Access and Analysis: the Ethics of Brain-Computer Interfaces

First Author
Krysta Chauncey
161 College Avenue
Medford, MA 02155 US
Krysta.chauncey@tufts.edu

Second Author
Evan Peck
161 College Avenue
Medford, MA 02155 US
Evan.peck@tufts.edu

CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.
ACM 978-1-4503-0268-5/11/05.

Abstract
In this paper, we argue that physiological computing shares a major ethical issue with other areas of computing: determining and enforcing access to the information driving physiological computing. However, the subset of physiological computing that examines brain data has one major and unique ethical issue: unlike other types of information-gathering system, how the neural data is analysed is much more meaningful than what information is gathered. This unique property of brain data must shape any discussion of the ethical use of such data.

Keywords
Physiological computing, ethics, privacy, brain

ACM Classification Keywords
H5.m. Information interfaces & presentation: Misc.

Introduction
Physiological computing can use two basic data streams: one from the peripheral nervous system (galvanic skin response, heart rate, pupil dilation, etc), and one from the brain (electroencephalography, near-infrared spectroscopy, magnetic resonance imaging, etc). There is a crucial difference in the ethical calculus between these two streams. There is currently no

indication that it is possible to extract operation-specific cognitive or affective information from data exuded by the peripheral nervous system. However, any aspect or level of mental experience can be indicated by at least one method of neuroimaging. This complicates the ethical picture: there is no other single species of data that can be used to reveal so many different things—some of which may be ethical in some circumstances, and some of which are certainly not.

Peripheral Nervous System

The peripheral nervous system is best at indicating the current degree of physiological arousal—how engaged someone with their current environment. This is a very general index, however: it cannot generally be specified whether the participant is positively engaged (enjoying a challenge, for example) or negatively engaged (frustrated with a failing interface); it also generally cannot distinguish between physical arousal (increased heart rate due to exertion) from mental arousal (increased heart rate due to an impending achievement). This means that the ethics of using such data is much simpler than those of using the more flexible and wide-ranging data from the brain. The primary ethical issue in using peripheral nervous system data is access. Who gets to see this data? This question can best be answered for a specific system in a specific context, and should be considered as part of the design phase for any such system.

Mapping Data Type to Purpose

Most data in the world have a simple mapping between data type and purpose: locational data can only indicate where someone is—although it may be extrapolated to include what they are doing, this must necessarily be a fairly weak inference. Similarly, physiological data from the peripheral nervous system can only be used to indicate arousal levels, regardless of how it is analysed. However, brain data differs because, if the data gathered is complete enough, there is no limit to the subtlety, type, or degree of mental state that can be gleaned from it. All changes in mental state co-occur with changes in brain state, and it is only a matter of time before these co-occurrences are well enough understood to yield neural correlates of complex mental states. This means that that the important ethical question is no longer, what data should be collected and who should have access to it? But, what analysis techniques on that data are permissible and what use should be made of these analytical conclusions?

An Example

Suppose a brain-computer interface exists that gathers electrophysiological data, and rewards cognitive focus in a game-like setting in order to train ADHD participants (in fact, such an interface does exist [1]). Clearly, once the participant has been informed about the purpose of the game and given their consent, it is ethical for this program to analyse the participant's data to determine their current degree of cognitive focus. But is it ethical then for this program to analyse this data as a block, and determine the degree of severity of this particular patient's ADHD? How should that information be shared, or not shared? It is also currently possible to analyse this same data for evidence of clinical depression or epilepsy, as well as for incident-specific deception, comprehension, or expectations; in the future, it may also be possible to analyse this same data to find out completely extraneous information: the subject's opinion of the therapy administrator, for example. Some of these analyses are clearly out of the scope of a neurofeedback-type therapy for ADHD, but nothing in the data itself precludes such analyses. It is currently axiomatic in the privacy world that once a consumer has given permission for data to be collected, they lose control of how that data will be used. This must not extend to neural data, because what the data can say depend almost wholly on how they are analysed, not what type of data is collected.

Differences in Methods

It is worth considering if the different methods of neuroimaging have different ethical implications. This is of course an impossible question to answer at the moment, since the full capabilities of these methods are in flux and will probably remain so for years. However, there are some basic properties of these methods that may never change, and if they do, will change far in the future. These properties can give us an idea of what is likely to become possible, and what is not. The most common methods of neuroimaging used in physiological computing or applied neuroscience are electroencephalography (EEG), functional magnetic resonance imaging (fMRI), and functional near-infrared spectroscopy (fNIRS).

