From Scanning Brains to Reading Minds: Talking to Engineers about Brain-Computer Interface

Nick Merrill, John Chuang
BioSENSE, UC Berkeley School of Information
Berkeley, CA, USA
{ffff, john.chuang}@berkeley.edu

ABSTRACT

We presented software engineers in the San Francisco Bay Area with a working brain-computer interface (BCI) to surface the narratives and anxieties around these devices among technical practitioners. Despite this group's heterogeneous beliefs about the exact nature of the mind, we find a shared belief that the contents of the mind will someday be "read" or "decoded" by machines. Our findings help illuminate BCI's imagined futures among engineers. We highlight opportunities for researchers to involve themselves preemptively in this nascent space of intimate biosensing devices, suggesting our findings' relevance to long-term futures of privacy and cybersecurity.

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g. HCI): Miscellaneous

Author Keywords

brain-computer interface; mind-reading; technology probe

INTRODUCTION

In 2017, both Mark Zuckerberg and Elon Musk announced efforts to build a brain-computer interface (BCI) [22]. One blog post enthusiastically describes Musk's planned BCI as a "wizard hat," which will transform human society by creating a "worldwide supercortex," enabling direct, brain-to-brain communication [31].

A slew of inexpensive brainscanning devices underwrite such utopic visions. 2017 saw a BCI for virtual reality gaming [24] and brainwave-sensing sunglasses [30] join the already large list of inexpensive, consumer BCIs on the market [22, 18, 14]. These devices, which are typically bundled with software development kits (SDKs), shift the task of building BCIs from the realm of research into the realm of software development. But what will software developers *do* with these devices?

CHI 2018, April 21–26, 2018, Montreal, QC, Canada

ISBN 978-1-4503-5620-6/18/04…$15.00

DOI: https://doi.org/10.1145/3173574.3173897

Figure 1: A participant uses our brainwave authenticator in his startup's office.

This study employs a technology probe to surface narratives, and anxieties, around consumer BCIs among professional software engineers. We provided a working brain-computer interface to eight software engineers from the San Francisco Bay Area. As brainscanning devices become more accessible to software developers, we look to these BCI "outsiders" as a group likely to participate in the future of brain-computer interface. Specifically, we provided participants with a brain-based authenticator, an application predicated on the notion that a BCI can detect individual aspects of a person, making it a potentially fruitful window into broader beliefs about what BCIs can reveal [27, 12].

Despite heterogeneous beliefs about the exact nature of the mind, the engineers in our study shared a belief that the mind is physical, and therefore amenable to sensing. In fact, our participants all believed that the mind could and would be "read" or "decoded" by computers. We contribute to an understanding of how engineers' beliefs might foretell the future of brain-controlled interfaces. If systems are to be built that read the mind in any sense, we discuss how such systems may bear on the long-term future of privacy and cybersecurity.

BRAIN-COMPUTER INTERFACES & PATHWAYS TO BROADER ADOPTION

BCIs allow people to interact with computers without muscular action. Instead, nervous system activity is translated to a discretized (digital) signal. BCIs can be categorized broadly as invasive (requiring implantation) or non-invasive (requiring only external, removable equipment). Non-invasive, consumer BCIs, are lightweight, require minimal setup, and do not require special gels. EEG (electroencephalography) is currently the most viable choice of sensing modality for consumer BCIs [6].

Historically, researchers have conceived of BCIs as accessibility devices, particularly for individuals with severe muscular disabilities. However, accessibility devices can sometimes provide routes for early adoption, and thus broader use. Speech-to-text, for example, was once a tool for individuals who could not type; eventually, it became adopted as a tool for computer input, now commonplace in IoT devices such as Alexa and Siri. Since accessibility devices can give rise to broader consumer adoption, we ask what such a pathway might look like for brain-computer interfaces. With an expanding array of inexpensive brainscanning hardware, many of which come bundled with engineer-friendly SDKs, the pathway to a future of consumer BCI increasingly becomes a matter of software engineering.

Thus, we look to software engineers in the San Francisco Bay Area. We use these engineers as a window into broader beliefs about "Silicon Valley," a term we use here to stand in for the technical, economic and political climate that surrounds the contemporary technology industry in the area [28]. While we do not believe only Silicon Valley engineers will influence the future of BCIs, historically, these engineers have a outsized impact on the types of technologies developed for mass consumption, especially with respect to software. As BCI hardware becomes more accessible, and therefore more amenable to experimentation as software, this group once again holds a unique role in devising a consumer future for this biosensor. Indeed, the Muse, and similar devices, have robust SDKs and active developer communities that are building and showcasing BCI applications [25].

However, we did not want our subjects to have first-hand experience in developing BCIs, as we did not want them to be primed by existing devices' limitations. Instead, we selected individuals who indicated they would be interested in experimenting with consumer BCI devices in their free time. This screening was meant to draw subjects likely to buy consumer devices and develop software for them. We believed that these engineers' professional expertise in software development afford a desirable criticality around our technical artifact.

What brain scans can tell

Brain scanning holds a unique *charisma* [3], not only among researchers in related fields [27], but among non-experts as well [2]. Ali et al (2014) found university undergraduates believed a brain scanning device (a fake one, unbeknownst to them) could reveal intimate details of their thoughts, even after receiving a lecture about the limitations of brain scanning technologies [2]. In that study, participants saw scans of the *brain* as informative with regard to the *mind*, a distinct entity that is potentially more expansive than the brain [9, 15].

This entanglement of mind and brain has been explored by past work in science and technology studies. For example, Dumit's (2004) study of positron emission tomography (PET) explores utopian (and dystopian) visions of diagnosing mental illness, or even criminality, from scans of a person's brain [12]. The idea of the mind's "legibility" via computational technologies has been concretely explored by Rose (2016) [27], who ties together a number of efforts across neuroscience and cognitive science to argue that specific technical implementations from these fields (along with their rhetoric around, and beliefs about the brain) allow *the mind* to be "read" or "decoded."

