INSTRUCTOR GUIDE

SIMULATION: NEUROTECHNOLOGY[1]

OVERVIEW

In this simulation, participants will be called upon to make a series of decisions in response to three morally charged cases. Participants are asked to take seriously how they can justify their positions, as well as how those justifications can be effectively communicated to others who may disagree with them.

CORE QUESTIONS

Technology Specific:

1. Is the therapy/enhancement distinction ethically important? Is it conceptually tenable?
2. Does neuromodulation threaten autonomy? If so, what might ethical governance of the technology require?
3. How do we choose (consciously or not) who determines how controversial technologies get developed and deployed?

General:

4. What is an ethical principle and how does it differ from an ethical value?
5. How are we to judge between competing principles?
6. How can awareness of ethical principles strengthen or impair leadership?

LEARNING GOALS

The scenarios are designed to emphasize the complexity of many ethical decisions and the deep values at stake—and often in conflict—in ethical disagreements. These aspects are emphasized to set up the following takeaways:

(1) Moral values, like "equality," "justice," "integrity," etc., are the unstated premises underlying moral reasoning. However, they are too vague to be action guiding. It is necessary to learn to formulate principles or rules that put values into practice in a consistent way.

(2) People who share the same values may nonetheless offer competing principles for applying those values. The strength of a principle is determined by how consistently it can plausibly apply the value in question to different scenarios, as well as the plausibility of the mechanism by which the principle works.

[1] Created by Brian Palmiter for the Scientific Citizenship Initiative at Harvard University, Spring 2020.

(3) Being able to articulate our values in the action-guiding form of concrete principles (which are essentially decision-rules) is an important leadership skill. It allows a leader to get everyone on the same page in a way that vague "values" in a mission statement cannot.

Scenario 1 offers participants an example of a principle in the form of the therapy/enhancement distinction. The principle is something like "therapeutic use of technologies like DBS is morally permissible, but using the same technology to enhance an otherwise-healthy individual is morally suspect." Debating the validity of the therapy/enhancement distinction calls participants to weigh the merits of a principle (Goal 2).

Scenario 2 focuses on a value (autonomy) that everyone is likely to endorse, but it does not provide a principle to specify the content to the value. Participants are forced to articulate a principle of autonomy in order to assess whether autonomy is threatened by Smart DBS. The difficulty of this process should reveal the problematic vagueness of values compared to the specificity of principles (Goal 1).

Scenario 3 asks participants to consider what viewpoints are most essential when establishing regulations for a technology. The values and principles that might guide this decision are left up to the participants.

RUNNING THE SIMULATION

Class Time Breakdown

- Intro/set up: 10 minutes
- Simulation: 60 minutes
- Debrief: 50 minutes
- Total: 120 minutes

Materials

- Copies of the scenarios for each participant

Procedures

Divide participants into groups of 3-5 individuals for the simulation. To begin, pass out the Background and Setting sheets and give participants time to read them. Then pass out scenarios 1-3, one at a time, giving groups approximately 15 minutes to discuss their response to each case. Decisions are to be made as a group according to the following procedure:

- 5 minutes of private reflection.
- Approx. 10 minutes of group discussion with the goal of arriving at a consensus, but consensus that does justice to the moral complexity of the cases.
- If consensus cannot be reached, majority rules.

Once the discussion period is over and each group has made its decision, the facilitator will call on the groups to report their decisions. Throughout this process, each participant should keep a careful record of the reasons the group considered in their discussion, the votes taken, and the decision the group arrived at. This information will be important when the simulation ends and the activity moves to the "debrief" phase.

THE DEBRIEF

After the simulation is over, the facilitator will bring all the groups together for a common debrief. This is the most important element of the activity, pedagogically speaking, so it should not be given short shrift. In the debrief, the facilitator asks a series of questions that work through the decisionmaking process groups just underwent. The goal is to use metacognitive reflection to inductively achieve the learning goals. During the debrief the facilitator should incorporate some brief didactic segments where appropriate in order to distill and formalize the lessons that emerge organically from the debrief conversation.

