Entropic Brain-computer Interfaces
Using fNIRS and EEG to Measure Attentional States in a Bayesian Framework

Samuel W. Hincks[1], Sarah Bratt[2], Sujit Poudel[2], Vir V. Phoha[2], Robert J. K. Jacob[1], Daniel C. Dennett[1] and Leanne M. Hirshfield[2]

[1] *Tufts University, Medford, MA, 02155, U.S.A.*

[2] *Syracuse University, Syracuse, NY, 13210, U.S.A.*

Keywords: BCI, Brain-computer Interface, fNIRS, EEG, Workload, Implicit Interface, Attention, Task-positive Network, Default Mode Network, Entropy, Physiological Computing, Entropic Brain-computer Interface, Bidirectional Brain-computer Interface, ADHD, Meditation.

Abstract: *Implicit Brain-Computer Interfaces* (BCI) adapt system settings subtly based on real time measures of brain activation without the user's explicit awareness. For example, measures of the user's cognitive profile might drive a system that alters the timing of notifications in order to minimize user interruption. Here, we consider new avenues for implicit BCI based on recent discoveries in cognitive neuroscience and conduct a series of experiments using BCI's principal non-invasive brain sensors, fNIRS and EEG. We show how Bayesian and systems neuroscience formulations explain the difference in performance of machine learning algorithms trained on brain data in different conditions. These new formulations posit that the brain aims to minimize its long-term surprisal of sensory data and organizes its calculations on two anti-correlated networks. We consider how to use real-time input that portrays a user along these dimensions in designing *Bidirectional BCIs*, which are *Implicit BCIs* that aim to optimize the user's state by modulating computer output based on feedback from a brain monitor. We introduce *Entropic Brain-Computer Interfacing* as a type of *Bidirectional BCI* which uses physiological measurements of information theoretical dimensions of the user's state to evaluate the digital flow of information to the user's brain, tweaking this output in a feedback loop to the user's benefit.

1 INTRODUCTION

In contrast to *direct brain-computer interfaces* (BCIs), which attempt to build a brain-based substitute for mice, keyboards, and other direct input methods, implicit BCIs (Poel et al., 2012; Zander et al., 2014; Zander et al., 2010) strive to portray dimensions about the user that are otherwise invisible to a computer, and use brain data to update system parameters whose values are better left beyond direct user control or attention but might still be usefully adjusted in response to the user's cognitive profile. In this paper, we attempt to sketch a process for automatically inferring whether or not a system property is a useful target for adaptation based on how it alters brain activation along two dimensions. In order to do this, we first explain the concept of entropy as it relates to brain functioning. In brief, Bayesian cognitive science posits that information processing in the brain increases when the brain fails to account for system input; and this increased prediction error, randomness, and computation in the brain correlates with the introspectable richness of experience (Carhart-Harris et al., 2015). Therefore, a user interface that tracked the cognitive burden that its constituent elements placed on a user, and adapted these *information scores* based on brain entropy measures, could learn how to customize itself to optimally enrich user experience.

In this paper, we compare the performance of machine learning algorithms calibrated on data from fNIRS on EEG when detecting different levels of cognitive workload as well as transitions between task and rest. Based on the Bayesian formulation that computation and energy expenditure in the brain depends in part on the novelty of information, we hypothesize that the performance of machine learning algorithms will degrade in the second session of an experiment. In the three sections of this paper, we explain how this property can be exploited as a feature not a bug of BCI.

Three interdisciplinary research questions under-

Hincks, S., Bratt, S., Poudel, S., Phoha, V., K. Jacob, R., Dennett, D. and Hirshfield, L.
Entropic Brain-computer Interfaces.
DOI: 10.5220/0006383300230034
In *Proceedings of the 4th International Conference on Physiological Computing Systems (PhyCS 2017)*, pages 23-34
ISBN: 978-989-758-268-4

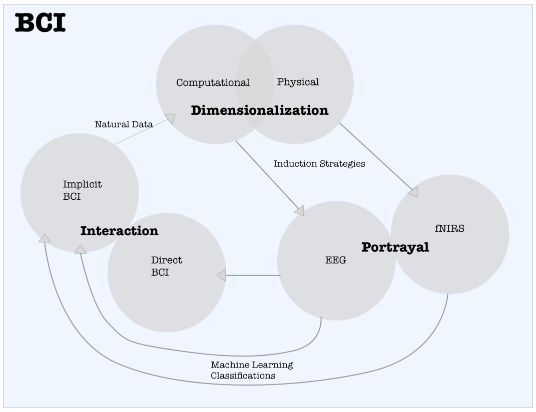

Figure 1: The three challenges of BCI and how they interact.

lie the creation of implicit brain-computer interfaces. First, an appropriate set of dimensions for the user's state needs to be identified. Much of previous BCI research uses machine learning algorithms to classify the user's state; the dimensionalization component of BCI amounts to deciding what states about the user the machine learning algorithms will classify. For example, many implicit BCIs (Afergan et al., 2014b; Afergan et al., 2014a; Afergan et al., 2015; Yuksel et al., 2016; Solovey et al., 2012) attempt to parse the user's cognitive workload. We refer to the challenge of identifying dimensions with both neurobiological meaning and relevance to interface design as the *dimensionalization* challenge of BCI. Second, for HCI, these dimensions must somehow be extracted on the basis of measurements collected from non-invasive and low-cost sensors, which we call the *portrayal* challenge of BCI. With known dimensions and methods of extracting them using practical sensors, the final challenge, what we call the *interaction* challenge, is to map classification interactively onto dynamic system parameters that drive adaptations to the user's benefit, and to measure such benefits. We illustrate these challenges and how they interact in Figure 1.

There are currently two brain sensors sufficiently promising for ordinary human-computer interaction scenarios: EEG (electroencephalography) and fNIRS (functional near infrared spectroscopy). They each have different advantages and disadvantages, complement each other quite well, and have been studied separately many times. EEG and fNIRS both index a user's brain activation. EEG relies on a chemically induced electrical charge that occurs at a neuron's action potential (the fundamental unit of computation in the brain). The EEG infers the aggregate of many such events by measuring voltage changes on the user's scalp. The detected voltage at the scalp oscillates according to specific patterns, which conveys information about the activity of neurons, especially near the cortex (Teplan, 2002). fNIRS relies on a neuron's need for oxygen to execute an action potential. When a given region is enlisted for computation, it consumes the oxygen it has available locally and requests more. By introducing near-infrared light from sensors by a user's forehead, which penetrates skin and bone but is sensitive to hemoglobin in the blood, a nearby light detector can infer changes to the amount of oxygen present, as oxygenated and deoxygenated blood absorb and scatter the light differently (Ferrari and Quaresima, 2012). Unlike EEG, whose most studied signals oscillate at frequencies ranges above 1 Hertz, fNIRS takes seconds to register changes in state. This prevents fNIRS from indexing the high frequency patterns powering the EEG signal but opens the device to measure longer term patterns in brain activation (Pfurtscheller et al., 2010).

