Brain-Augmented Wearable Computers: Concepts, Designs, Tools, and Implementations

1st Author Name
Affiliation
Address
e-mail address
Optional phone number

2nd Author Name
Affiliation
Address
e-mail address
Optional phone number

3rd Author Name
Affiliation
Address
e-mail address
Optional phone number

ABSTRACT

Head-mounted wearable computers (e.g. Google Glass) support seamless integration with lightweight brain monitors such as functional Near Infrared Spectroscopy (fNIRS). A brain-augmented wearable with minimal peripheral hardware could model its user along several cognitive and emotional dimensions with neural correlates at lower exterior regions of the brain. It could use these classifications to better acquaint itself with the user, time interruption, modulate notification detail, and alter content. We distribute source code for Neuracle, an online toolkit for calibrating machine-learning-based physiological classification algorithms and linking them bidirectionally and in realtime with wearable computers. We validate Neuracle's calibration process in a small experiment, as well as use it to power adaptation in a wearable turn-by-turn navigation system and in software that associates images with time, space, and mental state.

Author Keywords

BCI; passive brain-computer interface; implicit interface; fNIRS; near-infrared spectroscopy; workload; default mode network; wearable; head-mounted display; physiological computing

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

By providing on-demand real estate in the user's visual periphery, head-mounted wearables (e.g. Google Glass) support applications that broadcast information with little cost of access on their bearers behalf. This new genre of human-computer interaction creates an intimate bond between user and computer that can also be disruptive, especially if the device mistimes its output, clutters and confuses, or in any way interrupts the user as she engages with other demands imposed by the real world. Well-designed applications err on the side of not interrupting the user. Even better would be applications that leave this decision to in-the-moment probabilistic calculation, and throttle notifications when the user's present situation or mental state appears unconducive to benefiting from its content. In this paper, we prototype two user interfaces that rely on passive input to alter notifications and content in consumer-grade wearable computers, as well as distribute source code for a data-processing toolkit designed to broadcast mental state classifications in realtime.

Physiological sensors offer valuable data towards the calculation to interrupt the user or not, and, positioned on the forehead, head-mounted computers can probe the richest source of the user's state, their brain, with little additional hardware as far as the user is concerned. But brain monitors tend to be fickle instruments – suitable for laboratory measurement, but underprepared for investigating a user in motion in the real world. Ideally, a brain-probing wearable computer should be visually indistinguishable from its unaugmented counterpart: light, cheap, aesthetically tolerable, and not dependent on conductive gel. These requirements rule out bulky brain monitors such as functional magnetic resonance imaging (fMRI). Electroencephalography (EEG) may work, although the constant noise of the motor cortex may overwhelm any informative signals [41].

Trimming a brain monitor to the point where its weight, aesthetic appeal and energy consumption renders it operable in an unpredictable non-laboratory environment, integrated with otherwise elegant consumer-grade technology, is no trivial challenge. fNIRS provides a promising realistic brain-monitoring technology for this purpose. In contrast to EEG and fMRI, fNIRS setup just requires the device to be placed against a user's head; it affords users the freedom to move their heads fairly naturally without affecting the recorded brain signals [48] and preserves relevant signals even when the user is in motion [20]. fNIRS relies on fundamentally cheap technology (a set of light-sources and light-detectors) – parts that can be miniaturized into light, hidden components of the device. In practice, efforts to build a portable, low-cost fNIRS are already underway [45, 14, 43], and the fundamental technology is very basic. A light source shines near-infrared light that penetrates skin and bone before being absorbed and scattered by hemoglobin in the blood. Critically, light interacts differently with oxygenated and deoxygenated hemoglobin, and the intensity of light returning to a nearby optical detector reflects this distribution, allowing local neuronal activation to be inferred according to the same neurovascular coupling principle [49] behind fMRIs.

Because of the biologically necessary lag between neuronal activation and associated hemodynamic trace, fNIRS data functions best as passive input [53] in an implicit user interface. Implicit Interfaces adapt the content, structure, or presentation of hidden or visual elements without the user's explicit intention. From the perspective of the user, implicit input is free: an additional bandwidth of information from the user to the computer. The input tends to be probabilistic and error prone. Therefore, a well-designed implicit interface never makes drastic changes in the face of uncertainty; instead, it subtly alters the content or appearance of internal or external data in a manner that is completely innocuous when the user's intention or mental state is miscalculated [51].

Implicit interfaces rely on accurate real-time models of the user's state. Unfortunately, there does not exist an effective open-source for real-time fNIRS data processing. OpenVIBE is a useful tool for processing EEG data but, calibrated for high-frequency signals, does not work as effectively for fNIRS [42]. Brainput provides tools for real-time fNIRS processing, but the Matlab code is not open-source, and does not offer a graphical user interface [47].