However, there exists an opportunity to investigate how pervasive such beliefs are among those who are not neuroscience experts, yet nonetheless technical practitioners. Given the recent shift of brain scanning equipment from research tool to consumer electronic device, we ask what software engineers, newly able to develop applications around brain scanning, might build. Answers to this question could have far-reaching consequences, from marketing, to entertainment, to surveillance. In particular, we aim to center how engineers' ideas about the mind, especially its relationship to the brain and body, inform and constrain their beliefs about what BCIs can (and should) do.

A BCI technology probe

In this study, we use a technology probe to examine the beliefs of software engineers about what BCIs can reveal about the mind. Technology probes are functional apparati intended to both collect data *in situ* from participants, and to inspire participants to reflect on the probes, and on their beliefs more generally [17].

Probes have a long and diverse history within HCI, often referring to a variety of different practices [5]. In the context of our study, our probe seeks primarily to answer research questions, rather than to figure as one step in an iterative design process. Unlike some probes in past work ours was not intended for longitudinal deployment. Instead, we aimed to gather beliefs about particular technologies and domains through a session of open-ended interaction with a device [21].

Our probe's unfinished appearance was intended to invite critique and playful experimentation [10, 21]. However, unlike a mock-up or provocation, our probe did function as advertised, allowing participants to interact with the devices in an exploratory and unconstrained way (indeed, many engineers tested that the device's feedback was real). We designed our probe to steer participants away from providing narrow feedback about the interface at

hand, and toward sharing their broader beliefs about the brain and mind.

Brain-based authentication

Our study employs a brain-based authenticator as a research probe to elicit engineers' beliefs about BCIs (and the mind and/or brain they purport to sense). This section explains how brain-based authentication works, and why we chose this application for our study.

Authentication (i.e., logging into devices and services) entails a binary classification problem: given some token, the authenticator must decide whether or not the person is who they claim to be. These tokens typically relate to one or more "factors": knowledge (something one knows, e.g. a password), inherence (something one is, such as a fingerprint), or possession (something one has, such as a device) [8]. Brain-based authentication relies on signals generated from individual's brains to uniquely authenticate them, which has a number of potential advantages over other authentication strategies (see [23] for a review). First, brainwaves are more difficult to steal than biometrics fingerprints, which are externally visible, and left in public as one's hands touch objects in the environment. Brainwaves also change over time, making theft even less likely. Second, brain-based authentication requires no external performance, making it impervious to "shoulder-surfing attacks" (e.g., watching someone enter their PIN).

We chose to build a brain-based authenticator for our study for a few reasons. First, having participants use a functioning system helped them imagine how they might use BCIs themselves. Second, the system is a plausible one, backed by peer reviewed research, thus we expected our participants to judge its claims credible. Third, the system embeds particular assumptions about what brain scanners are able to capture. Our system embeds ideas that our Muse headset can capture aspects of individual brains that are unique; as such, we expect that a working, brain-based authenticator will encourage participants to reflect not only on how a BCI applications might be adopted by the broader public, but also on what BCIs may be able to reveal about the mind and brain, and to critically examine the limits of what BCIs in general are able to do.

BUILDING THE BCI AUTHENTICATOR PROBE

Implementation

Since we wanted our technology probe to appear portable enough for use in the real world, we decided to use a pre-existing consumer EEG device to build our authenticator. We settled on the Interaxon Muse (Figure 1), a $299 headband that can be worn easily, transmits data wirelessly, and requires no gel to maintain contact between the scalp and electrodes [18]. Using a system that required conductive gel would have signaled to the participants that the technology is still limited to lab settings, and not yet ready for the real world, which could have influenced their responses.

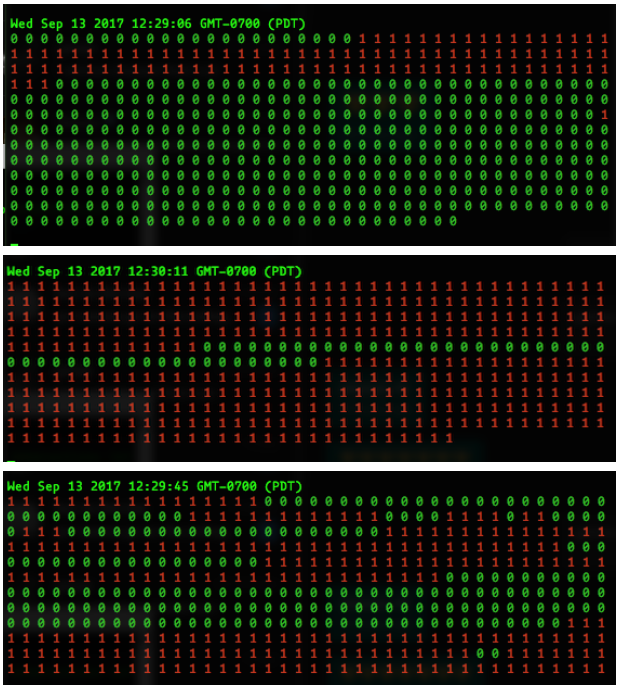

Figure 2: Our probe's visualization of 1's and 0's gave our engineers a "raw" view of the authenticator's behavior. Pictured, the UI (a) accepting someone, (b) rejecting someone, or (c) presenting mixed, ambiguous feedback.

Although the Muse's signal likely contains noise, a perfectly clean signal was not necessary to elicit beliefs from subjects in the context of our technology probe. Further, despite the Muse's small form-factor and dry electrodes, past studies have verified its signal is sufficient quality for some neuroscientific research [20].

Due to the device's battery life and intermittent connectivity when walking, the Muse headband did made a longer-term study impractical. Thus, we opted to perform a study over a short time and in a controlled environment, drawing on past technology probe studies with similar constraints [10, 19].