Complementing each other's weaknesses and supplementing each other's strengths, fNIRS and EEG seem to invite integration into a singular input device that delivers a suite of predictions about the user's cognitive, emotional, and intentional state. EEG has poor spatial resolution (Berka et al., 2004), meaning it is difficult to resolve from where a given signal is originating; but it has good temporal resolution, meaning a given measurement is temporally very close to the phenomenon it endeavors to portray. Conversely, fNIRS has good spatial resolution but poor temporal resolution (Ferrari and Quaresima, 2012). Because of its good temporal resolution, EEG can capture a brief episode of mental activity and translate that activity into a command. For example, a user might imagine body movements (moving left or right hands) which would produce a machine learnable EEG signal, which then could be mapped to a deliberate command (Guger et al., 2003). fNIRS is very difficult to use for direct input because it would necessarily take seconds for the command to register. But EEG's sensitivity to short term electrical fluctuations leaves it vulnerable to noisy inputs. For example, an eye blink and other movement also produces an electrochemical effect that drowns the neurological signal underlying state classification (Berka et al., 2004; Hoffmann and Falkenstein, 2008). fNIRS does not suffer as severely from noise (Maior et al., 2015; Solovey et al., 2009).

In this paper, we make three contributions to implicit brain-computer interfacing, in dimensionalization, portrayal, and interaction, and show how these three pieces fit together using the concept of entropy to lead to the design and implementation of bidirectional brain-computer interfaces:

Dimensionalization: In the second section, we connect the concept of cognitive workload to system entropy in the Bayesian framework. Cognitive workload is the standard term for cerebral engagement, and is the user dimension that existing implicit BCIs (Afergan et al., 2014b; Solovey et al., 2012; Yuksel et al., 2016) have generally targeted for adaptation. Here, we ground cognitive workload with information-theoretical principles in Bayesian cognitive science and suggest that BCI supplement the one-dimensional workload attribute with a second user dimension, which resolves which spatial network is executing the cognitive workload.

Portrayal: In the third section, we run two sets of identical two-session experiments, where the user undergoes a series of computer-based task transitions. We use the EEG and fNIRS datasets to build machine learning algorithms that classify users along the dimensions proposed in section two, and evaluate their performance internally using standard cross-fold validation techniques. In the process, we relate the measurement of the two brain sensors (EEG and fNIRS) which are the obvious candidates to be the mouse and keyboard of next generation brain-computer interfaces but whose performance is seldom evaluated in concert. We suggest that EEG and fNIRS are better suited for detecting different dimensions of a user.

Interaction: Section four uses theory from section two and data from section three to propose a way for interactive systems to automatically discover and adjust what elements can be updated in order to pull the user's state in directions that improve the quality of attention, leading to designs for future user interfaces.

2 DIMENSIONALIZATION

Good dimensionalization is essential to implicit BCI because without a link between mental states and the neurobiological machinery that computes it, it should be impossible to build an algorithm that predicts the mental state from brain sensors. If dimensionalization is done properly, then implicit BCI has a small library of dimensions that are true enough to the basic operation of the brain to show an effect to non-invasive brain sensors but far enough away from its mathematical calculations to transmit meaning to a user interface designer. Physiological computing research attempts to measure dimensions of the user's mental states, e.g. cognitive workload (Fairclough et al., 2005; Venables and Fairclough, 2009), task engagement (Fairclough et al., 2009), stress (Fairclough, 1993), and various emotions (Picard, 1997). These dimensions are born from psychological research, and provide the basis for our previous adaptive systems (Afergan et al., 2014b; Afergan et al., 2015; Afergan et al., 2015). In these systems, we train a machine learning algorithm to separate a user's cognitive profile during different levels of the n-back, a psychological induction scheme that varies the strain on a user's attention and short term memory. Cognitive workload classifications then drive system adaptations, such as controlling interruptions (Afergan et al., 2015) or changing difficulty levels (Afergan et al., 2014a). In this section, we reinterpret these user dimensions from cognitive and systems neuroscience literature. The two most practical dimensions we find in this investigation provide an information theoretical formulation of cognitive workload, stress, and task engagement. In section 3, we show how these formulations are useful for explaining the behavior of machine learning algorithms; and in section 4, we consider new avenues for adaptive design based on these dimensions.

There is compelling evidence in recent neuroimaging literature to suggest that the brain is consistent in how it organizes the goals of its information processing (Friston, 2010). According to the Bayesian Brain Hypothesis (BBH), the basic goal of the brain is to actively and parsimoniously predict and suppress external sensory signals using the knowledge of internal models, and to update these models so that prediction error is minimized in the future (Friston, 2010). The brain might be summarized as a hierarchical prediction and error correction machine (Clark, 2013), in which information processing proceeds bidirectionally, so that statistically informed prediction flows from the top-down and prediction error modifies internal statistics from the bottom-up. Thus, computation at large (cognitive workload) ought to increase when the information content of the task or environment dictates interest and modification in more top-down predictive machinery. When circumstance is novel relative to the brain, more prediction error propagates through the hierarchical Bayesian filter, causing data flows to branch and spread up the information processing hierarchy (Carhart-Harris et al., 2015); and it is to this broadly defined event which we hypothesize non-invasive brain sensors positioned at the outer edges of the brain are approximately sensitive.

Entropy: In general, entropy indexes the average amount of surprisal in an information processing system, and in the brain, entropy is high when internal models fail to account for system input. High entropy states include infant consciousness (characterized by the lack of concepts needed to suppress external stimuli), early psychosis (characterized as a bug to top-down reality testing), near death experiences (a situation that is unusual and interesting), psychedelics (which bind to receptors in hierarchically central top-down information throttlers) and creativity (where existing models are penetrated and recalibrated by original material) (Carhart-Harris et al., 2015). The association between the reported richness of these experiences and their consistent grounding in terms of information processing that we hypothesize EEG and fNIRS can detect opens new avenues for design; it implies that appropriately designed implicit BCIs can coerce rich and creative experience by intelligently modulating information content in response to brain activity.

