Prindle AI Ethics
Depauw University
April 2024

Taraneh Wilkinson

BCIs, Relational Ethics, and the Habitus of Ableism

In late 2022 Elon Musk's Neuralink disseminated a video of a macaque monkey playing a video game using a device surgically implanted in its brain. As Musk described it, it was like "replacing a piece of skull with, like, you know, a smart watch" (Wall Street Journal Video 2022). The company's interest in animal testing was to engineer new technology for human prostheses. Yet, many who watched the video wondered about the future of cognitive autonomy. Further, while news of the Neuralink monkey sparked conversation as to the future applications of brain-computer interface technologies (BCIs), the ultimate fate of the macaque test subjects was shadowed by rumors of death caused by the BCI implants, rumors that Musk strongly denied. The FDA initially denied approval for testing of the implant on human subjects, but in May 2023 Neuralink received a green light (Wired 2023).

The Neuralink example illustrates two ways in which progress in BCI technology deserves our ethical scrutiny. One, BCIs are increasingly in a position to break down Cartesian barriers between mind and matter, challenging assumptions about the limits of technology to shape our inner mental worlds. Although smart phones, the internet, and mass entertainment certainly have an outsized influence on our mental worlds, we have only just begun to imagine the possibilities and risks of allowing human-made technology to bypass our digits and voice signatures and plug directly into the brain. Two, as with much of the hype around AI these days, the lower tech costs of progress, such as the suffering of animal test subjects, is often overlooked in the fervor to imagine the future of new technological horizons.

Abeba Birhane has tried to bring greater attention to the human costs of the wider industry and its lack of regulation, often overlooked in sci-fi friendly speculation about the future of artificial cognition and the fates of those most likely to benefit from the positive gains of advanced AI. This paper applies Abeba Birhane's relational ethics approach to an ethical reflection on the ableist assumptions of many developing BCI technologies. I argue that, while there are valid rehabilitation applications of BCIs, they also present a growing risk of perpetuating and deepening ableist discrimination of non-normate bodies and brains. To do this, I draw on Arseli Dokumaci's concept of the "habitus of ableism" (2023).

In what follows, first I present Birhane's relational ethics as a criterion for ethical investigation of AI technologies. Then, I take up Dokumaci's habitus of ableism, applying it to different cases of recent BCI applications. I make the case that, even in a world invested in the systematic regulation of AI technologies, the lack of situational and ethical awareness generated by the habitus of ableism render it all the more likely that some risks of using BCIs will be overlooked.

Birhane's Relational Ethics and Algorithmic Fairness

Birhane (2021) argues that recent work on algorithmic fairness, while important, is not enough to address the systemic biases that impact and guide our use of AI by simply addressing bias in datasets. Beyond the datasets and algorithms, more work is needed to expose and address biased assumptions that shape data sets, along with the lived histories of injustice and ongoing power asymmetries.

Birhane posits relational ethics as a way of centering the needs and concerns of those most impacted by the recent trends in machine learning, moving away from the pernicious false neutrality of the technocratic problem/solution approach. Relational ethics and the ongoing critical reflection on systemic injustice in technology, on this view, is a skilled habit that must be intentionally practiced. I will argue that practicing this skill is a necessary (but not sufficient) practice to combat the potential deleterious effects of the habitus of ableism embodied in powerful and rapidly developing AI technologies.

BCIs and Risks to Health and Fairness

BCI technologies as we know them today got started at UCLA in 1973 (Nam et al., 2022, p. 35). BCI technologies can be invasive or non-invasive, i.e. implanted in the brain like a chip or fitted over the head like an EEG cap. Today, BCIs rely heavily on models and methods in artificial intelligence to read brain waves, usually EEG, to classify and interpret biometric signals from the brain. BCIs have numerous potential applications. These range from a monkey playing a computer game with its brain (Nam et al., 2022) to strategic interventions in rehabilitation (Butt et al., 2022). BCI use now encompasses both medical and non-medical applications from neuro-rehabilitation to human cognitive augmentation (Gao et al., 2021). Many BCIs involve the use of EEG with motor imagery techniques to control external devices such as spelling devices and wheelchairs (Subasi et al., 2022).