Data from the Muse was collected via the device's native OSC interface, and stored in a timeseries database. Queries from this database were used to provide training data for a machine learning classifier. In a preprocessing step, we performed a fast Fourier transform (FFT) to generate frequency-domain data from the time-domain data. In the machine learning step, we split a corpus of readings (and labels) into train and validation groups. Using XGBoost [7], we trained a binary classifier on seven different splits of the train group. After the classifier was produced, we validated its performance on the withheld validation set.

Given a target participant to classify, our classifier used any reading from this participant as a positive example, and any reading *not* from this participant as a negative example. Negative examples also included signals with

poor quality, and signals from which the device was off-head or disconnected. Ideally, the resulting classifier should produce "authenticate" labels when the device is on the correct person's head, and "do not authenticate" labels at any other time. This classifier could output its labels to a simple user interface (UI), described in the next section.

Interface

As the device produces data, the classifier outputs labels of "accept" or "reject." Our interface displays these labels as a square of 0s and 1s, which filled up as data from the device rolled in (Figure 2).

Several considerations motivated this design. First, the UI represents the probabilistic nature of the classification process. Individual signals may be misclassified, but over blocks of time, the classifier should be mostly correct (represented as blocks of mostly 0s by our interface). Thus our simple UI makes visible both the underlying mechanism of binary classification, and its probabilistic nature. Second, because our UI provides potentially ambiguous feedback (as opposed to unambiguous signals of "accept" or "reject"), it allows for potentially richer meaning-making and explanatory work [29]. Toward this end, the UI's real-time reactivity ("blocks" of 1s and 0s filled in over time) allows participants to experiment actively with the device, forming and testing hypotheses as to what makes classification succeed or fail.

Finally, our UI gives the probe an "unfinished" appearance. We believed this interface would cause our participants to activate their "professional vision" as tech-workers [13], and critique or test the device as if it were a design of their own. Ideally, we hoped participants would intentionally stress-test the device, or find playful ways of misusing it. These misuses could allow participants to form hypotheses about why and how the device succeeds and fails.

METHODS

We recruited participants by word of mouth. A recruitment email explained that subjects would interact with a working BCI, and be asked their opinions about the device, and about BCI broadly. We screened respondents by their current occupation and stated interest in experimenting with BCIs in their free time.

A total of eight people participated, three of which were women. Participants' ages ranged from 23 to 36. We met with subjects for a single, one-hour session in which we trained and tested a brain-based authenticator, allowing them to interact with it in an open-ended way.

These sessions were designed as a semi-structured interview, interspersed with conversation between the researcher and the participant. Our study protocol was approved by our institutional IRB. Interviews were recorded, and later transcribed. We performed an "issue-focused" analysis of the transcriptions [32], allowing topics and themes to emerge during analysis. Subjects names were changed to pseudonyms to protect their anonymity. The remainder of this section describes in detail how subjects interacted with the device during sessions.

Wearing the device

The interviewer began by explaining that participants would wear a BCI, which we would train to work as an authenticator, answering participants' questions about how the device would work. Subjects were told that they would be asked about their opinions on BCIs generally, and that their anonymized voice and EEG data would be collected.

The interviewer asked participants to place the EEG headband themselves, and to assure that the device fits comfortably, at which point the interviewer would begin recording signals from the device. Next, the interviewer would ask participants how they felt about having the EEG device on their head. This question would typically begin a short, open-ended exchange about their past experience with brain-scanning devices, and prior knowledge, if any, of BCIs. This exchange would segue into a broader discussion about the participant's use and relationship with technology, in personal and work life.

After this initial conversation, the interviewer would perform a brief *calibration* step with the participant, in which data are collected to train a custom classifier for use in authentication. Participants would perform a number of tasks, or *mental gestures*, prompted by a stimulus presentation program. These tasks provide a more diverse corpus of an individual's signals, which should enable a more robust (and accurate) classifier. After this calibration procedure, which usually lasted about ten minutes, the interviewer would perform a semi-structured interview with participants. The interviewer would continue to record data from the Muse throughout this interview.

Using the authenticator

At this point, the interviewer would explain to participants that the data collected thus far would be used to train an custom authenticator for them. The interviewer would explain roughly how the authenticator would work: the probe should *accept* readings when the participant is wearing the device, and *reject* readings in any other case.

Next, the interviewer would run a script that trained our XGBoost classifier (Section 3.1). Participants could watch the training process run, if interested (a few were). After the training process completed, the researcher would set up the UI (Section 3.2) and allow participants to view the classifier's output in real-time using live data from the participant's Muse device. Participants would then see the probe's *accept* or *reject* classifications using the live data from their headset.

After allowing participants to acclimate to the output, and answering any preliminary questions, the interviewer would encourage the participant to experiment with the authenticator, and share any impressions, reactions or ideas. The open-endedness of this session was meant to encourage participants to explore the device's capabilities

and limitations, free of particular tasks to accomplish. However, we suspected that our participant population would be particularly prone to "hypothesis-testing," exploring the devices limitations by building theories about how it might work. We structured the session around this assumption, preparing to ask participants to think aloud as they explored the device's capabilities.

After some free-form exploration (usually involving some back-and-forth with the participant), the interviewer would transition into a semi-structured interview, which would occur with the device still active. The interviewer would ask participants to unpack their experience, and lead them to explore what they felt the device could reveal about them. After some discussion, the formal interview would conclude, and the participants would remove the Muse device from their head.

EXPERIENCING THE AUTHENTICATOR

In general, we found particular reflections to come at different points in the interview protocol. Critiques (and questions) about the device narrowly tended to come as soon as engineers placed the device on their heads. Reflections on the BCI broadly, and its future trajectories, tended to come after viewing the probe's feedback for some time. As these conversations progressed, participants naturally tended to reflect on what future BCIs could do. Subjects would typically relate the capacities of the probe, and of possible future technologies, to their ideas about the mind, body or brain. The probe continued to run during these discussions. Toward the end of the interview, the researcher would prompt participants to reflect on any anxieties they might have about the future of BCIs (interestingly, only one participant raised this subject on their own). The remainder of this section is organized to roughly mirror this common order of participants' reflections during interviews.