Predominant Network: When the informational exchange between human and computer is too low or high, the mind tends to retreat into an internal mode of cognition (Csikszentmihalyi, 1996; Raichle et al., 2001). In this mode, the brain grants more resources to endogenous systems, which are not yoked to the external environment (Raichle, 2010). Instead of computing on data that originates from the senses, these *resting networks* (the most prominent of which is the default mode network) (Raichle et al., 2001) operate on endogenous data or memories – planning, reflecting, and fantasizing. Because these networks generally operate even when the organism is otherwise liberated to take a break, the net oxygen consumption of the brain decreases by less than 5% at rest (Raichle, 2010), implying that a BCI that attempts to maximize user experience must supplement its cognitive workload or entropy index with a secondary dimension that informs the user interface the space in which computation is occurring in the brain. Conveniently for BCI, these endogenous *resting networks* are anti-correlated with the exogenous *task-positive networks*, making them especially suitable candidates for detection by a spatially well-resolved brain monitor, since nodes from the opposite network can be used to constrain false positives and most regions have a known and consistent bias towards involvement in either one of these two networks (Glasser et al., 2015). Since time spent mind wandering predicts reported present and future happiness (Killingsworth and Gilbert, 2010), implicit BCI that learned how to apply system adaptations that minimized endogenous retreat could impact user happiness beyond the immediate experience.

Together, *entropy* and *attentional orientation* explain the user dimensions already posited in psychological and physiological computing literature. For example, task engagement and stress are both high entropy, but engagement may involve the more consistent application of exogenous computation, whereas stress may imply periodic retreat into introspective networks. Designing for cognitive neuroscience dimensions and not psychological dimensions is worthwhile for implicit BCI, since these systems need to base user classification on physiological measurements from brain sensors. We predict that the greater the distance between tasks along these two dimensions, the easier it is for a machine learning algorithm to separate instances of those tasks.

The problem for BCI is that each user's brain is unique, and therefore poses unique external requirements for coercing it into different states. Furthermore, brains change over time based on new input. For these reasons, we think it is critical to dimensionalize the user with the intuitions of Bayesian mathematics (Perfors et al., 2011). If the energy expenditure of the brain is proportional to the degree to which it must update belief and action in order to integrate evidence from the environment and if non-invasive brain sensors are principally sensitive to physiological correlates of that energy expenditure, then the novelty of a task in relation to a control (e.g. resting state) is likely low hanging fruit for a machine learning algorithm operating on either fNIRS or EEG data to classify. To test this, we have repeated our experiment twice, and evaluated the change in performance of machine learning algorithms from a first to second session.

Bayesian cognitive science argues that the brain seeks to minimize its entropy but that it generates more computation and experience (Carhart-Harris et al., 2015) when it is in a state of high entropy. Many humans struggle to reconcile their brain's natural urge to rewire itself towards lower entropy with the desire for rich experience. A well-designed user interface can assist in this endeavor. In section 4, we describe *entropic BCIs*, which attempts to judiciously modulate user-mediated information in order to pull the brain into exogenous high entropy states. Figure 2 simplifies the user as always belonging to one of four states, the desirability of which depends on the current scenario. A rigid and mild endogenous brain indicates boredom or tiredness; a random and intense endogenous brain indicates creativity or anxiety; a rigid and mild exogenous brain indicates relaxation; and a random and intense exogenous brain indicates flow or stress. Each state has specific user affordances, tasks which suits it, controlled and incidental means to calibrate machine learning algorithms that detect it, as well as guidelines for how to transform the state into another one.

In our experiment, we hypothesize that fNIRS, which has good spatial resolution, is better than EEG at classifying when user's transition from resting to task states, but that EEG, which has good temporal

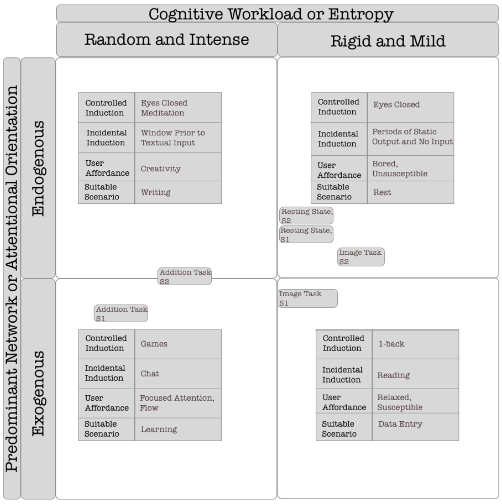

Figure 2: A simplification of users into four states along two dimensions, which we expect to be low hanging fruit for detection via fNIRS and EEG. The *Cognitive Workload* or *Entropy* dimension indexes the degree of computation and surprise in the user's brain, and the *Attentional Orientation* or *Predominant Network* dimension indexes the degree to which this computation is engaged in a goal-directed loop with its sensory environment.

resolution, is better than fNIRS at indexing a user's entropy or workload, since information processing occurs quickly in the brain.

3 PORTRAYAL

In typical BCI experiments at the portrayal level, the subject undergoes a series of task transitions under simultaneous interrogation of brain monitors (Girouard et al., 2009; Hirshfield et al., 2009; Hirshfield et al., 2011). These experiments and many others package results as the performance and behavior of machine learning algorithms, where the brain data of trials of the same condition receive the same label and are translated into a statistical feature-space. Machine learning algorithms, such as support vector machines, build models for relating the feature-space to the label, and, if the brain data differs between the two conditions, then machine learning algorithms should be able to predict the labels of trials that were not a part of its training model. The performance of a machine learning algorithm can be evaluated using a cross-fold validation scheme, where a model is built many times, each time excluding from its training set a new set of trials for its testing set. This machine learning method will be the basis for how we distribute results in this section. Because the novelty of the experiment is that we have repeated it four times using two portable

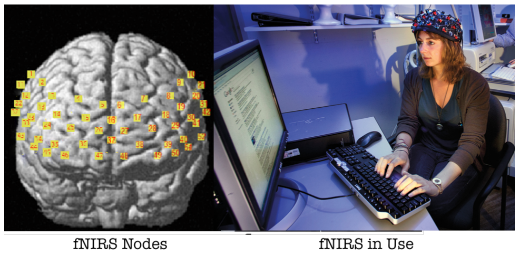

Figure 3: Hitachi fNIRS equipment.

brain sensing instruments (EEG and fNIRS), we can observe what dimensions of user information processing are better left to the jurisdiction of either sensor, in the hope of specifying how an fNIRS and EEG-integrated could jointly classify the user's state.