Research in Brain-Computer Interfaces generally underestimates both its field's challenge and possibility. If a system can portray the state and intention of its user, then half the challenge of human-computer interaction is solved. The standard approach taken in the last decade - inviting a random group of participants into a laboratory and evaluating their brain activity in aggregate - has not brought the field of BCI significantly closer to its exalted goal. Each user, by virtue of having unique prior experience, has a somewhat different functional setup to their brain, primarily with respect to the rigging of their attentive faculties. For some participants, a state induction task (such as n-back) effectively grabs their conscious mental workspace, whereas others run parallel mental processes while completing the task even though their scores and verbal report of the task could be identical [26]. Note that some of the most successful applications of brain-computer interfaces (brain-control of wheelchairs [31] or speech [32] for disabled users) does not suffer these issues, since summoning a particular intention is easier for a user than summoning a state.

There may therefore be a case for steering BCI away from studies on a large and randomly selected number of subjects. These studies do not fully establish ground truths for the states they endeavor to induce or measure. Instead, we suggest researchers adopt a fail-fast rapid-prototyping mindset to its inquiry, and run more controlled studies on a small number of participants who have cultivated skills of state control and introspection.

In this paper, we describe, test, and distribute Neuracle, open source software, which provides the user interface, visualization, and statistical tools needed to empower fNIRS-based state exploration by organizing state induction tasks, data streams, and manipulation techniques together in the same interface. We discuss Neuracle in the context of implicit interfaces for fNIRS-augmented wearable computers, and make six contributions:

- In section 1, we consider the space of cognitive/emotional states relevant to the design of wearable implicit interface. We ground these mental states to their associated neural correlates, and, based on a combination of fMRI and fNIRS literature, speculate which of these states a head-mounted computer with minimal attached brain probing could plausibly detect.
- In section 2, we expand upon the utility of knowing the user's mental state for the design of various categories of wearable applications. We portray the available design space, and provide concrete use cases and best practices for adaptation. We highlight the utility of applications that know when the user is at mental rest, engaging their brain's default mode network.
- In section 3a, we distribute source code on Github at (LINK REMOVED) for our new physiological sensor analysis toolkit, called Neuracle, specialized for rapid testing of new mental states, bidirectional realtime communication with wearables, and human-centered machine learning evaluation and training. We evaluate its user interface in an experiment.
- In section 3b, we investigate the effect of adaptive filtering, a promising technique for processing fNIRS data in a wearable computer context.
- In section 4a, we implement a brain-adaptive turn-by-turn navigation system for Google Glass, and evaluate it in a proof-of-concept experiment.
- In section 4b, we implement a cognitive heatmap (an application which associates images and pictures) for Google Glass, and evaluate in a proof-of-concept experiment.

1. COGNITIVE STATES WITH NEURAL CORRELATES NEAR A HEAD-MOUNTED DISPLAY

The head-mounted BCI we envision houses a sleeve of near-infrared-light source-detector pairs in a perimeter around its existing hardware. Its capability to extract user state information would depend on how far the fNIRS components extrude from the device. In this section, we highlight cognitive states with known neural correlates near the zone of easy detection with fNIRS. We therefore dismiss states that do not implicate computation in exterior regions of the prefrontal cortex (PFC), temporal lobe, or occipital lobe, and flag states that inhabit these regions but remain somewhat above above a baseline perimeter, formed by a wearable computer resting like glasses on the user's brow. We summarize these findings in figure 1.

1a. Cognitive Workload:

An engaged working memory (WM) spends explicit cognitive resources to sustain a conscious mental buffer that stores and manipulates data [4]. Continuous usage of working memory implicates the dorsolateral prefrontal cortex (DLPFC; Brodmann (B.) regions 9 and 46) [34]. With its relatively exterior location, the DLPFC can be probed using near-infrared spectroscopy [29, 3, 22, 33], and the n-back task, which requires participants to recall an item presented n iterations ago,

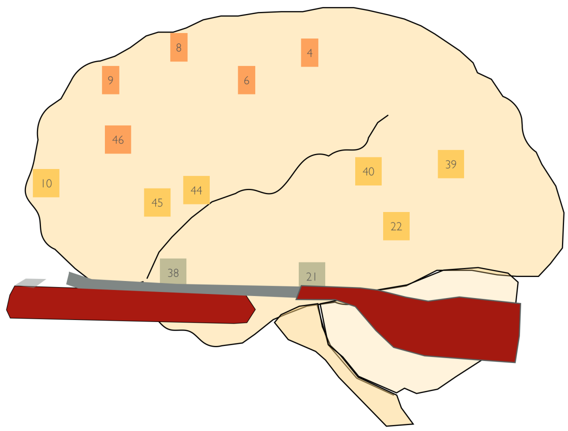

Figure 1. Summary of targets for brain probing with fNIRS-integrated Google Glass. The image reflects Glass's true location relative to the brain. Numbers refer to Brodmann areas (Brodmann, 1909). Red nodes may be beyond a minimally intrusive wearable; yellow nodes are within grasp if fNIRS components protrude beyond existing hardware; grey nodes are within direct reach. Brodmann areas 9 and 46 comprise the DLPFC (working memory). Brodmann areas 21, 10, 40, and 39 partake in the default mode network.