While some of these technologies may prove effective for rehabilitation, they also risk exacerbating existing social disparities. For instance, even in the best-case scenario of a BCI that has proven itself as safe and effective tool for neuro-rehabilitation, there are still questions of equal access to such technology. Beyond that, however, it is important to look at how the rise of such technologies fuel what Dokumaci (2023) has dubbed a "habitus of ableism," increasing the possibility and expectation of bodily and neurological homogeneity, often to the detriment of those already marginalized by ableist norms.

I now offer an account of Dokumaci's habitus of ableism. Using examples of recent BCI technologies, I will then sketch how the habitus of ableism may apply to the use of AI-guided BCIs.

Habitus of Ableism

Arseli Dokumaci argues for an ecological understanding of disability. She introduces the "habitus of ableism" to capture how social norms make invisible alternate ways of moving through the world (2023, 21). These alternative ways of inhabiting space are deemed less likely and relegated to matters of access, rather than seen as valid trajectories of embodied living. Not only that, the habitus of ableism is a socially embodied practice that appears necessary and inevitable to those caught up in it (77). Think of modern travel. We have grown so accustomed to moving by car or plane to different parts of the world or even our towns and cities, that for many, if not most of us, the idea of travelling across a country on foot or through town on horseback are quasi-unthinkable oddities. This tendency goes for how we eat, how we dress, and what are expectations for movement are as we complete other mundane tasks at home or in public.

On this view, there is one environment, but "multiple ontologies" or ways of existing meaningfully in the environment (43). Bodies that are differently abled are active in carving out their own niches for meaningful embodied existence through what Dokumaci calls "activist affordances." Activist affordances capture the way differently abled bodies resist the limiting effect of the habitus of ableism and find ways to move through the world despite the world being tailored to normate bodies. However, due to the outsized influence of mainstream or normate ways of being in the world, often the livable ecological niches hard won by those differently abled are destroyed through the habitus of ableism.

Combatting the habitus of ableism is not a simple matter of shedding biases. Following Pierre Bordieu, a habitus is something lived, something in the body, something that we are not usually consciously aware of. At the level of conscious awareness, there may be no intention to overlook accessibility needs, but the habitus of ableism blinds normate bodies to the embodied experiences and needs of others. Its effect is to render even minimal accessibility requirements as structural extravagances in the eyes of much of the public and city planners:

> The habitus of ableism is what prevents city officials, store managers, or any other agent from noticing the lack of access or even form being able to identify this lack as a problem in the first place. (Dokumaci 2023, 82)

In one example, Dokumaci tracks the daily habits of a blind inhabitant of Montreal who regularly uses public transportation. Although he walks with a cane to allow him to navigate the subway safely without vision, those around him thoughtlessly invade his space, curbing his ability to effectively use his cane to find his way around the station (72). The issue was not the lack of accessible space, but the fact that normate bodies were taking up his space in such a way that prevented him from being able to move in it effectively.

How does this apply to BCIs? BCI applications have largely focused on the use of prostheses and forms of neuro-rehabilitation. In other words, BCIs are being designed for non-normate bodies and embodied minds. What Dokumaci's habitus of ableism tells us is that a good attitude is not sufficient to prevent harm or injustice to non-normate embodied experience. Further, it should warn us that being shaped by the habitus of ableism may even blind us to the needs, injustices, and ecological niches of others. For a technology whose limits and applications are still unknown, adding this additional layer of unknowing sharpens the call for ethical scrutiny.