Electroencephalography measures electrical potentials at the scalp, which gives an indication of collective activity by populations of neurons. Since this method uses the transmission of electricity, it indicates activity as fast as the activity itself happens; however, since human tissue transmits electricity relatively well, its spatial resolution is and probably will remain much lower than that of other methods. This means that EEG will probably never give the best answers to questions centering on where in the brain things happen—so abuse of an EEG system by spatially-crucial analyses is unlikely. For example, people with a great deal of navigation experience have larger hippocampi than people with less experience [2]; EEG data certainly can't give any indication of this currently, and are extremely unlikely to be able to in the future. However, EEG data can be analysed to indicated mood and level of concentration, as well as a large and expanding catalog of cognitive markers, some of which could be spuriously used to indicate general intelligence.

Functional magnetic resonance imaging measures distortions in magnetic fields created by the differing magnetic properties of oxygenated and deoxygenated hemoglobin, from which neuronal activity can be inferred. The spatial resolution of fMRI is very high (around 2-3 mm), and continues to improve. The temporal resolution has historically been lower—around five seconds [3]. Although this also continues to improve, since fMRI fundamentally measures blood flow, and blood flow changes relatively slowly, it is likely that at some point the temporal resolution will plateau short of the capabilities of EEG. This means that abuse involving time-based analysis—say, using the length of activation for a particular task to extrapolate a participant's intelligence—unlikely. fMRI will probably be able to indicate truthfulness or deception at some point, although current methods are unreliable.

Similarly, functional near-infrared spectroscopy measures the differential refraction of near-infrared light from oxygenated and deoxygenated hemoglobin. Since light refraction happens much faster than the manipulation of magnetic fields in fMRI, fNIRS is a real-time measure of a slow process. Most fNIRS systems have very few channels, but in theory, they could have spatial resolution equivalent to fMRI in some parts of the brain. fNIRS is unlikely to be able to image subcortical areas of the brain, most important for emotion, motivation and reward, so abuse involving unauthorized extraction of emotional state is unlikely to become possible. Additionally, the same information unavailable to fMRI—high-speed changes in neuronal activity—will probably remain unavailable to fNIRS as well.

Although the capabilities of these methods will continue to change, some of their most basic properties in each case preclude certain types of abuse and make others more likely. These parameters can usefully guide the discussion of possible abuse for a given system.

A Change in Perspective

Thus far, we have focused on the ethical concerns of system users. What are the ethical concerns of system developers?

Developers working with brain data, like developers using other types of data, are likely to be concerned on how to restrict access to data efficiently, transparently, and fairly. However, the flexibility and multi-purpose nature of brain data poses a new and two-fold concern for developers. First, how can they restrict the type of analysis done on neural data to the user-approved and application-appropriate? This concern could be addressed by a system of permissions that only allows approved analysis, or by a signal-processing approach that attempts to extract only information relevant to approved analysis types. Second, how can developers communicate to possible system users what will be done with their data, what could be done with their data, and how analysis will be restricted to consented techniques only? This is more of a social engineering problem, and has no simple solution. Procedures and materials must be developed to ensure that subjects understand the data that is being collected, its possibilities, and its security.

Conclusions

In many ways, the ethical dilemmas of physiological computing are the same as any new technology: we wonder what the possibilities for misuse are, and how can they be prevented effectively. However, unlike other new data types we have discovered, the possibilities for the use and misuse of physiological data and brain data specifically are diverse and far-reaching and must be considered at the design phase of any system using such data.

References

[1] Lubar, J.F., Swartwood, M.O., Swartwood, J.N., & O'Donnell, P.H. Evaluation of the effectiveness of EEG neurofeedback training for ADHD. *Applied Psychophysiology and Biofeedback,* 20 (1995), 83-99.

[2] Maguire, E.A., Gadian, D.G., Johnsrude, I.S., Good, C.G., Ashburner, J., Frackowiak, R.S.J., & Frith, C.D. Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences,* 97(2000), 4398-4403.

[3] Menon, R.S., & Kim, S-G. Spatial and temporal limits in cognitive neuroimaging with fMRI. *Trends in Cognitive Neuroscience,* 3 (1999), 207-216.