provides a reliable method for eliciting cognitive workload [2, 1]. In a typical fNIRS trial investigating activation in the user's prefrontal cortex, the cerebrally demanding 3-back task typically accompanies a steady escalation of the oxygenation concentration of hemoglobin, starting ten seconds into the trial, before leveling off at 30 seconds [19]. The modality of working memory content can be verbal (words and sounds) or spatial (locations in an environment), or object-related (shape, color, texture). Although all housed by the PFC, these systems have somewhat different underlying neural implementation. Verbal working memory (induced by an oral n-back) preferentially activates left-hemisphere speech areas (B.44) Spatial working memory activates the right premotor cortex (B.4), and object storage inhabits the ventral PFC at (B.10) [46, 36]. Current research has not been able to isolate significant neural distinctions between the task of storage and broader executive processes (e.g planning, problem solving, and decision making) [35].

1b. Emotion and Preference
The primary nodes for emotional processing inhabit deeper regions of the brain (the limbic system and its amygdala), concealed from light-based probing [28]. But the PFC also plays a role in emotion, and fMRI [6, 50] as well as fNIRS [15] shows higher vPFC (B.10) activation when emotive pictures are passively viewed and reappraised [37] and when a brain makes preference judgments about mundane items (e.g. coke vs pepsi) [38], indicating this region's regulatory and generative participation in human emotion and preference.

1c. Mind Wandering
Until the late 90's, neuroscience research mostly neglected the brain at rest, instead debating the usage of passive states as controls. Out of this debate emerged the discovery that the brain at baseline in fact involved a consistent network of brain regions which together perform the mental meandering that emerges at the cracks of task-positive experience [9]. Most notably, the default mode network (DMN) appears to show most activation at rest compared to when the brain engages a task [52]. Although there may be adaptive value in the brain's default state of mind wandering, this state carries an emotional price. Based on data taken from a smartphone survey of 5000 individuals, stimulus-independent thought comprises 46% of human cerebration. Time spent indulging the DMN precedes unhappiness, regardless if the content is pleasant [25], and recent research suggests DMN activation can be actively combated by the continued practice of meditation [8]. The main nodes of the default mode network include the lateral temporal cortex (B.21), vPFC (B.10), the posterior cingulate cortex, the precuneus, and the inferior parietal cortical regions (B. 40/39) (see Figure 1) [9].

2. IMPLICIT BRAIN-COMPUTER INTERFACES FOR HEAD-MOUNTED DISPLAYS
Based on this review, we posit that an fNIRS-integrated wearable computer could, with minimally obtrusive hardware, portray information about the user's emotional state and task engagement. With a few protruding optodes, it could quantify the engagement of working memory. Figure 2 summarizes the design space for an fNIRS-integrated wearable computer, and we elaborate further on related work and thoughts on how to use fNIRS input to passively learn about the user and as a means to adapt the when, how, and what of presented information.

2a. The when, how, and what of wearable notification streams
Unlike traditional computers that 'own' the user's attention, wearable computers are companions to the real world: digital assistants, whose utility hinges on successfully relaying context-specific information [7]. At any given moment, an active application must judge whether present circumstance warrants introducing information to the user's visual field, and if so, its content and format. There have been many efforts to investigate how to manage interruption according to estimates of the user's cognitive workload from physiological sensors such ECG [12], pupil dilation [21], EMG [11], EEG [10], and fNIRS [39] in the context of ordinary computer usage but not in a wearable context.