BCIs as Potential Perpetuators of the Habitus of Ableism

BCIs show great medical potential, but they also raise ethical concerns. With the habitus of ableism in mind, I will offer several examples of how deep learning applied to both EEG and fMRI that may allow us to make significant advances in healing and helping sectors, such as "brain prints" for multi-factor log-ins (Bialas et al., 2022), use of EEG for prostheses (Butt et al., 2022), increased empowerment of tetraplegic patients (Korik et al., 2022), and generating images from fMRIs. Then I will draw on Dokumaci's habitus of ableism to point out several potential risks that go along with AI in BCIs. By framing the growing application of BCIs as an ethical and societal question, my position, following Birhane, asks the reader to look not only at the technology itself, but also the kinds of goals, assumptions, and risks that go with said technology.

BCIs and Prosthetics

BCIs have wide and growing applications in prosthetics. For instance, AI can be used to understand the correlation between EEG and signals sent from the plantar nerve and a prosthetic correlate (Butt et al. 2022). This is an example of a positive benefit of using AI to read brainwaves. Applying AI and BCI to prostheses could mean developing prostheses that deliver realistic pressure signals to amputees' brains, improving function and reducing trauma. This sounds like a winning situation.

Yet, it is important to consider the broader range of social expectations around prosthetics. While many prostheses could be a form of activist affordance, and a life affirming mode of moving through the world, not all prostheses necessarily serve such a purpose. Some prostheses serve a social function in rendering a differently embodied experience invisible. In the famous example of Audre Lord's *Cancer Journals*, Lorde recounts her own decision not to wear a prosthesis after mastectomy and the backlash she received from healthcare providers and society for that choice. She recounts her experience as one of stark ostracism:

> It hurts when even my sisters look at me in the street with cold and silent eyes. I am defined as other in every group I'm a part of. The outsider, both strength and weakness. (*Cancer Journals*, 10/3/79)

BCIs can also be used to restore or enable greater social participation in some patients. In another instance of recent BCI advancement, a tetraplegic patient trained with an EEG-based BCI to remotely control an avatar in a virtual race (Korik et al. 2022). Although the patient had trained for 10 years, performance dipped for the Cybathlon championship, likely due to the fact that stress factors shift EEG signals, making the BCI less effective under stress. This sounds like a potential opportunity for liberation, allowing new forms of social interaction and participation. Yet, again, we should still ask the question as to how the social space enabled by the BCI is structured. Who has designed this race? Would tetraplegic patients use this same technology for other social ends if allowed?

The potential applications of BCIs in the lives of those differently abled are many. Drawing on previous work on multi-scale hierarchical classification of EEG readings for BCI, there are prototype BCIs for controlling multiple devices as a potential aid for differently-abled individuals (Mehdi et al. 2022). They use a hybrid CNN convolutional filter for AI reading of EEG. The method produced a 97% accuracy rate, but there is more study needed on how to improve accuracy for practical use. As AI improves, there will be increasing accuracy in reading human biofeedback under diverse conditions. This leaves BCI engineers in a position of great responsibility. Yet, if Dokumaci is right about the habitus of ableism, many of these engineers may be blind to the actual needs of different bodies.

Even breakthroughs in the so-called mindreading capacities of BCIs might benefit those with symptoms of bodily paralysis. While proposed applications include cybersecurity authentication (Bialas et al. 2022) and generating images from brain signals (Naish 2022), it is not hard to fathom such technology being used to communicate with locked-in syndrome (LIS). However, we should also ask how devices created, trained, and standardized by normate experiences might also unintentionally shape the brain chemistry and internal experiences of those using it, particularly those who might have very few other routes for communication with others beyond the technology.

In fact, some have warned that the as yet understudied ability of BCIs to shape human brain function represents a vulnerability for totalitarianism. Rafferty (2021) draws attention to the advanced state of BCIs and their ability not only to read imagery, motor, and emotional data directly from brains, but also their ability to shape brain function. Until now, this has mostly had a therapeutic application, but the author warns of the dangers of this technology becoming a tool used to promote global totalitarianism. He dubs some of the new advancements in BCIs (but not all BCIs) as a "global existential risk." Others have deemed BCIs safe for human use but also documented various potential misuses of BCI technology for inappropriate purposes (Williams et al. 2022).