Using the BCI probe

Our working authenticator elicited diverse reactions from the engineers in our study. Almost all participants cracked jokes after putting on the headband (three subjects commented that they felt like they were "from Star Trek"). All participants except Joanna said they would not wear the device in public, though a few conceded that they might if the headsets were more common. Terrance commented, "If most people are doing it, then it's fine. Sort of like stock speculation."

Perceptions of the authenticator's accuracy were mixed. Four participants found that the authenticator worked well for them. For these participants, the authenticator consistently rejected blocks when the headset was off of their head, or worn by the researcher.

On the other hand, four participants found the probe consistently rejected every reading, whether it came from them or the researcher (i.e., they experienced false rejections, but not false acceptances). These subjects often tried to remedy the situation by attempting tasks they had rehearsed, typically with mixed success. Most of these subjects concluded that there was not enough training data to produce reliable classification, but that such a system would work with a larger corpus. In contrast, Alex, a 30 year-old founder of an indoor agriculture startup, blamed himself, saying "I must not produce very distinguishable thoughts."

Those participants who felt the probe's authentication was reliable tended to center their explanations on why it worked. Participants who experienced less consistent accuracy with the authenticator tended to center their explanations on how the device might be improved, e.g. with better or more comprehensive sources of data. This impulse to "fix" likely speaks to our participants' general tendency to engineer working systems, which extended in our case even to this experimental technology.

As we hoped, the engineers engaged critically with the technical implementation of the probe. In general, engineers asked about the machine learning infrastructure underlying the authenticator, and several participants (particularly John, Mary and Alex) asked specific questions, and made specific recommendations, diagnosing issues with the authenticator by thinking about the diversity and size of the training set. Almost all participants noted the authenticator worked better when they were not looking at the visual feedback from the user interface. Participants generally theorized that this might occur because they were not viewing feedback when training the classifier. In these cases, the engineers appeared to apply their domain knowledge to their observations in using our technology probe.

Reflecting on the future of BCI

Our technology probe caused almost all of our participants to speculate on the future of BCIs generally. To most participants, the future of BCIs seemed to be largely pre-determined. One of our participants, Terrance (a 24 year-old software engineer at a small transportation startup), removed the headband to inspect it, and commented on its awkward visibility. In doing so, he reflected on the future of BCIs, speaking in no uncertain terms about a future of computer-mediated "telepathy."

> Things just get progressively smaller until they disappear. And one day this'll just be an implant in my brain, doing crazy things. It'll be interesting socially, how people come to terms with it, when it's just an implant, or at least very pervasive ... I could send you a message, and it could be like you're thinking it yourself, even if you're on the other side of the Bay. *(Terrance)*

Terrance believed that BCI *will* become more prevalent: not just that smaller sensors will lead to more effective or usable BCIs, but that they will also result in greater uptake of the technology. While he references the social dimension of their adoption, he indicates that people will need to "come to terms with" the developments, rather than providing direct agency to users who may choose to adopt the technology or not.

Two participants felt less sure that such a future of pervasive BCI would ever come to pass. Elizabeth, a 30 year-old front-end engineer, noted skepticism about signal quality, or usefulness outside of persons with disabilities. Mary, a 27 year-old software engineer at a large company, pointed to social reasons for her skepticism. In reflecting on the relative accuracy of the probe's authentication performance during her session, she commented that "90 plus percent" of people would be "totally freaked out" by brain-computer interfaces generally. She continued to say that companies may themselves stop BCIs from becoming too pervasive or advanced.

> I feel like those companies, even if this were feasible, there's a moral quandary they philosophically have not figured out. They will not let the research get that advanced ... I just don't imagine them being like, "okay computer, now read our brains." *(Mary)*

While the probe was effective in spurring subjects to talk about issues around BCIs, its accuracy as an authentication device did not seem to alter participants' belief in BCI's future as a widespread technology. Unsurprisingly, the four subjects who experienced reliable authenticator accuracy all expressed that BCIs would become commonplace in the future. However, only Joanna connected the device's poor performance in her session with a probability of ongoing accuracy issues for BCIs in the future. The other three subjects who felt the device did not perform accurately all offered explanations as to why, and explained that future devices would fix these issues.

Mind, brain, body

During their interactions with the probe, almost all of our subjects discussed their deeper beliefs about the nature of the mind, and its relationship to the brain and body. Since participants discussed the future trajectory of BCIs led to discussions while the probe continued to work (or fail), the subject often arose of what BCIs might be able to detect, even theoretically. As one example, John, a 26 year-old software engineer at a small chat startup, noticed that the authenticator only worked when he was speaking, but not when he was listening to the researcher. He offered an explanation for the discrepancy.

> There's probably some kind of fundamental difference between creating thoughts and consuming thoughts. You're still making thoughts, right, but it's almost like programming versus being programmed. *(John)*

When pressed on how strictly he meant his metaphor of programming, John confirmed that he meant it quite literally, saying, /"I think we are just computers that are way more sophisticated than anything we understand right now."/ We return to this strictly computational account of the mind as "just" a computer in the discussion.

Mary gave a computational account of mind that was more metaphorical than John's, drawing on comparisons between machine learning and the mind. She cited the many "hidden layers" in deep neural networks, and that, like in the brain, "information is largely distributed." While she believed deep learning models and the brain were "different systems foundationally," she said "there are patterns" that relate the two to one another, and indicated that advances in deep learning would spur a greater understanding of the brain.

Although six of our participants provided a largely computational account of mind-as-brain, not all did. Joanna, a 31 year-old engineer who previously completed a PhD in neuroscience, felt that the mind was "the part of the brain I am aware of, the part that is conscious." She believed that neurotransmitters throughout the body have a causal relationship to what happens in the mind, but do not constitute the mind themselves; the contents of mind occur physically in the brain, and the brain alone. In other words, her account is one of "mind as conscious awareness," and while unconscious phenomena affect mind (e.g. the body, environment), they are not part of the mind *per se*. Interestingly, the probe did not work well for Joanna, and she felt confident that its poor performance was due to contaminating signal from her body (a theory she tested, and validated, by moving around and observing the probe's feedback).

