"Can I Trust the Device in My Head?" Closed-loop DBS and Navigating Threats to Agency

Timothy Emmanuel Brown 1

1 Department of Philosophy, University of Washington

The Promise of Closed-Loop Stimulation

Deep-Brain Stimulators (DBS) have become an effective means of treating a variety of neurological conditions: Parkinson’s, essential tremor, dystonia, epilepsy, and (experimentally) mood disorders.

Most current systems are open-loop (OL-DBS): they apply stimulation constantly, at a level determined by clinician and patient. Next-generation Closed-Loop DBS (CL-DBS) read signals from the nervous system and use them to adjust stimulation. [1]

These new systems could personalize therapy, improve treatment effectiveness, and reduce battery consumption.

CL-DBS in the "decisional loop"?

Where OL-DBS relies on a clinician to make decisions about stimulation parameters, CL-DBS places autonomous systems in patients’ decisional loops. [2][3]

These systems are “in the loop” insofar as they:

- Record/monitor neural activity in that are active in ways that track symptoms of conditions or users signals.
- Analyze and classify neural recordings according to predefined categories / criteria.
- Initiate or adjust stimulation parameters to treat symptoms.

Can users trust CL-DBS?

CL-DBS may make life difficult for patients by acting—in the above three domains—in ways that are at odds with user’s interests or deeply-held beliefs.

CL-DBS should operate with an "enduring commitment to acting in a morally respectful way toward us," with “actions to accord with that commitment.” [4]

Evaluating the impact of CL-DBS

A variety of oppressive structures may play a role in way CL-DBS behaves—causing it to betray its users. We must ensure that CL-DBS behaves in a way that minimize the possibility of moral harm.

We must ask:

- What areas of the brain should not be targets? What types of data should not be recorded?
- Who/what should have access to these data once they are collected? —and for what?
- Will algorithms classify data using essentializing categories that harm marginalized people?
- Will data be analyzed (or will stimulation be initiated) in ways that are the result of systemic biases?
- Will CL-DBS’ treatment decisions harmfully constrain or complicate users’ agency?
- Will users’ had difficulties with the experience of real-time stimulation adjustment?

References