NEUROMODULATION

BACKGROUND[2]

Neuromodulation is a branch of functional neurosurgery that aims to treat chronic neurological or psychiatric diseases by surgically targeting deep brain nuclei and pathways involved in the mediation of the symptoms in order to stimulate, inhibit, or otherwise modify, i.e., modulate, pathological activity…

The modern era of neuromodulation started with the publication in 1987 of a paper by Benabid et al. on the use of deep brain stimulation (DBS) for suppressing tremors of Parkinson disease, and the field is still dominated by deep brain stimulation (DBS). Thanks to modern imaging, including functional imaging, and improved surgical techniques, neurosurgeons are now able to implant DBS electrodes virtually anywhere in the brain with a high degree of accuracy and relative safety. Increased understanding of the neural circuits involved in various neurological, psychiatric, cognitive, and behavioral disorders makes it tempting to use the nondestructive stereotactic technique of DBS to modulate these circuits in the hope of alleviating symptoms. The success of this approach on the motor symptoms of Parkinson disease (the most common indication for DBS) has led to enthusiasm for applying DBS beyond movement disorders, in the realm of psychiatry, behavior, and cognition.

To date, DBS trials have targeted no fewer than 40 different brain sites for at least 30 clinical indications. The common denominator of these investigational applications of DBS is their intention to treat symptoms of illnesses and diseases that are refractory to nonsurgical management, be they tinnitus or obesity, depression or dementia, epileptic seizures or phantom pain.

The general euphoria surrounding the potential of the technique of neuromodulation can arouse fears of a “slippery slope” in its future use for enhancement of normal functions or for “indications” beyond disease and illness. The ethics of neuromodulation use in psychiatric and behavioral illnesses is much debated today and the debate now stretches to address potential DBS applications beyond pathology.

[2] *Virtual Mentor.* 2015;17(1):74-81. doi: 10.1001/virtualmentor.2015.17.1.oped2-1501.

SETTING:[3]

You are members of a research team working on a novel treatment for Obsessive Compulsive Disorder (OCD) utilizing deep brain stimulation (DBS). While most patients with OCD eventually respond to treatment with medication and/or behavioral therapy, a small minority do not improve following all conventional treatments. For this small minority, one of the few remaining options is neurosurgery, including lesion procedures or deep brain stimulation.

DBS, the approach your team employs, involves placing electrodes in targeted areas of the brain. Once the electrodes are in place they are connected by wires under the skin to pulse generators under the skin (usually just below the collarbone). The pulse generator or “implantable neurostimulator” contains a battery for power and a microchip to regulate the stimulation. These pulse generators are very similar to those implanted under the skin (also usually just below the collarbone) for patients with cardiac pacemakers. The biggest difference is that in DBS the electrodes are in the brain instead of in the heart (as is the case with cardiac pacemakers). DBS requires opening the skull, but it does not require destroying any brain tissue. In the main alternative, lesion procedures, there is a fixed amount of brain tissue that is destroyed. DBS allows for different amounts of electrical charge, giving the doctors a wider range of treatment.

DBS has shown promise as a treatment for OCD for many years now, but the degree of relief provided by existing therapies is limited. Additionally, DBS only works at all for about 60 percent of treatment-resistant OCD patients that try it. However, the treatment your team has developed, “Smart DBS”, is significantly more promising. In Phase 3 trials it proved effective in over 95 percent of patients, and it offers markedly more control over OCD symptoms. The key innovation behind your team’s treatment is that it pairs the traditional implanted electrodes for DBS with a proprietary machine learning algorithm that can adaptively modulate the stimulation provided by the implants. By learning to recognize each patient’s neural patterns leading up to when he or she is experiencing a compulsive thought, the algorithm can come to anticipate rising compulsions and preempt the compulsion from forming via the DBS implants.

In addition to the relief this promises to provide those suffering from OCD, your team believes there is potential to easily adapt this technique to other indications. Because Smart DBS works by learning to recognize and then preempt certain thoughts—in the case of OCD, compulsive thoughts—before the patient becomes fully conscious of them, it may be possible to train the intervention to recognize and preempt other unwanted thoughts.

[3] The description of how DBS works and its potential applications in cases of extreme OCD is adapted from Dougherty and Greenberg, “Deep Brain Stimulation for Highly Treatment-Resistant OCD,” International OCD Foundation 2009. https://iocdf.org/expert-opinions/expert-opinion-dbs/.

