseases. At the same time, these methods will provide access to the core mechanisms that underlie human identify, memories, emotions, personality and, more generally, our minds. As such, they could have profound consequences for human identity and society.

Spurred by a recent workshop that brought together neurotechnologists and bioethicists (http://ntc-symposium.org), in this Commentary we highlight the need for strong advocacy toward developing and funding neuroethical work to accompany these advances in neuroscience and to guide the development of neurotechnologies. We think that the time is ripe, given that neural interventions currently being studied in many animal models already demonstrate the capacity to decode imagery and intentions, stimulate or alter sensory perceptions, enhance and combine brain processing power, and alter animal behavior. Even scientists working directly in these areas can recognize the rich opportunities to better understand neural processing while still expressing some trepidation about the kind of future we may be bringing about. This situation calls for the development of a clear set of guidelines, similar to the Belmont Commission Report, to integrate the development and use of these technologies with our core societal and human values. These guidelines could be developed by multidisciplinary panels, composed of scientific, medical, bioethical, and legal experts, as well as representatives of the citizenry at large. Such panels would help to ensure that neurotechnology is developed in ways that are sensitive to some of the profound qualities that serve as the condition of human experience: private mental life, agential action on the world, and an understanding of individuals as bounded by their bodies.

Novel Neurotechnologies and Their Future Use

The BRAIN initiative, sponsored by the White House in 2013, is a large-scale, 12-year-long project aimed at creating tools to interrogate and alter neural circuits in experimental animals and humans with unprecedented detail (https://www.whitehouse.gov/BRAIN). Similar initiatives are now underway in the European Union, Japan, Korea, Canada, Australia, Israel, and China. As a consequence of these projects, more than two hundred laboratories around the world are now funded to systematically develop new methods for recording and manipulating the activity of neural circuits, including awake behaving animals or human subjects. These methods use optical, electrical, or chemical platforms, aided by novel computational tools that help decipher this trove of neural activity.

While we are still far from properly "breaking the code" that any nervous system uses to generate behavior or mental states, decoding efforts are progressing swiftly. For example, using the relatively low spatio-temporal resolution method of fMRI, researchers can start to predict the activity patterns associated with complex visual stimuli that a human subject has been exposed to (Kay et al., 2008), and the ability to access high-quality recordings of neural activity will make these efforts more powerful. Once encoding models are accurate, researchers may be able to decode visual imagery from active human subjects, enabling a technology-based kind of "mind reading." Visual imagery is just one of many possibilities for decoding: other scenarios such as the decoding of speech, thoughts (including lies), and dreams, or even the decoding of internal states of animals, can be envisioned. Brain decoding could become ubiquitous. Human subjects using portable decoders could also perhaps covertly communicate with each other without the need to speak, sign, or type their thoughts.

In addition, powerful methods can start to manipulate neuronal activity with single-cell precision. In experimental animals, such as worms, *Drosophila*, zebra fish, or mice, optogenetics is routinely used to alter behavior (Yizhar et al., 2011), including triggering of memories (Ramirez et al., 2013). Two-photon optogenetics enables researchers to implant, and later replay, patterns of activity into the cerebral cortex of awake mice (Carrillo-Reid et al., 2016). While optogenetics is unlikely to be applied to humans in the near future because it involves genetic manipulation, other optical stimulation methods using optochemical compound or nanoparticles do not require any genetic manipulation and could be more easily adaptable to humans. Such interventions offer considerable promise for managing unwanted patterns of thought and behavior in human patients but also raise daunting possibilities for control.