Meanwhile, in one subject's account, the mind extended beyond the confines of the body. Terrance felt that there was "no meaningful difference" between the body and brain, nor between the body and the physical environment at large, saying that "you can't have one without the other." He believed that all three of these entities constitute the mind in a mutually-dependent way. However, Terrance indicated that the mind is still strictly physical, as are these three entities. Although Terrance did not provide details on how exactly the mind extended beyond the body, it is interesting to note this position's similarities to Clark's (2013) account of the extended mind [9], or Edward Hutchins's work on distributed cognition [16], though Terrance was familiar with neither.

Participants also offered differing levels of confidence in their beliefs about the nature of the mind. Joanna (who has a background in neuroscience) reported that "we do not know everything we need to know" about how the mind works. Three other subjects reported similar beliefs. However, those subjects with a computational account of mind tended to feel more confident that their account was substantially accurate.

> I think the consensus is that the body is mostly like the I/O of the brain. *(John)*

John's account here implies that a sufficiently high-resolution brain sensor would accurately capture all of a person's experiences. John confirmed this explicitly, saying *"if you could 3D print a brain, and apply the correct electrical impulses, you could create a person in a jar."* In this computational metaphor of I/O (input/output), the body itself does not have agency; instead, the body actuates the brain's commands (output), and senses the environment, sending data to brain for processing (input).

Reading the mind

As discussed in the previous section, every participant's account of mind was strictly physical, rooted mostly in the brain, in a few cases in the body, and in one case extending beyond the body to the physical world. With this physical understanding of the mind, it is not overly surprising that all participants believed it would someday be possible for a computer to read or decode the contents of the human mind. No participants expressed hesitation when asked about such a proposition.

For example, Alex did not feel comfortable providing a specific physical locus for the mind. Although he did not feel the probe was accurate for him, he took great pains to express his belief that such a device could work, though not necessarily by sensing the brain.

> We're driven by single-celled organisms in ways we don't really yet understand, but... there's got to be some sort of physical storage of memories or experiences. We just haven't quite learned how to read it yet. *(Alex)*

Though it leaves open room for a variety of interpretations about the exact nature of mind, Alex's view is explicit that thoughts are physical, therefore *can* be read, and *will* be read with some future technology.

There was a great deal of heterogeneity in the way this belief was bracketed or qualified. Joanna felt that there would "always be parts of the mind that can't be seen." She likened the question to the way that other people can know some parts of another person's mind, e.g. through empathy; their perspective, however, would always be partial, and she felt the same would be true for machines.

However, some participants did not bracket their belief that machines would someday read the mind. Participants for whom the authenticator worked reliably typically said that a mind-reading machine was "absolutely possible" (Mary) or "just a matter of the right data" (Alex). Participants who did not feel the authenticator was accurate described current state-of-the-art as "crude" (John) or "low-granularity" (Elizabeth).

Even Terrance, who believed the mind extended beyond the confines of the body, felt that the mind was readable by machine. After he stated his personal belief in a mind that extended to the physical environment, the experimenter asked what consequence this belief might have for the future of BCIs.

> Practically, it has no implication. We could still devise an authentication tool that does the job, and it doesn't matter. Maybe in some way there could be this ESP thing where you could somehow read my thoughts... If we want to do something, we will find a way. *(Terrance)*

Terrance's language here belies broader narratives of positive technological progress (notions of "[moving] forward," and that "we will find a way"). Despite his personal beliefs about the "true" nature of the mind, he felt that engineers would manage to build the systems they intended to build, even ones with a much higher specificity than those available today (e.g. an "ESP device").

BCIs for everyone?

Generally, participants stated (implicitly or explicitly) that BCI technologies would become smaller, less expensive, more accurate, and therefore become prevalent as a consumer device. Only Mary raised the question of how institutions exert agency over the artifacts they create. Where most subjects indicated BCIs become smaller and thus more pervasive, Mary indicated that companies have beliefs, which affect what devices and technologies they produce. Specifically, Mary spoke of a "quandary" between advancing technology on one hand, and systems' autonomy on the other. She viewed this reluctance to allow systems to become more autonomous as a signal that certain technologies, potentially including BCIs, may *not* be developed for ethical, moral or philosophical reasons.

Interestingly, the other seven engineers in our study expected a future in which BCIs are pervasive, in spite of their unwillingness to wear our probe's headband in public. Some subjects believed the device's awkward, outward visibility might be mitigated by future miniaturization. Other subjects felt that social norms may simply change if the device became pervasive. This latter attitude is reminiscent of those around Google Glass, which shared an awkward (and, in practice, often stigmatizing) visibility [33]. Future work might draw out the relationship of Google Glass's imagined future to that of BCI, perhaps as a way of learning lessons about possible commercial failures, and how engineering communities may have failed to foresee them.

BCI anxieties

An important counterpoint to emerging technologies is the anxiety that rises along with them [26]. Interestingly, engineers in our study expressed no strong anxieties regarding the development of BCIs, for the most part. Regardless of their experiences with our probe, participants felt that BCIs would be developed, and would improve people's lives. Participants mentioned domains such as work, safety, and increased convenience in the home.

Only Mary reported existential anxiety about the possibility of machines that could read the human mind. She reported a technology to be "absolutely possible," and referenced the probe's continuing high accuracy as we spoke. However, in stark contrast to Terrance, Mary *feared* such a development would occur sooner rather than later.

> I hope it's fifteen years out, but realistically, it's probably more like ten. *(Mary)*

Despite Mary's prior statement about the power of institutions to change the course of technical developments, here she seems to indicate that such course changes will

not occur, or that they will converge on machines that can read the mind.