3.1 Equipment

The EEG used in this experiment was *Advanced Brain Monitoring's b-alert X10*, a 9 channel wireless EEG system with a linked mastoid reference, sampling at 256hz. The EEG headset was placed on users using standard 10-20 measurement set-up techniques (Homan et al., 1987). Figure 5 shows the nine regions of the brain measured by each channel.

The fNIRS device used in this experiment was the *Hitachi ETG-4000* fNIRS device with a sampling rate of 10Hz. The fNIRS probe (Figure 3) was a 3x11 probe with 17 light sources and 16 detectors, resulting in 52 locations measured on the head.

3.2 Method

Twenty-three subjects (8 female) between the ages 18 and 49 participated in the experiment. Upon arrival, subjects consented to the experiment and were fitted with the fNIRS or EEG sensors. The prepackaged *b-alert* and *Hitachi* software calibrated itself to the detected connection with the user's scalp. Then, the subject alternated between 8 instances of an arithmetic task and 8 instances of an image-matching task, performing each task for 35 seconds, with 15-second controlled rest periods in between. For the image-matching task, users indicated whether sequences of images matched each other, as in an n-back (Gevins and Smith, 2003) with *n* permanently set to 1, similar to the low cognitive workload condition used in previous implicit BCI work (Afergan et al., 2014a). For the arithmetic task, users added two two-digit integers to each other, entering the response into a text-box. We included workload in two separate modalities in order to, potentially, induce two separate states of workload In figure 4, we show the computer output for these tasks and in figure 2, we plot these tasks (and com-

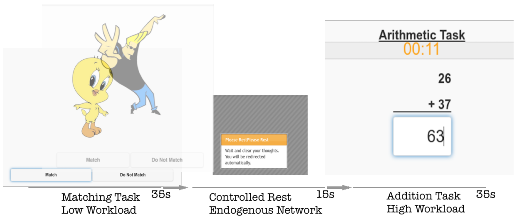

Figure 4: The three conditions of the experiment.

parisons between them) along the spectra of cognitive workload and attentional orientation.

We were thus interested in the transition between task and rest, and how that would change in the first and second session for both EEG and fNIRS. Our interest in inter-session comparison was born out of a consideration for how machine learning algorithms might decay over time if they did not account for updates to the user's cognition. We were especially interested in corroborating a previous small long term pilot (Hincks et al., 2016), where we tracked one of the author's fNIRS data over the course of several months as he made himself an expert at the cognitive workload tasks typically used in implicit BCI (the n-back). In later sessions, we noticed that fNIRS failed to register a strong effect unless the difficulty placed him at the edge of his ability, and we hypothesized that the task had over time generated too efficient top-down schemes for solving it, thereby leaking less prediction error, the information theoretical construct to which we hypothesize these cortically-sensing instruments are principally sensitive. If that is the case, then machine learning algorithms operating from either EEG or fNIRS data should degrade in a second session of the experiment. The experiment was thus repeated four times in two sessions for each device on two separate days. The second session for a given participant was at most 27 days later and at least 23 days later.

For each of the 23 subjects, we therefore had four datasets (two sessions for each device). Each dataset included 8 trials for the math task, 8 for the image task, and 16 for the resting task. We were interested in whether or not we could build a machine learning algorithm to separate the image and math task using both fNIRS and EEG data, and how this algorithm might change in the second session. We built a new machine learning algorithm, for each subject for each session, and for each neuroimaging device. We tried one approach for feature design using specifications we optimized elsewhere (Treacy Solovey et al., 2015), and for ease of communication, we let this algorithm be identical for both EEG and fNIRS, leaving the critical preprocessing components to the software distributed by *Hitachi* and the *b-alert* EEG.

For fNIRS, the input to our data mining scheme was a matrix of 52 channels that had been converted from light intensity into oxygenation measurements according to the Beer-Lambert Law, and bandpass filtered, leaving only the components of the signal that fluctuated between 0.01 and 0.5 hz. In our analysis, we omitted deoxygenation measurements since these values largely convey the same information as oxygenation. For EEG, the raw data was processed by *Advanced Brain Monitoring's* proprietary acquisition software, which includes artifact decontamination algorithms for eye blinks, muscle movements, and environmental/electrical interference. After decontaminating the raw data, the input was a matrix of 90 channels, consisting of the average power spectral density, averaged together into one second time periods, at each of the nine channel locations. Power spectral density was computed for the ten frequency bands of delta (1-3Hz), theta slow (3-5Hz), theta fast (5-7Hz), theta total, (3-7Hz) alpha slow (8-10Hz), alpha fast (10-13Hz), alpha total (8-13Hz), sigma (12-15Hz), beta (13-30Hz), and gamma (25-40Hz), at each of the channel locations.

For each instance, we computed the mean, linear slope, and standard deviation of the entire time-series of values for each channel. Thus, for inter-task (A1, A2) comparisons on the fNIRS data, the 350 *readings* x 52 *channel* windows became 2 *condition* x 8 *trials* x 156 *feature* instances and for EEG, the 35 *readings* x 90 *channels* window became 2 *condition* x 8 *trials* x 180 *feature* windows. For the *task* vs. *rest* comparisons (B1, B3), the same transformation occurred but the first 150 *readings* of the task were extracted, and compared to the 150 *readings* of resting data. We fed these feature sets into *Matlab's Statistics and Machine Learning Toolbox* implementation of the linear kernel support vector machine (SVM) and did not change default parameters (since the goal was to discern to what fundamental dimensions the machine learning algorithms were most sensitive and how that differed between devices, and not to maximize machine learning performance). We evaluated each machine learning separation using 10-fold cross validation (Refaeilzadeh et al., 2009), training the machine learning algorithm on all but an approximate tenth of the data, changing what tenth was omitted from the dataset and using that set for testing the trained classifier in ten separate tests. For all tests, evaluation instances were drawn from the same subject and session as the training instances that drove the machine learning algorithm. Next, we report on the averaged 10-fold cross validation classification accuracy for each test of interest.