SCENARIO 1: THERAPY VS. ENHANCEMENT

Shortly after a big profile on your team's work on Smart DBS as a treatment for OCD came out in the New York Times, you are approached by Dr. Ian Singleton, an assistant professor of virology at your university, with a request for help. You see, Dr. Singleton is having trouble focusing on his research because it is baseball season and he is a huge sports fan. Singleton reports that he cannot seem to go more than five minutes without thinking about upcoming games, checking ESPN for stories about the last game, tweaking his fantasy baseball lineup, etc. His inability to focus is a problem because 1) he is in the midst of a time-sensitive study that promises to greatly enhance our understanding of vaccine development, and 2) he only has two years left on his tenure clock and he is a few publications short of being a plausible candidate for tenure. Singleton spoke to his therapist about his trouble focusing and they have been working on focusing techniques for several months now, but with little effect. He does not meet the DSM 5 criteria for an OCD diagnosis or for any related medical condition. Nevertheless, he would like to receive the DBS treatment you've developed "off-label" to help him block out the distracting thoughts that are making it difficult for him to focus on his research right now.

Although the therapy you developed is meant to treat OCD, there's no reason to think it would be any less safe or effective at regulating repeated, intrusive thoughts of any kind. Additionally, this kind of off-label use of the therapy is not illegal [if the lab files a compassionate use request.] Dr. Singleton knows DBS is not without danger (it requires opening the skull), but the surgical procedure is relatively routine. With full knowledge, he is willing to accept all risks. Additionally, Singleton's wealthy parents have volunteered to foot the bill for his procedure *and* donate $3 million to the team's research lab as a gesture of gratitude.

Some at the lab have expressed concern that providing Singleton with Smart DBS to help him focus would be unethical because it amounts to using the technology as an enhancement instead of a therapy. Others accept that the therapy/enhancement distinction is ethically meaningful, but believe that using Smart DBS in Singleton's case would be therapeutic even though his unwanted thoughts are not officially recognized as OCD. And yet others don't believe the therapy/enhancement question is relevant at all when deciding what to do about his case. What do you think? Is there an ethical difference between the therapeutic use of a technology and use for enhancement? And all things considered, should the team grant Dr. Singleton's request?

A. Yes, the therapy enhancement distinction matters, and therefore the team should respectfully refuse to help Dr. Singleton.
B. Yes, the therapy enhancement distinction matters, but the team should help Dr. Singleton anyway.
C. The therapy enhancement distinction does not matter and the team should help Dr. Singleton.
D. The therapy enhancement distinction does not matter, but the team should respectfully refuse to help Dr. Singleton anyway.

SCENARIO 2: AUTONOMY

Smart DBS has become an increasingly common treatment for OCD and other psychiatric disorders, and users are thrilled by the relief from their symptoms it provides. However, while everyone is glad to no longer experience pathological thoughts, user reactions to their newfound states of mind are mixed. Some people feel like they're finally free to be their true selves now that they're freed from foreign impulses. Other users find themselves questioning their identity and autonomy while on Smart DBS. Most of the second group's concerns are new versions of old questions raised by psychotherapeutics generally. For example, "Am I still me if a pill or a program is modifying my pathological thoughts?" And "To what extent can my choices be said to be 'freely willed' if they are the product of chemical or electrical processes altered by my treatment?"

In addition to these perennial concerns, Smart DBS adds a few new wrinkles arising from its mechanism of action. The technique works by learning to recognize and anticipate a building (presumably undesirable) impulse and then using the implanted electrodes to interrupt it before the patient ever becomes aware of the impulsive thought. One concern this has raised among some Smart DBS users is that *there are some thoughts they are no longer capable of having* while on Smart DBS. Not only that, but they have no way of introspectively recognizing what those inaccessible thoughts are or when they would have arisen.

This is compounded by the nature of the machine learning on which Smart DBS relies to recognize pathological thoughts and preempt them.[4] Machine-learning systems work by identifying patterns in oceans of data. Using those patterns, they hazard answers to fuzzy, open-ended questions. Provide a neural network with labelled pictures of cats and other, non-feline objects, and it will learn to distinguish cats from everything else; give it access to medical records, and it can attempt to predict a new hospital patient's likelihood of dying. Give it brain states and resulting actions, and it can associate brain states with compulsive actions. And yet, most machine-learning systems don't uncover causal mechanisms. They are statistical-correlation engines. They can't explain why they think someone is about to engage in a compulsive behavior, because they don't "think" in any colloquial sense of the word—they only answer. Because most machine-learning models cannot offer reasons for their ongoing judgments, there is no way to tell when they've misfired if one doesn't already have an independent judgment about the answers they provide. Misfires can be rare in a well-trained system, but how can we know if your Smart DBS implant is preemptively blocking other, non-pathological thoughts you would otherwise have had? Can someone with a Smart DBS implant be held moral responsible for anything they do when they don't even know if their action was produced by the Smart DBS or themselves?