Studies even include research on the effects of BCIs on children due to their high neuroplasticity (e.g. Sun et al. 2021). Children, already recognized by IRBs as a vulnerable population, are especially vulnerable to the negative (and potentially totalizing) effects of BCI technology due to their higher neuroplasticity. They, developmentally speaking, require access to explore the kinds of "multiple ontologies" mentioned by Dokumaci. The ethical harm wrought might be more than measurable side-effects, but also the narrowing of the window of lived experience via calibration to a BCI device likely to reflect some features of the habitus of ableism.

Further Considerations: What to do with Ignorance?

One might be tempted to think that addressing bias in the training of models is a sufficient solution for preventing AI from further ensconcing social equalities. Yet, as Birhane has stressed, this downplays the human role in designing, training, and implementing AI, giving insufficient attention to AI design and implementation in various applications. Further, as Birhane notes, ongoing critical reflection on systemic injustice in technology is a skill to be practiced, not a one-time solution. As I have argued so far, Dokumaci's habitus of ableism is an important conceptual resource for cultivating said critical reflection.

If there is one theme that is both a tool for critical reflection and an obstacle, it is the theme of ignorance. Dokumaci's habitus of ableism posits that normate bodies, even well-intentioned ones, will often fail to see the embodied practices or socio-ecological niches of others, often to their detriment. This is an ethically harmful kind of ignorance. And it is not limited to Dokumaci's treatment.[1] Unfortunately, in the face of such ignorance, positive ethical solutions seem fated to miss the mark, at least sometimes. This is perhaps where Birhane's relational ethics is most salient. By working through ignorance in relation with others who share different experiences, we are more likely to cultivate the kind of awareness needed to design responsible AI-driven devices, such as BCIs.

Finally, some scholars have proposed a three-stage model for the development of BCIs (Gao et al. 2021). The first stage involves straightforward device control, i.e. direct channeling for disabled patients. This is largely what was covered here today. The second stage is "closed-loop BCI systems," which uses a closed loop between device and patient brain to not only control an external device but restore cognitive functions. This second stage was only briefly touched upon. The third stage is the generalization of integrating brain/body and AI systems in the world. This third stage is what I hope we will think about as we consider how a habitus of ableism might lead us to imagine this integration. As one might imagine from such a comprehensive model, applications in disability are only the beginning. Without careful ethical scrutiny, their potential for totalizing tendencies might not only affect the non-normate but the way each of us, regardless of ability, interface with the world.

[1] David Kinney and Liam Kofi Bright (2023) argue that privileged groups are expected to ignore information informing them of their privilege, and this is expected on Buchak's model of risk averse rationality (2010, 2013). Namely, those with privileged social status are expected to ignore information that informs them of actual social inequalities that benefit them. This holds even for those who have done work to condition themselves to accept the potential emotional cost of learning of their privilege, because on their model the ignorance stems not from intention or an active bias but from how a risk averse agent is expected to behave. If someone expects to be scrutinized less by police and other authority figures, they may engage in riskier behaviors that would otherwise have a higher expected payoff. To avoid this risk, elite agents simply "tune out" so as not to learn about social advantages that might rationally motivate them to take greater societal risks. From this they argue that informational DEI trainings are costly and ineffective.

References

Bialas, K., Kedziora, M., Chalupnik, R., & Song, H. H. (2022). Multifactor Authentication System Using Simplified EEG Brain–Computer Interface. *IEEE Transactions on Human-Machine Systems, Human-Machine Systems, IEEE Transactions on, IEEE Trans. Human-Mach. Syst*, *52*(5), 867–876.

Birhane, A. (2021). Algorithmic injustice: A relational ethics approach. *Patterns* (New York, N.Y.), 2(2), 100205-100205.