Table 1: SVM Machine Learning Accuracies for Matching vs. Addition Comparison (A1,A2). m denotes the mean classification accuracy for all 23 subjects in 10-fold cross-validation and s refers to the standard deviation. For each row and column, classification accuracies have been compared in a *paired samples t-test*, and the *p-value* is reported.

S	EEG	fNIRS	p
1	m=73%, s=21%	m=72%, s=18%	0.7358
2	m = 80%, s = 29%	m = 71%, s =17%	0.0525
p	0.2430	0.8900	

3.3 Results

3.3.1 Task Comparison: A1,A2

For each session, we made two comparisons for each device, first distinguishing between the two task conditions and then separating the two tasks. Table 1 shows classification accuracies in cross-fold validation for the inter-task separation. In the previous section, Figure 2 shows how we expect these two tasks to differ from each other with respect to the workload and attentional orientation dimensions for the different sessions. The addition task presumably poses a greater burden to the user's cognitive workload than the matching task. We have compared the mean classification accuracies for each of the 23 subjects between both device and session, and Table 1 reports the probability that the null hypothesis is true in a *paired sample t-test*. There were no significant effects in these comparisons, but the EEG-based machine learning algorithms trended towards better performance. The results highlight that both devices can effectively parse the user along the cognitive workload dimension.

3.3.2 Rest Versus Task Comparison: B1,B2

Table 2 shows an identical analysis but for the comparison between the two tasks and rest. Since rest periods were shorter than task periods, we truncated the task trials so that they only included the first 15 seconds of data. Table 1 indicates that this comparison primarily manipulates whether or not the user has engaged an endogenous or exogenous network. For fNIRS, machine learning performance in the first session (m = 84%, s = 9%) is significantly better than machine learning performance in the second session (m = 75%, s = 13%) (p =0.0096, N = 23). Similarly, for EEG, machine learning performance decays significantly from the first session (m = 79%, s = 12%) to the second session (m = 68%, s = 17%) (p =0.0270, N = 23).

Table 2: SVM Machine Learning Accuracies for Task vs. Rest Comparison (B1,B2).

S	EEG	fNIRS	p
1	m = 79%, s = 11%	m = 84%, s = 9%	0.11
2	m = 68%, s = 17%	m = 75%, s =13%	0.11
p	0.0270*	0.0096**	

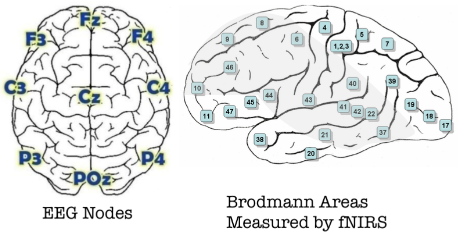

Figure 5: Channel locations for fNIRS and EEG. For fNIRS, only highlighted Brodmann Regions are measured.

3.4 Discussion

It is interesting that EEG (m = 76%) outperformed fNIRS (m = 71%) at separating the mathematical and image recognition task, which manipulates the user along the cognitive workload dimension but fNIRS (m = 80%) outperformed EEG (m = 74%), which manipulates whether or not the user is engaging a *task-positive* or *task-negative network*. Even though these differences are not significant, the results are consistent with the hypothesis that these two devices complement each other, covering the other's weakness. fNIRS is generally regarded as supporting better spatial resolution whereas EEG has better temporal resolution. Since every region in the brain is better described as belonging to either a task-positive or task-negative network and these two networks are anti-correlated, the fNIRS-based features (which are not as confused as EEG about the tissue they measure) might provide the information the SVM needed to discern the notion of anti-correlated networks, and robustly predict the user's state.

In general, as BCIs attempt to portray finer dimensions about the user, fNIRS may prove the better instrument for detecting spatial dimensions where different states of the dimension imply the enlistment of different neuroanatomic space. Beyond *predominant network*, an additional spatial dimension might include parsing the *modality* of input (spatial vs. motor vs. auditory) (Glasser et al., 2015). Conversely, EEG may prove better at extracting information processing phenomena. Beyond entropy or cognitive workload, this may include the *abstraction* (explicit vs. implicit) (Kahneman, 2011), *continuity* (static vs. dynamic) (Baddeley and Hitch, 1974), and *direction* (top-down vs. bottom-up) (Pinto et al., 2013) of cerebral data flows. Investigation into how to induce, portray, and design for these four other fundamental dimensions is an area of future research for implicit BCI.

It is also interesting that in both fNIRS and EEG experiments for the separation rest versus task, classification accuracy in session 1 reduces significantly in the second session for both fNIRS (p = 0.0096) and for EEG (p = 0.0270), but not for the separation between the two tasks, where classification accuracy is in fact better in the second session for EEG and approximately the same for FNIRS. But this is not surprising if the brain is dimensionalized according to the Bayesian framework. Both fNIRS and EEG measure brain activation principally at the outer-tips of the brain, its cortex. In a Bayesian framework, the outer-tips of the brain's hardware presumably carry out computations very high in the information processing hierarchy. In the second session, subjects had already been exposed to the tasks; thus, the second session cognitive makeup likely included internal representations that solved the input-output relations dictated by the task at a more primitive point in the information processing hierarchy, reducing the prediction error and its associated corrective events to penetrate the higher level regions of the brain under interrogation by the brain sensors.

Specifically, we attribute the relatively increased difficulty of the SVM to predict transitions between task and rest in the second session to the greater difference in system entropy between the resting and task conditions in the first session than in the second session. Since prediction error dictates the entire operation of a Bayesian brain, this may be another way of expressing that, in the second session, the user had absorbed efficient probability distributions for the task, enabling competing endogenous resting state inputs (which draw from the same finite pool of oxygen supply) to flourish and thus produce a profile that better matched the resting state. In simpler terms, the user's brains had figured how to efficiently solve the task in the second day of the experiment, but not how to efficiently rest. With this interpretation, machine learning accuracy did not change significantly for the inter-task comparison since the user had previously engaged both tasks, making so that task-induced entropy would decrease equally in both conditions.

The difficulty to control for system entropy between tasks and sessions is a feature not a bug of brain-computer interfacing so long as it is acknowledged by the designer. By fusing the user dimension of cognitive workload with system entropy, there is an opportunity to build a device that uses brain classification to manipulate the quality of user experience. In the next section, we show how to use the dimension of entropy in a bidirectional brain-computer interface.