As the team that pioneered the first successful Smart DBS therapy, you have been invited by *Science* to coauthor an editorial responding to these concerns. What are the key points you wish to raise in your piece?

E. In general, does therapeutic control of OCD symptoms increase or decrease patient autonomy?
F. Is there anything about Smart DBS's method of treatment that makes it more or less threatening to autonomy, compared to other forms of treatment (e.g. medication, cognitive behavior therapy)?

---

[4] The description of how machine learning systems operate is quoted (with contextual adaptations for this scenario) from Jonathan Zittrain, "The Hidden Costs of Automated Thinking," *The New Yorker*, 2019.

SCENARIO 3: PRIVACY & BIAS

Despite its demonstrated success treating intransigent mental illnesses like OCD and its potential to enhance the abilities of neurotypical individuals, Smart DBS remains controversial. In addition to the fears that it invites a blurring of the therapy/enhancement line and that it may be a threat to identity and autonomy, a third set of worries involves the functional imaging data the Smart DBS algorithm relies on for its predictions. Some critics have privacy concerns about the intrusively detailed portrait of an individual's cognition the Smart DBS dataset contains. (Lending credence to this concern, your lab has been receiving an increasing number of calls from advertisers looking to use the Smart DBS dataset and algorithm to predict consumer behavior and from political consultants looks for greater insight into voter behavior.)

Additionally, concerns have been voiced about the potential for bias in Smart DBS. The algorithms your lab develops are trained on a proprietary dataset and then honed on patient-specific observations. Critics worry that the dataset may train biases into the algorithms, either because the dataset is based on an unrepresentative sample or because the algorithm ends up making predictions based on correlations caused by structural biases in society. The fact that the dataset is proprietary and that machine learning is something of a black box only magnifies these apprehensions.

In light of all these issues, your PI has been asked to help establish a regulatory board to govern the use of neuropredictive data and technologies. Given that the board is to consist of only seven members, what criteria should guide the board's composition?

G. List as many of the aptitudes, interests, identities, or other qualities board members should possess
H. Rank the five most essential traits in order of importance.

DECISION RECORD SHEETS

Decision 1: Therapy Vs. Enhancement

A. Yes, the therapy enhancement distinction matters, and therefore the team should respectfully refuse to help Dr. Singleton.
B. Yes, the therapy enhancement distinction matters, but the team should help Dr. Singleton anyway.
C. The therapy enhancement distinction does not matter and the team should help Dr. Singleton.
D. The therapy enhancement distinction does not matter, but the team should respectfully refuse to help Dr. Singleton anyway.

Option	# of Votes	Reasoning
A		
B		
C		
D		

Decision 2: Autonomy

E. In general, does therapeutic control of OCD symptoms increase or decrease patient autonomy?
F. Is there anything about Smart DBS's method of treatment that makes it more or less threatening to autonomy, compared to other forms of treatment (e.g. medication, intensive care programs, cognitive behavior therapy, etc.)?

Considerations counting in favor of concluding therapeutic control of OCD leads to INCREASED autonomy:

Considerations counting in favor concluding therapeutic control of OCD leads to DECREASED autonomy:

Conclusion E:

Therapeutic control of OCD symptoms [DOES | DOESN'T] increase patient autonomy.

Considerations in favor of concluding Smart DBS is MORE threatening to autonomy than alternative therapies:

Considerations in favor of concluding Smart DBS is LESS threatening to autonomy than alternative therapies:

Conclusion F:

Therapeutic control of OCD symptoms **with Smart DBS** [DOES | DOESN'T] increase patient autonomy.

G. List as many of the aptitudes, interests, identities, or other qualities board members should possess
H. Rank the five most essential traits in order of importance.

	Desired qualities for board members
Aptitudes	
Interests	
Identities	
Other	

Top five qualities board members should possess:

1. ____
2. ____
3. ____
4. ____
5. ____