Manipulating neural circuits to alter behavior in human subjects is not new. Deep brain stimulation (DBS) devices are routinely used as treatment for many neurological diseases with some success, though, in some cases, they lead to changes in personality and behavior. In others cases, subjects report feeling the need to use DBS in order to ''become themselves'' (Vlek et al., 2012). In addition, brain computer interfaces (BCI) are routinely used in a wide variety of clinical applications to record neuronal activity and connect it via computer systems to control robotic or prosthetic limbs or devices. With BCIs, subjects can be trained to operate robotic prostheses with their thoughts. Such interfaces have provided the ability for paralyzed patients to move prosthetic limbs (Hochberg et al., 2006). Again, once high-quality recordings on neural activity become routine, BCI will become much more powerful. In particular, if neural recordings are performed in a non-invasive fashion, BCI could become widely used to augment the physical or intellectual capabilities of humans. In addition, several human subjects could jointly train connected BCIs to perform a joint task, blurring the identity (and responsibility) of individual users.

Ethical and Societal Issues Raised by Novel Neurotechnologies

While some would say that these scenarios are still in the realm of science-fiction, we believe that they are coming down the pipeline, given ongoing research. Thus, they warrant the question of how the application of novel neurotechnologies to humans should be properly guided and regulated. This problem is not novel, since humankind has been experimenting with methods to alter and manipulate brain activity for thousands of years, for example with alcohol, recreational drugs, or pharmacology. But while those methods normally affect the brain in a relatively coarse fashion, the more we learn about neural processing and the more powerful neuromodulation methods become, the more profound will be the effect of these manipulations on mental states and behavior.

Ethical, legal, and societal issues arising from novel neurotechnologies are many and could unfold progressively: some issues may arise in the very near term, whereas others are farther out and will depend on how research in the intervening period proceeds. Here, we review just a few issues on the horizon, given topics within currently funded neuroscientific research.

One of the more important ethical and philosophical issues ahead is the possibility of substantial changes in the concept of self. While at present we tend to identify ourselves as relatively separate, private entities bounded by our bodies (and even on more dynamic, relational views, like Baylis [2013], we still have private consciousness and distinct agency), the use of novel neurotechnologies may lead to a partial dissolution of traditional ideas of self. Access to neural activity will also call into the question the privacy of our internal lives (Farah and Wolpe, 2004), while the capacity to neuromodulate that activity may alter our sense of agency (Haselager, 2013), along with issues of moral and legal responsibility for our thoughts and actions (Farahany, 2012). If our neural activity can control devices beyond our bodies through thought alone (via BCI devices), our internal body schemas may stretch to encompass the devices under our immediate conscious control, essentially extending our sense of self beyond the boundaries of our bodies. BCI users can already learn to control a robotic arm just by thinking; what if they could send it out of the room to collect a needed item? *Where* is the user in this scenario?

In the other direction, we can also envision changes to our sense of self through the possibility of collecting multiple BCI users into a shared task in a way that results in a common action. Brain-to-brain interface experiments (Stocco et al., 2015), while still primitive in many respects, demonstrate the possibility of sending neural activity from one person's brain directly into another person's brain. Some experiments have demonstrated the capacity to connect neural activity of even three different experimental animals for a common purpose (Pais-Vieira et al., 2015). In these scenarios, assignments of responsibility and understandings of self and agency will be complicated. The boundary of self (as body) becomes permeable to another's consciousness and control.

These issues are philosophically profound and morally and legally momentous. Traditional understandings of who (and where) we are and what we are responsible for are foundational for our legal systems and our moral interactions with each other. Proper protections for private internal spaces and agential identity will need to be integrated into our understandings of human rights. The use of methods that may substantially alter one's personality, thoughts, and sensorimotor experience requires attention to individual and societal protections.

In addition, the ability to augment one's physical or mental performance raises a host of issues about fairness and justice regarding how those augmenting technologies should be accessed or regulated. Are they intended for mass consumption, or should they be restricted to human users who have identifiable impairments (Aas and Wasserman, 2016)? How should we think about the standards of ''normal functioning'' and ability expectation (Wolbring and Diep, 2016) in the context of technologies that suggest a kind of ability revolution? If such technology is expensive, as it is bound to be at the beginning, it could generate or exacerbate societal divisions among the population or among inhabitants of different countries.