4 INTERACTION

The interaction challenge of implicit BCI is to map machine learning classifications onto system variables whose values are better set with the knowledge of the user's cognitive state. There are a variety of such implicit BCI implementations reported in human-computer interaction venues for fNIRS. In previous work, we built an engine for adjusting task difficulty by training machine learning algorithms to separate fNIRS data pertaining to easy and hard versions of the n-back task (Afergan et al., 2014a). Once trained, the support vector machine provided a traffic monitoring simulation predictions about the user's workload, which then removed planes from user jurisdiction when they were deemed to be overworked, and added planes when they were underworked or bored. Users performed better at the task when adaptation was driven by changes in their brain as opposed to randomly. Other implicit BCIs have followed a similar principle, updating system properties based on brain activity, and comparing user performance to constant or random properties (Afergan et al., 2014b; Afergan et al., 2015; Solovey et al., 2012; Yuksel et al., 2016).

We refer to a BCI which attempts to alter the user's state as a *Bidirectional BCI* (Hincks et al., 2017), emphasizing the two-way channel in a system which outputs physical events which affect the state of the brain, measures this effect using non-invasive sensors, and uses this measurement to modify future interventions in a feedback loop. An *Entropic BCI* is a *Bidirectional BCI* which uses information theoretical models of cognition to interpret data from non-invasive sensors such as fNIRS and EEG in order to in real-time dimensionalize users along spectra which describe how the brain is currently processing information (e.g. system entropy, high vs. low cognitive workload, or attentional orientation). An *Entropic BCI* determines how to modulate the flow of information to the user based on these classifications in order to coerce more desirable attentional states. *Entropic BCIs* design for a brain that will retreat into endogenous processing when the informational exchange with the computer is too low or too high, and they acknowledge that the brain seeks to minimize its entropy but often relishes high entropy, learning, and the rich experiences that occur when the brain must adapt to new input. With current technology, *Bidirectional BCIs* can use four categories of intervention

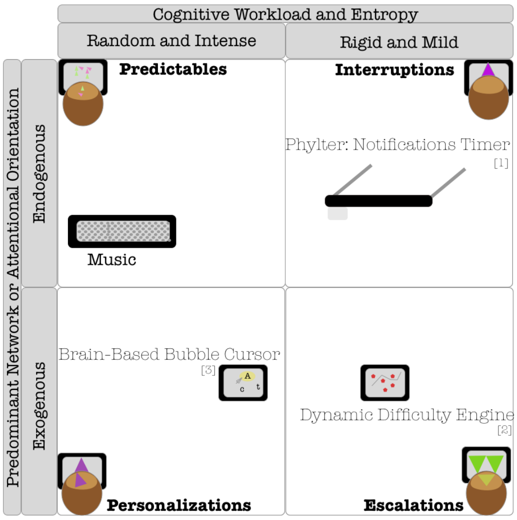

Figure 6: Four methods for implicitly modulating the informational exchange between the user and computer, and the state in which it is most suitable to apply them.

for manipulating the user's state, but new modalities, e.g. mixed reality and brain stimulation (Hincks et al., 2017) will expand this arsenal. In previous work, we have tested prototypes for three of these four categories (see figure 6), and we reinterpret these prototypes within an *Entropic BCI* framework.

Interruptions (e.g. text messages, email, or social media notifications) can be framed positively as methods for breaking malproductive rumination provided that they occur when the user has retreated into a mild task-negative endogenous state. In previous work, we built a system that triggered more wearable notifications when the user was deemed to be in a state of low cognitive workload, but in the associated experiment, performance measures did not improve dramatically (Afergan et al., 2015). An improved interruption engine might recognize the predominant network dimension and time notifications to break endogenous day-dreaming instead of low cognitive workload.

Escalations increase the difficulty or complexity of a task, and can make an underworked exogenous state more engaged. For example, one of our previous experiments increased the number of unmanned aerial vehicles under user jurisdiction when the user was predicted to be underworked, which improved overall performance in (Afergan et al., 2014a). Other BCIs that might be framed as escalations include a musical interface which adapted the challenge of musical scores (Yuksel et al., 2016) and a simulation which controlled the degree of automation in a semi-automated robot (Solovey et al., 2012). In general, if an HCI task is more strictly exogenous (meaning that it, like a computer game, aims to instantiate a swift input-output bond with the user), the more suitable it is to apply *escalations*.

Personalizations do not aim to modify the user's state, instead aiming to facilitate and design for the state that has unfolded. For example, we built a brain-based bubble cursor (Grossman and Balakrishnan, 2005), which modified the ease of selecting graphical user interface elements with a mouse dependent on the user's state (Afergan et al., 2014b). Like *escalations*, adaptations that cater to the user's current state can apply to the system's semantic layer (its internal values and parameters) or to its syntactic layer (its inputs and outputs) (Treacy Solovey et al., 2015).

Predictables are brain-adaptive computer output that convey no essential information to the user, existing only by virtue of the fact that the brain is a prediction machine that will faithfully dedicate processing and manipulate the user's state given unpredictable stimuli. We know of no BCI implementations that would qualify as *Predictables* but potentially, the world of unexplored ideas has many novel opportunities, making it uncharted area of BCI research. Music is noise organized by a musician to occupy a predictive sweet-spot for the listener, where the parameters of the sound are predictable enough to enlist top-down circuitry but unpredictable enough to necessitate adaptation and simulate vivacious experience. In an *Entropic BCI*, the sound would need to modify in response to brain activation. One simple alteration would be to modify the sound's origin in 3-dimensional space. This would function as a system knob for controlling the user's implicit computation since low level regions of the brain would need to predict and adapt for the moving sound. Potentially, this phenomenon could be replicated digitally by assigning a unique motion pattern to each constituent instrument or electronic track of a song. That sound could then be delivered by 3D audio headphones, adapting the amount of motion based on brain activity.