When pressed on downsides, the participants who did not volunteer any anxieties about BCI initially did mention security (especially the "leaking" of "thoughts") as a concern. For example, Elizabeth did not report any particular anxieties about BCIs in general, "if the proper protections are in place." Pressed on what those protections might look like, she cited encryption as a solution to privacy concerns. Terrance, who expressed wanting BCIs to become more widespread, described in deterministic terms the cybersecurity issues such devices might pose.

> If there are security holes - which there almost certainly will be - then what happens when I'm leaking my thoughts to someone? What if I'm thinking about the seed phrase for my Bitcoin wallet... and then you put it in this anonymized dataset ... and I lose all my coins? What then? *(Terrance)*

Even alongside his concern, Terrance very much wanted a mind-reading machine to exist. He mentioned a desire for a programming assistant that would somehow speed up the process of software development. Since Terrance's conception of BCI presents high stakes with regards to privacy and security (he variously mentioned "telepathy," and an "ESP device," implying a high degree of specificity with regard to what BCIs can resolve), it is telling that he thought primarily of using BCIs to become a more efficient engineer, rather than concerns around privacy or potential harm. Later in the discussion, we unpack further how larger cultural tendencies in Silicon Valley might shape the way engineers build BCI systems.

DISCUSSION

We find that engineers hold diverse beliefs about what the mind is, what the brain is, and about the relationship between these entities. However, all of these engineers shared a core belief that the mind is a physical entity, one that machines can and will decode given the proper equipment and algorithms (Section 6.1). Despite this belief, engineers did not largely express concerns about privacy or security (Section 6.2). As BCI startups continue to grow, we propose further work within technical communities, with a sensitivity toward emerging narratives, so that we may instill criticality among this emerging technical practice (Section 6.3). We conclude with avenues for future work focusing on different communities of technical practice (Section 6.4).

Physical mind, readable mind

Although our engineers broadly believed BCIs would become pervasive as consumer devices, we found no consistent visions of what such a future might look like. Instead, and to our surprise, we found a shared belief that there exists a physical mind that can be "read" or "decoded" by machines, despite participants' heterogeneous beliefs about its exact nature. Interestingly, only one participant shared any anxiety about this prospect with the researchers; the other participants reported looking forward to such a possibility.

Crucial to beliefs about the machine-readable mind were frames of the mind as physical, and therefore amenable to sensing. In many cases, subjects would use analogies to computation in making this point. For example, John observed an anomaly in the authenticator's performance (it did not work when he was listening to the experimenter speak). He theorized that the states are distinguishable, because speaking "is like programming" and listening to someone speak "is like being programmed". In this case, John's observations about the BCI met with his pre-existing notions of the mind, producing a hypothesis for what "brain states" might exist *and* what states Muse headset might be able to detect. Hypotheses such as these could be consequential, as they might provide ideas or starting points for engineers looking to build systems. Our results highlight the importance of both pre-existing beliefs and particular interactions with BCIs in structuring engineers' understandings.

Broadly, engineers' beliefs about the mind-as-computer metaphor (Section 5.3) could provide starting points for engineers to build BCIs in the future. This computational view of mind has been popular among engineers at least since the "good old-fashioned AI" (GOFAI) of the 1950s. While much work has critiqued this stance from various angles [1, 15], those same critiques have acknowledged the role these metaphors have played in the development of novel technologies: If the mind is a machine, then those tools used to understand machines can also be used to understand the mind. Here, we see this metaphor return, its discursive work now focused on biosensing rather than on artificial intelligence. Of course, these metaphors illuminate certain possibilities while occluding others [15]. As such, future work should follow past research [1] in understanding what work this metaphor might do in its new domain of computational mind-reading.

Even those participants who did not subscribe to computational theories of mind still believed the mind to be strictly physical. These subjects all agreed that computers could someday read the mind, precisely because of its physical nature. While our results indicate that engineers believe the mind to be machine-readable, some work indicates that non-engineers may share this as well [2]. Future work could further investigate this claim more deeply in the context of consumer BCIs. If so, a machine designed by engineers and purported to read the mind might find acceptance among a broader public audience.

Those subjects with a computational account of mind tended to feel more confident that their account was substantially accurate. John referenced "the consensus" in justifying his beliefs about the mind being equivalent to the brain. It is worth asking whose consensus this might be: that of neuroscientists, philosophers of mind, cognitive scientists, or engineers? In any of these cases, engineers' confidence in their beliefs could have implications for what types of systems are considered buildable,

and where engineers might look to validate their implementations. As products come to market, professionals in the tech industry must find ways of claiming their devices to be legitimate, or working, to the public (consumers), to potential investors, and to other engineers. These claims of legitimacy could prove to be a fruitful window for understanding the general sensemaking process around these devices as their (perceived) capabilities inevitably evolve and grow alongside changing technologies.

A future for privacy and security

Since the engineers in our study believed the mind to be readable, an important question remains around the consequences for the future of consumer privacy and security. Our participants largely acknowledged that "leaking" thoughts through security holes was a valid concern, and one participant claimed that these exploitable holes will "almost certainly" exist. However, the types of threats that engineers referenced may not square with the notion of BCIs as a device for the masses. For example, Terrance's concern about someone stealing his Bitcoins through some BCI-based attack involves a technology which for now remains niche. This imagined scenario demonstrates how the security (and privacy) concerns of engineers may not match that of the general public. Such mismatches could have consequences for the types of systems that are designed, and whose needs these systems will account for.

Crucially, discussions about privacy and security concerns did not cause any participants to reflect further on the consequences of pervasive BCIs, nor did they deter enthusiasm for the development of these devices. These findings indicate either that engineers are not be inclined to prioritize security in the systems they build, or that they have resigned themselves to the inevitability of security holes in software. In either case, our findings suggest a long-term direction for cybersecurity concerns. These devices carry potentially serious security and privacy consequences. If our engineers will try to build devices that make judgments about the inner workings of a person's mind, future work must critically examine how to protect such systems, and the people who use them.