Pertinent to this discussion are the mobile ''smart'' devices, such as cell phones, that are increasingly part of our decision making and that could be interpreted as a basic, low-tech BCI devices. The current generation of smart phones is the precursor of future, more powerful prosthetic devices, which could be worn or implanted. The profound impact that smart phones have had in our society and culture over the last decade presages an even larger effect that brain-controlled wearable electronics devices may have. Importantly, in addition to enhancing our access to information, these interactive neural devices incorporate machine-learning algorithms (e.g., in the service of facial recognition software) that may incorporate and reproduce implicit bias, reinforcing negative stereotypes among racial groups or minorities.

Although many of the novel neurotechnologies are currently being explored in the context of medical research and patient care, or in the frame of assistive devices, as with all technology, they are also likely to have commercial applications that would, of necessity, involve different forms of regulation and oversight. Furthermore, one can easily envision military or security applications. As with issues such as chemical, biological, or nuclear weapons or, more recently, cyber warfare, this raises the concern as to the rules under which novel neurotechnologies should be used, if at all, in human conflagrations.

Developing Ethical Principles for Neurotechnologies

We think it is fair to say that, to date, we lack any agreed-upon guidelines to responsibly shape the development and application of these novel neurotechnologies. How could one approach the task of building such ethical guidelines? One possibility is to follow the case of medicine as a natural example, given that medical technology also represents a set of methods that interfere with the function of the human body. Indeed, in medicine, the use of technologies to monitor and alter the capabilities of the human body has always been guided by a common humanistic goal: to help patients in need or, more generally, to promote, without borders, the health of the entire population of the world. Over millennia, at least since the Hippocratic Oath, medicine has developed a corpus of ethical rules that have formed a deontology. These medical ethics principles are taught in medical schools and adhered to closely by practitioners of medicine throughout much of the world and across history. Principles of medical ethics are also respected by society as a whole, including scientists, governments, the private sector, and the military. After the Second World War, modern medical ethics were institutionalized by the Belmont Report, a document generated by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research in the US (http://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html). This report proposed ethical principles and guidelines for research involving human subjects under three core principles: respect, beneficence, and justice. Issues such as informed consent, assessment of risks and benefits, and selection of subjects are dealt with in a practical fashion. The Belmont report is widely respected and constitutes the core set of values underlying modern medical practice.

We think that the time is ripe for a similar set of principles, a Belmont report for neuroethics. These principles should offer guidelines for the effective protection of human subjects and values in conjunction with the ongoing research in this new area. But who should generate such a set of ethical and societal guidelines for the new technologies? Scientific experts who are developing these methods should be involved, given that they know better than anyone the current and likely future capabilities of the technologies. In addition, medical practitioners can contribute their experience interacting with patients seeking help for neurologically related disorders and people interested in assistive devices. Bioethicists, who work at the interface between ethical issues and biomedicine, are also clearly needed. Legal experts can provide important insights regarding human rights protections and approaches to the integration of these technologies in the legal codes of society. Finally, we think that representatives of the citizenship at large should also be involved, including of disabled people who are members of the likely early target populations for BCI and neuromodulation experiments designed to address impairments or offer assistive devices. The progress of science, while it should be freely pursued for curiosity's sake as the best way of enhancing knowledge, should also be informed by the needs and circumstances of the society that, after all, funds and supports all scientific work (Kitcher, 2011). Representatives of the citizens could also play a critical role in translating to the wider society the importance of these methods and discussions and the guidelines that are recommended. Given the amount of hype that abounds in the press and in science-fiction literature and movies, it is of great importance that citizens clearly understand, without any exaggeration, the potential benefits of these technologies, as well as their potential dystopian outcomes.