Beyond audio spatialization, music has many more opportunities for adapting sound and interfacing with the brain in an *Entropic Brain-Computer Interface* since any parameter governing how sound is produced in a synthesizer could be subject to optimization. In practice, such audio brain-computer interfaces might proliferate if headphones augmented with brain sensors (such as the Kickstarter project *Mindset*) became commercially available and if songs were distributed as programs which adapted based on input instead of as mp3s. The proliferation of adaptive music depends on musicians writing electronic songs using the Web Audio API (Rogers, 2012) and open source javascript software that extends it (Choi and Berger, 2013; Mann, 2015) since songs written using web synthesizers can be distributed online for playback on any browser, where they could change elements of a song in response to input as any other web application. Discovering other *predictables* is productive future work, especially in mixed and virtual reality scenarios that access more of the user's sensorium. With the *Entropic Brain-Computer Interfacing* framework, any system which specified a set of variables governing the information presented to the user could be tweaked by a common algorithm for automatically inferring when and how to modify settings to these variables in order to steer the user into desirable states. Such bidirectional BCI UI software would formalize the notion of *user relative information* and customize it over time depending on the user's cognitive state. The space of interaction between the user and computer could be encoded in a high level object referred to as the *information space*. The information space consists of *objects*, chosen to represent the independent set of elements that best describes the user's current exchange with the computer, as well as a set of *transitions* between different versions of that *object*. The novelty of this interface is that both *objects* and *transitions* have an *information score*, which describes the average *entropy burden* the *object* or *transition* poses on the user. These values are initially set randomly (or pre-calibrated according to expectations or the cloud), but adjust incrementally (up or down) based on concurrent brain-based entropy calculations. Over the long run on many users, we expect these *information scores* to be accurate even if the underlying classification algorithm is not because of regression to the mean. The set of *transitions* with low inferred entropy are defined as *implicit*. Depending somewhat on the goals of the interaction, a Bayesian implicit BCI could execute *low entropy (implicit) transitions* into an *information space* with a higher sum to the overall *information score* when the user had engaged a mild *resting state network*, implying that the cerebration demanded from the user's *task-positive network* was sufficiently low that the competing, internally ruminating branch of cognition was granted access to the energy supply and allowed to subtract the user from external experience. The effectiveness of that intervention would itself be evaluated by how the prediction of the user's state switched on the *task-positive* versus *task-negative* dimension, internally rewarding the transitions that succeeded and punishing the one's that failed.

5 CONCLUSIONS

In this paper, we have given theoretical and empirical arguments for why fNIRS and EEG can jointly estimate the user's state along the dimensions of *attentional orientation* and *cognitive workload*. We synonymize *cognitive workload* with *entropy* in order to make a bridge between implicit BCI and a growing Bayesian literature that is positing an equivalence between entropy and the introspectable temperature of experience (Carhart-Harris et al., 2015). We sketch four interface designs for pulling the user into high temperature externally oriented flow states based on brain activation, and sketch algorithm for how these interventions might improve automatically if the user interface formalizes, adjusts and adapts *user relative information*. We expect BCI to mature alongside virtual, augmented, and mixed reality technology. In due time, the computer output under the jurisdiction of mixed BCI technology will then extend beyond merely notifications and other system settings to the complete sensorium molding the brain's state. With concurrent measures of user entropy and attentional orientation, it may be possible to customize such a virtual reality so that it interactively creates arrangements of sensory information that maximize its user's experience and joy. A working implementation of this technology would warrant a careful ethical consideration of how to judiciously wield this new power over the brain.

ACKNOWLEDGEMENTS

We thank Beste Yuksel, Daniel Afergan, Evan Peck, Erin Solovey, Leanne Hirshfield, Tomoki Shibata, Anzu Hakone, Ronna ten Brink, Min Bu, Sonal Chatter, Yan Huo, Eun Youb Lee, Bushra Alkadhi, Feiyu Lu, Mary Skitka, Nick Sempere, Maya DeBellis, Tal August, James Carney, Nik Liolios, Beibei Du, and Calvin Liang who are students and alumni of the HCI group at Tufts. We thank Remco Chang, Erika Hussey, Tad Brunye, Sergio Fantini and Angelo Sassaroli from Tufts University. We thank Stuart Hirshfield from Hamilton College. We thank Google Inc. for support of this research. Vir V. Phoha was supported in part by National Science Foundation Award SaTC Number: 1527795.

REFERENCES

Afergan, D., Hincks, S. W., Shibata, T., and Jacob, R. J. (2015). Phylter: a system for modulating notifications in wearables using physiological sensing. In *International Conference on Augmented Cognition*, pages 167–177. Springer.

Afergan, D., Peck, E. M., Solovey, E. T., Jenkins, A., Hincks, S. W., Brown, E. T., Chang, R., and Jacob,

R. J. (2014a). Dynamic difficulty using brain metrics of workload. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 3797–3806. ACM.

Afergan, D., Shibata, T., Hincks, S. W., Peck, E. M., Yuksel, B. F., Chang, R., and Jacob, R. J. (2014b). Brain-based target expansion. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*, pages 583–593. ACM.

Baddeley, A. D. and Hitch, G. (1974). Working memory. *Psychology of learning and motivation*, 8:47–89.

Berka, C., Levendowski, D. J., Cvetinovic, M. M., Petrovic, M. M., Davis, G., Lumicao, M. N., Zivkovic, V. T., Popovic, M. V., and Olmstead, R. (2004). Real-time analysis of eeg indexes of alertness, cognition, and memory acquired with a wireless eeg headset. *International Journal of Human-Computer Interaction*, 17(2):151–170.

Carhart-Harris, R. L., Leech, R., Hellyer, P. J., Shanahan, M., Feilding, A., Tagliazucchi, E., Chialvo, D. R., and Nutt, D. (2015). The entropic brain: a theory of conscious states informed by neuroimaging research with psychedelic drugs. *Psychoanalytical neuroscience: Exploring psychoanalytic concepts with neuroscientific methods*, page 140.

Choi, H. and Berger, J. (2013). Waax: Web audio api extension. In *NIME*, page 499502.

Clark, A. (2013). Whatever next? predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(03):181–204.

Csikszentmihalyi, M. (1996). Flow and the psychology of discovery and invention. *New Yprk: Harper Collins*.

Fairclough, S. (1993). Psychophysiological measures of workload and stress. *Driving future vehicles*, pages 377–390.

Fairclough, S. H., Moores, L. J., Ewing, K. C., and Roberts, J. (2009). Measuring task engagement as an input to physiological computing. In *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, pages 1–9. IEEE.

Fairclough, S. H., Venables, L., and Tattersall, A. (2005). The influence of task demand and learning on the psychophysiological response. *International Journal of Psychophysiology*, 56(2):171–184.