Butt, A. M., Alsaffar, H., Alshareef, M., & Qureshi, K. K. (2022). AI Prediction of Brain Signals for Human Gait Using BCI Device and FBG Based Sensorial Platform for Plantar Pressure Measurements. *Sensors* (14248220), 22(8), N.PAG.

Dado, T., Güçlütürk, Y., Ambrogioni, L., Ras, G., Bosch, S., van Gerven, M., & Güçlü, U. (2022). Hyperrealistic neural decoding for reconstructing faces from fMRI activations via the GAN latent space. *Scientific Reports*, 12(1).

Dokumaci, A. (2023). *Activist Affordances: How Disabled People Improve More Habitable Worlds*. Duke University Press.

Gao, X., Wang, Y., Chen, X., & Gao, S. (2021). Interface, interaction, and intelligence in generalized brain–computer interfaces. Trends in Cognitive Sciences, 25(8), 671–684.

Kinney, D. & Liam Bright, K. (2023). Risk Aversion and Elite-Group Ignorance. *Philosophy and Phenomenological Research.* 106(1): 35-57.

Korik, A., McCreadie, K., McShane, N., Du Bois, N., Khodadadzadeh, M., Stow, J., McElligott, J., Carroll, Á. & Coyle, D. (2022). Competing at the Cybathlon championship for people with disabilities: long-term motor imagery brain–computer interface training of a cybathlete who has tetraplegia. *Journal of NeuroEngineering and Rehabilitation*, 19(1), 1–22.

Lorde, A. (1980). *The Cancer Journals*. Penguin Books.

Medhi, K., Hoque, N., Dutta, S. K., & Hussain, M. I. (2022). An efficient EEG signal classification technique for Brain–Computer Interface using hybrid Deep Learning. *Biomedical Signal Processing and Control*, 78.

Mhlambi, S. (2020). From rationality to relationality: ubuntu as an ethical and human rights framework for artificial intelligence governance. In *Carr Center for Human Rights Policy Discussion Paper Series*, 2020-009.

Naish, J. (2022, Aug 23). Are these pictures proof that brain scans can read our minds? Researchers showed volunteers pictures of faces, while a special type of MRI translated their brainwaves into images... with uncanny accuracy. *Daily Mail* [London (UK)].

Nam, C. S., Jae-Yoon Jung, Sangwon Lee, Sanghyun Choo, Traylor, Z., Hoonseok Park, Donghyun Park, & Sangyeon Kim. (2022). Human-computer interface, deep learning and explainable AI: BCI technology based on brain waves data has varied potential uses. *ISE: Industrial & Systems Engineering at Work*, 54(6), 34–39.

Rafferty, J. (2021). Brain Computer Interfaces: A New Existential Risk Factor. *Journal of Futures Studies*, 26(2), 51–65.

Subasi, A. (2022). Artificial Intelligence in Brain Computer Interface. *2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA),* 1–7.

Sun, J., He J., Gao X. (2021). Neurofeedback Training of the Control Network in Children Improves Brain Computer Interface Performance. *Neuroscience*, 478, 24-38.

Wall Street Journal (2022, Dec 1). Neuralink Monkey Types with Brain Implant, Elon Musk Says Human Testing Coming. https://www.wsj.com/video/neuralink-monkey-types-with-brain- implant-elon-musk-says-human-testing-coming/28516D82-E6B5-4D57-88A3-F5D86BE4BC3D

Williams, S. C., Horsfall, H. L., Funnell, J. P., Hanrahan, J. G., Schaefer, A. T., Muirhead, W., & Marcus, H. J. (2022). Neurosurgical Team Acceptability of Brain–Computer Interfaces: A Two-Stage International Cross-Sectional Survey. *World Neurosurgery*, 164, e884–e898.

Wired (2023, Sep 20). Elon Musk PCRM Neuralink Monkey Deaths. https://www.wired.com/story/elon-musk-pcrm-neuralink-monkey-deaths/