Implications for the design of mind-reading machines

Our findings do not indicate a singular path for the future of BCIs. Instead, they indicate an undercurrent of belief among Silicon Valley engineers in the possibility of technologies that can read the contents of the human mind. Crucially, our study revealed narratives not just around BCIs, but around the nature of the brain and mind generally, which in turn legitimize narratives about the possibility of mind-reading machines.

Despite these beliefs about what BCIs are capable of, only one participant in our study reported that ethical issues around privacy or security might deter their development. We hope engineers will become more reflexive about these beliefs around BCI, and more critical about their downstream potential for harm (e.g. surveillance). Much as utopian dialogues around the potential of the World Wide Web missed risks to privacy and security, so might similarly utopian ideals of mind-reading machines.

Since the engineers in our study believed BCIs could perform this potentially invasive "mind-reading," why did they largely want such BCIs to be built? Explanations might be found by relating the narratives we uncover to existing social and economic value systems within Silicon Valley communities. Biohacking, for one example, has become an established part of Silicon Valley culture, through dieting (e.g. Soylent, fasting), or more extreme forms of body modification (e.g. chipping) [11]. Underlying all of these cultures is a mechanical model of the body, which facilitates notions of optimization and experimentation. How might BCIs (especially ones that purport to read thoughts) work their way into these already-established cultural patterns? We note that existing consumer BCIs already situate themselves in this context: the Muse headset we used in this study markets itself primarily as a meditation trainer (its advertising copy claims to "remove the uncertainty from meditation") [18]. Examining how BCIs perform discursive work in engineering communities will allow us to better understand engineers' intents as these devices begin to emerge, and help us trace these intents forward as devices are re-imagined, remixed and repackaged for other groups of users in the future.

In the nascent field of consumer BCI, researchers and designers should remain in touch with the beliefs of engineers. We pinpoint beliefs about the mind, and its readability by emerging biosensing devices, as especially an critical facet. Doing so will allow design to remain preemptive rather than reactive as software for consumer BCI emerges. Designers and researchers must not remain on the sidelines; as devices come to market, we must become actively engaged in engineers' beliefs (and practices). These systems hold the potential for exploiting an unprecedented level of personal data, and therefore present an high potential for harm. As such, the area presents a new locus for researchers and designers to engage critically with technical developments.

Future work

Software engineers are a diverse group, and the geographic confines of Silicon Valley do not describe all communities worldwide. Future work could explore communities in different places. Engineers in non-Western contexts may hold different cultural beliefs about the mind, which could lead to vastly different findings.

Professionals who work in machine learning could present another participant pool for future work. Machine learning is a critical component of BCIs, and many contemporary techniques, particularly deep learning, use neural metaphors to interpret and designing algorithms [4]. Thus, practitioners of these techniques may be inclined to draw metaphors between the brain and the algorithms they employ, which could color their understanding how and why BCIs work or fail.

Future work could allow participants to take an active, participatory role in the analysis of their data, and/or in the design of the BCI system. Although our participants had the technical expertise required to perform data analysis and systems engineering themselves, we did not have participants do any such analysis for this study. This participatory approach will also help us expand our understanding from engineers' beliefs to engineers' practices, as they relate to the emerging domain of consumer brain-computer interfaces. Participants might form their own interpretations of what the data mean (or can mean), building understandings that could differ from those we observed in this study.

CONCLUSION

As engineers in the San Francisco Bay Area, the participants in our study sit at an historical site of techno/political power. Our technology probe indicates these engineers believe the mind is physical, and therefore amenable to sensing. What are the consequences for the rest of us? We hope our study will encourage engineers to closely examine the potential of these devices for social harm, and encourage researchers to remain closely attuned to this emerging class of consumer biosensor.

ACKNOWLEDGMENTS

We would like to thank the reviewers, members of the Berkeley BioSENSE group, for their comments and suggestions. We also thank Morgan Ames, Richmond Wong, Noura Howell and Anne Jonas for their invaluable feedback. This work was supported in part by the Berkeley Center for Long-Term Cybersecurity and the William and Flora Hewlett Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Hewlett Foundation or Center for Long-Term Cybersecurity.