Ferrari, M. and Quaresima, V. (2012). A brief review on the history of human functional near-infrared spectroscopy (fnirs) development and fields of application. *Neuroimage*, 63(2):921–935.

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138.

Gevins, A. and Smith, M. E. (2003). Neurophysiological measures of cognitive workload during human-computer interaction. *Theoretical Issues in Ergonomics Science*, 4(1-2):113–131.

Girouard, A., Solovey, E. T., Hirshfield, L. M., Chauncey, K., Sassaroli, A., Fantini, S., and Jacob, R. J. (2009). Distinguishing difficulty levels with non-invasive brain activity measurements. In *IFIP Conference on Human-Computer Interaction*, pages 440–452. Springer.

Glasser, M., Coalson, T., Robinson, E., Hacker, C., Harwell, J., Yacoub, E., Ugurbil, K., Anderson, J., Beckmann, C., Jenkinson, M., et al. (2015). A multi-modal parcellation of human cerebral cortex. *Nature*.

Grossman, T. and Balakrishnan, R. (2005). The bubble cursor: enhancing target acquisition by dynamic resizing of the cursor's activation area. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 281–290. ACM.

Guger, C., Edlinger, G., Harkam, W., Niedermayer, I., and Pfurtscheller, G. (2003). How many people are able to operate an eeg-based brain-computer interface (bci)? *IEEE transactions on neural systems and rehabilitation engineering*, 11(2):145–147.

Hincks, S. W., Afergan, D., and Jacob, R. J. (2016). Using fnirs for real-time cognitive workload assessment. In *International Conference on Augmented Cognition*, pages 198–208. Springer.

Hincks, S. W., Debellis, M., Lee, E. Y., ten Brink, R., and Jacob, R. J. (2017). Towards bidirectional brain-computer interfaces that use fnirs and tdcs. In *PhyCS*.

Hirshfield, L. M., Gulotta, R., Hirshfield, S., Hincks, S., Russell, M., Ward, R., Williams, T., and Jacob, R. (2011). This is your brain on interfaces: enhancing usability testing with functional near-infrared spectroscopy. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 373–382. ACM.

Hirshfield, L. M., Solovey, E. T., Girouard, A., Kebinger, J., Jacob, R. J., Sassaroli, A., and Fantini, S. (2009). Brain measurement for usability testing and adaptive interfaces: an example of uncovering syntactic workload with functional near infrared spectroscopy. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2185–2194. ACM.

Hoffmann, S. and Falkenstein, M. (2008). The correction of eye blink artefacts in the eeg: a comparison of two prominent methods. *PLoS One*, 3(8):e3004.

Homan, R. W., Herman, J., and Purdy, P. (1987). Cerebral location of international 10–20 system electrode placement. *Electroencephalography and clinical neurophysiology*, 66(4):376–382.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Killingsworth, M. A. and Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science*, 330(6006):932–932.

Maior, H. A., Pike, M., Sharples, S., and Wilson, M. L. (2015). Examining the reliability of using fnirs in realistic hci settings for spatial and verbal tasks. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 3039–3042. ACM.

Mann, Y. (2015). Interactive music with tone.js. In *Proceedings of the 1st annual Web Audio Conference*. Citeseer.

Perfors, A., Tenenbaum, J. B., Griffiths, T. L., and Xu, F. (2011). A tutorial introduction to bayesian models of cognitive development. *Cognition*, 120(3):302–321.

Pfurtscheller, G., Allison, B. Z., Bauernfeind, G., Brunner, C., Solis Escalante, T., Scherer, R., Zander, T. O., Mueller-Putz, G., Neuper, C., and Birbaumer, N. (2010). The hybrid bci. *Frontiers in neuroscience*, 4:3.

Picard, R. W. (1997). *Affective computing*, volume 252. MIT press Cambridge.

Pinto, Y., van der Leij, A. R., Sligte, I. G., Lamme, V. A., and Scholte, H. S. (2013). Bottom-up and top-down attention are independent. *Journal of Vision*, 13(3):16–16.

Poel, M., Nijboer, F., van den Broek, E. L., Fairclough, S., and Nijholt, A. (2012). Brain computer interfaces as intelligent sensors for enhancing human-computer interaction. In *Proceedings of the 14th ACM international conference on Multimodal interaction*, pages 379–382. ACM.

Raichle, M. E. (2010). The brain's dark energy. *Scientific American*, 302(3):44–49.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences*, 98(2):676–682.

Refaeilzadeh, P., Tang, L., and Liu, H. (2009). Cross-validation. In *Encyclopedia of database systems*, pages 532–538. Springer.

Rogers, C. (2012). Web audio api. *s Draft [online specification], Version*, 1.

Solovey, E., Schermerhorn, P., Scheutz, M., Sassaroli, A., Fantini, S., and Jacob, R. (2012). Brainput: enhancing interactive systems with streaming fnirs brain input. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 2193–2202. ACM.

Solovey, E. T., Girouard, A., Chauncey, K., Hirshfield, L. M., Sassaroli, A., Zheng, F., Fantini, S., and Jacob, R. J. (2009). Using fnirs brain sensing in realistic hci settings: experiments and guidelines. In *Proceedings of the 22nd annual ACM symposium on User interface software and technology*, pages 157–166. ACM.

Teplan, M. (2002). Fundamentals of eeg measurement. *Measurement science review*, 2(2):1–11.

Treacy Solovey, E., Afergan, D., Peck, E. M., Hincks, S. W., and Jacob, R. J. (2015). Designing implicit interfaces for physiological computing: guidelines and lessons learned using fnirs. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 21(6):35.

Venables, L. and Fairclough, S. H. (2009). The influence of performance feedback on goal-setting and mental effort regulation. *Motivation and Emotion*, 33(1):63–74.

Yuksel, B. F., Oleson, K. B., Harrison, L., Peck, E. M., Afergan, D., Chang, R., and Jacob, R. J. (2016). Learn piano with bach: An adaptive learning interface that adjusts task difficulty based on brain state. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 5372–5384. ACM.

Zander, T. O., Brönstrup, J., Lorenz, R., and Krol, L. R. (2014). Towards bci-based implicit control in human–computer interaction. In *Advances in Physiological Computing*, pages 67–90. Springer.

Zander, T. O., Kothe, C., Jatzev, S., and Gaertner, M. (2010). Enhancing human-computer interaction with input from active and passive brain-computer interfaces. In *Brain-computer interfaces*, pages 181–199. Springer.