To complement and sustain the work of these panels, we recommend robust funding of neuroethics in order to develop it further as a subfield of bioethics. This could enable a vigorous academic and societal debate regarding the meaning and consequences of this technology and ensure that the guidelines reflect the best ideas generated by the community. The current NIH BRAIN initiative program to fund neuroethics work (https://grants.nih.gov/grants/guide/notice-files/NOT-MH-16-014.html), together with the recent creation of a Neuroethics working group (https://www.braininitiative.nih.gov/about/newg.htm), is a good start but still small steps relative to the significance of the issues at hand.

The Importance of Novel Neurotechnology for the Progress of Humankind

In closing, our aim here is not to contribute to or feed fear for doomsday scenarios but to ensure that we are reflective and intentional as we prepare ourselves for the neurotechnological future, which we believe can be a momentous positive change in our history. Indeed, novel neurotechnologies could serve as a liberating force for humankind. Humans have been defined by our tool making, and new technologies and the knowledge accumulated through them have enabled us to free ourselves from some of the tyrannies of prejudice and, in a way, of space and time. Scientific progress has enabled impressive opportunities for health and well-being, communication, and trade. Similar benefits probably lie ahead with the incorporation of novel neurotechnologies to our society. But these advances should be shared equitably, and a responsible society should aim to develop them in ways that are both broadly beneficial and sensitive to individual rights and needs. As we turn our impressive powers of investigation to the task of exploring and understanding the brain—the organ central to the physiological processes that make us who we are—we ought to think hard about the kinds of beings we might want to become and the kind of society we are building for our children. Public confidence in science should be anchored on responsible deployment of scientific advances. Technological advances must be shaped by our collective moral sensibilities in order to ensure that these advances are smoothly incorporated into our culture and indeed contribute to the common good.

REFERENCES

Aas, S., and Wasserman, D. (2016). J. Med. Ethics *42*, 37–40.

Baylis, F. (2013). Neuroethics *6*, 513–526.

Carrillo-Reid, L., Yang, W., Bando, Y., Peterka, D.S., and Yuste, R. (2016). Science *353*, 691–694.

Farah, M.J., and Wolpe, P.R. (2004). Hastings Cent. Rep. *34*, 35–45.

Farahany, N. (2012). Stanford Law Rev. *64*, 351–408.

Haselager, P. (2013). Minds Mach. *23*, 405–418.

Hochberg, L.R., Serruya, M.D., Friehs, G.M., Mukand, J.A., Saleh, M., Caplan, A.H., Branner, A., Chen, D., Penn, R.D., and Donoghue, J.P. (2006). Nature *442*, 164–171.

Kay, K.N., Naselaris, T., Prenger, R.J., and Gallant, J.L. (2008). Nature *452*, 352–355.

Kitcher, P. (2011). Science in a Democratic Society (Prometheus Books).

Pais-Vieira, M., Chiuffa, G., Lebedev, M., Yadav, A., and Nicolelis, M.A. (2015). Sci. Rep. *5*, 11869.

Ramirez, S., Liu, X., Lin, P.-A., Suh, J., Pignatelli, M., Redondo, R.L., Ryan, T.J., and Tonegawa, S. (2013). Science *341*, 387–391.

Stocco, A., Prat, C.S., Losey, D.M., Cronin, J.A., Wu, J., Abernethy, J.A., and Rao, R.P. (2015). PLoS ONE *10*, e0137303. http://dx.doi.org/10.1371/journal.pone.0137303.

Vlek, R.J., Steines, D., Szibbo, D., Kübler, A., Schneider, M.J., Haselager, P., and Nijboer, F. (2012). J. Neurol. Phys. Ther. *36*, 94–99.

Wolbring, G., and Diep, L. (2016). Cognitive/Neuroenhancement Through an Ability Studies Lens. In Cognitive Enhancement: Ethical and Policy Implications in International Perspectives, F. Jotterand and V. Dubljevic, eds. (Oxford University Press), pp. 57–75.

Yizhar, O., Fenno, L.E., Davidson, T.J., Mogri, M., and Deisseroth, K. (2011). Neuron *71*, 9–34.