REFERENCES

1. Phil Agre. 1997. *Computation and Human Experience*.
2. Sabrina S Ali, Michael Lifshitz, and Amir Raz. 2014. Empirical neuroenchantment: from reading minds to thinking critically. *Frontiers in human neuroscience* 8, May (may 2014), 357. DOI: http://dx.doi.org/10.3389/fnhum.2014.00357
3. Morgan G. Ames. 2015. Charismatic Technology. *Proceedings of the 5th Decennial AARHUS Conference* (2015), 109–120. DOI: http://dx.doi.org/10.1080/19447014508661941
4. Jimmy Ba, Geoffrey E Hinton, Volodymyr Mnih, Joel Z Leibo, and Catalin Ionescu. 2016. Using fast weights to attend to the recent past. In *Advances In Neural Information Processing Systems*. 4331–4339.
5. Kirsten Boehner, Janet Vertesi, Phoebe Sengers, and Paul Dourish. 2007. How HCI interprets the probes. *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07* (2007), 1077. DOI:http://dx.doi.org/10.1145/1240624.1240789
6. Francesco Carrino, Joel Dumoulin, Elena Mugellini, Omar Abou Khaled, and Rolf Ingold. 2012. A self-paced BCI system to control an electric wheelchair: Evaluation of a commercial, low-cost EEG device. In *2012 ISSNIP Biosignals and Biorobotics Conference: Biosignals and Robotics for Better and Safer Living, BRC 2012*. 1–6. DOI: http://dx.doi.org/10.1109/BRC.2012.6222185
7. Tianqi Chen and Carlos Guestrin. 2016. XGBoost : Reliable Large-scale Tree Boosting System. *arXiv* (2016), 1–6. DOI: http://dx.doi.org/10.1145/2939672.2939785
8. John Chuang, Hamilton Nguyen, Charles Wang, and Benjamin Johnson. 2013. I think, therefore I am: Usability and security of authentication using brainwaves. In *International Conference on Financial Cryptography and Data Security*. 1–16. DOI:http://dx.doi.org/10.1007/978-3-642-41320-9_1
9. Andy Clark. 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *The Behavioral and brain sciences* 36, 3 (2013), 181–204. DOI: http://dx.doi.org/10.1017/S0140525X12000477
10. Laura Devendorf, Joanne Lo, Noura Howell, Jung Lin Lee, Nan-Wei Gong, M Emre Karagozler, Shiho Fukuhara, Ivan Poupyrev, Eric Paulos, and Kimiko Ryokai. 2016. "I don't want to wear a screen": Probing Perceptions of and Possibilities for Dynamic Displays on Clothing. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. 6028–6039. DOI: http://dx.doi.org/10.1145/2858036.2858192
11. Markéta Dolejšová and Denisa Kera. 2017. Soylent Diet Self-Experimentation: Design Challenges in Extreme Citizen Science Projects. (2017), 2112–2123. DOI:http://dx.doi.org/10.1145/2998181.2998365
12. Joseph Dumit. 2004. Picturing Personhood: Brain Scans and Biomedical Identity. *Information Series* (2004), 251. DOI: http://dx.doi.org/10.1353/bhm.2005.0063
13. Charles Goodwin. 1994. Professional Vision. *American Anthropologist* 96, 3 (1994), 606–633. DOI: http://dx.doi.org/10.1525/aa.1994.96.3.02a00100
14. Mick Grierson and Chris Kiefer. 2011. Better brain interfacing for the masses. In *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems - CHI EA '11 (CHI EA '11)*. ACM Press, New York, NY, USA, 1681. DOI:http://dx.doi.org/10.1145/1979742.1979828
15. N. Katherine Hayles. 1999. Contesting for the Body of Information: The Macy Conferences on Cybernetics. In *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. 50–83.

16. Edwin Hutchins. 2005. Distributed cognition. *Cognition, Technology & Work* 7, 1 (2005), 5–5. DOI: http://dx.doi.org/10.1007/s10111-004-0172-0

17. Hilary Hutchinson, Benjamin B Bederson, Allison Druin, Catherine Plaisant, Wendy E. Mackay, Helen Evans, Heiko Hansen, Stéphane Conversy, Michel Beaudouin-Lafon, Nicolas Roussel, Loïc Lacomme, Björn Eiderbäck, Sinna Lindquist, Yngve Sundblad, Bosse Westerlund, Benjamin B Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, Nicolas Roussel, and Björn Eiderbäck. 2003. Technology probes: inspiring design for and with families. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)* 5 (2003), 17–24. DOI: http://dx.doi.org/10.1145/642611.642616

18. Interaxon. 2017. Muse: The brain sensing headband. (2017). http://www.choosemuse.com/

19. Katherine Isbister, Kia Höök, Michael Sharp, and Laaksoñahti Jarmo. 2006. The Sensual Evaluation Instrument: Developing an Affective Evaluation Tool. In *CHI 2006 Proceedings*. 1163–1172. DOI: http://dx.doi.org/10.1145/1124772.1124946

20. Olave E. Krigolson, Chad C. Williams, Angela Norton, Cameron D. Hassall, and Francisco L. Colino. 2017. Choosing MUSE: Validation of a low-cost, portable EEG system for ERP research. *Frontiers in Neuroscience* 11, MAR (2017). DOI: http://dx.doi.org/10.3389/fnins.2017.00109

21. Lucian Leahu and Phoebe Sengers. 2014. Freaky: performing hybrid human-machine emotion. *Proceedings of the 2014 conference on Designing interactive systems - DIS '14* (2014), 607–616. DOI: http://dx.doi.org/10.1145/2598510.2600879

22. Steven Levy. 2017. Brain-Machine Interface Isn't Sci-Fi Anymore. (sep 2017). https://www.wired.com/story/brain-machine-interface-isnt-sci-fi-anymore/

23. Nick Merrill, Max T Curran, and John Chuang. 2017. Is the Future of Authenticity All In Our Heads? Moving Passthoughts from the Lab to the World. *New Security Paradigms Workshop (NSPW)* (2017).

24. Neurable. 2017. Neurable: Power Brain-Controlled Virtual Reality. (2017). http://www.neurable.com/

25. NeurotechX. The international neurotechnology network. (????). http://neurotechx.com/

26. James Pierce. 2017. Dark Clouds , Io $ #! + , and ? [Crystal Ball Emoji]: Projecting Network Anxieties with Alternative Design Metaphors. (2017), 1383–1393.

27. Nikolas Rose. 2016. Reading the human brain: How the mind became legible. *Body & Society* 22, 2 (jun 2016), 140–177. DOI: http://dx.doi.org/10.1177/1357034X15623363

28. AnnaLee Saxenian. 1996. *Regional advantage*. Harvard University Press.

29. Phoebe Sengers and Bill Gaver. 2006. Staying open to interpretation: engaging multiple meanings in design and evaluation. *Proceedings of the 6th conference on Designing ...* (2006), 99–108. DOI: http://dx.doi.org/10.1145/1142405.1142422

30. Smith Optical. 2017. Lowdown Focus mpowered by Muse. (2017). http://www.smithoptics.com/us/lowdownfocus

31. Tim Urban. 2017. Neuralink and the Brain's Magical Future. (2017). https://waitbutwhy.com/2017/04/neuralink.html

32. Robert S Weiss. 1995. *Learning from strangers: The art and method of qualitative interview studies*. Simon and Schuster.

33. Richmond Y Wong and Deirdre K Mulligan. 2016. When a Product Is Still Fictional: Anticipating and Speculating Futures through Concept Videos. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. ACM, 121–133.