When the system has arrived at the decision to interrupt, it can further customize relevant parameters such as the content, level of detail, the minimality of design, and medium of presentation in accordance with the user's mental state. For example, it could swap between oral, written, and visual formats for route recommendations, capitalizing on the fact that working memory appears to support distinct (and parallel) processing units for these modalities [5]. In keeping with general design principles for implicit interfaces [51], these adaptive wearable applications should never shock the user; nor indeed, given their tendency to err, let the user in on the secret that probabilistic mental state estimate has changed some aspect of the interaction. Mission-critical information should never be held back nor erratically switch

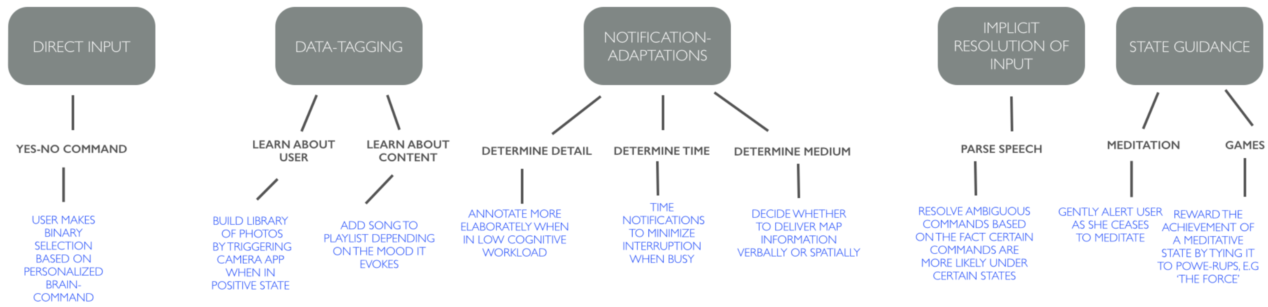

Figure 2. Five methods for using physiological data as input to head-mounted computers, with sample use cases.

formats. Instead, well-designed implicit interfaces would target non-essential or time/context insensitive information for adaptation. For example, a turn-by-turn navigation system would not drastically alter how or when it displayed an impending turn, but it could use its estimates to dictate the display of nearby gas stations, the looming presence of roadside attractions, or speed limits. In section 4a, we provide a prototype of a wearable turn-by-turn navigation system, which alters route recommendation based on measures of the user's cognitive workload.

2b. User Discovery

When navigating the web, a user leaves a rich cyber-trail that exposes their personality, mood, and preferences. Research into recommender systems has optimized methods for translating web interactions (e.g. clicking a link) into valuable models of the user, to the mutual benefit of the site and its customer. For example, Twittomender uses tweets and existing followers to recommend new followers [18] and Tagommender combines both user's search and rating history to infer a user's preference for tags and other movies [44]. fNIRS-based predictions of the user's preference have been used to augment movie recommendation algorithms [40]. Designed to avoid prolonged sessions of device interaction, wearable computers must incorporate other means than clicking links for user profiling. Wearables tend to remain active even when not used, as well as support a variety of sensors (e.g. camera, GPS, microphone), opening several channels for context sensitivity [13] . With concurrent physiological sensing, data from the environmental sensors could be timestamped with predictions about the user's mental state, and estimate the places, people, tasks and time periods associated with mental idleness, working memory engagement, positive emotion, focus, and mind wandering. These datapoints could inform targeted advertisement, recommendations, or could be related explicitly to the user in the form of a cognitive/emotional heatmap that associates space, time, and images with a current prediction of the mental state (see section 4b for a prototype implementation and demonstration of this idea).

The data analysis hub in charge of integrating the suite of physiological and context data should compute two simultaneous classes of models, the first predicting mental state given physiological data and context, and the second predicting profile from context and mental state. Machine learning with trustworthy confidence values should control whether physiological data drives the learning of the user's profile, their mental state, or both. Prior calibration from existing online templates as well as dedicated self-calibration are necessary steps for building a functioning, individualized classification system. The next section showcases Neuracle, which was built for these purposes.

3. NEURACLE

In this section, we present a new software system called Neuracle, validate its user interface in a small experiment, and explore its novel capability for functional connectivity analysis in a large experiment, investigating the brain's resting state default mode network. In the following section, we deploy its realtime output to alter information in glass or tagging user content.

In order to power adaptation in an implicit user interface, data from physiological sensors integrated on a wearable computer needs to be processed, filtered, and distilled into meaningful classifications about the user's state. Neuracle is an online interface for training and integrating mental state classification algorithms with wearable computers in realtime (see figure 3). It marries open-source Java-based statistics, filtering, and machine-learning libraries (such as Weka) with an easy-to-use interface for visualizing, tagging, manipulating, storing, and broadcasting data. The system makes the process of rapidly testing the detectability of a novel state an easy and satisfying endeavor. Source code can be found at GitHub (LINK REMOVED) under the MIT license. Documentation and tutorials are made available within Neuracle; the following sections describe novel interactions enabled by the software.

3a. Self-calibration Validation

We describe and test the following workflow for calibrating new physiologically-based machine learning algorithms, and refer to the interface's three components: (a) the data-view, which contains representations of active datasets (preloaded or streaming) and the tools for building machine learning algorithms from data, (b) the visualization-view, which shows live streams of the data and meaningful visualizations of it,

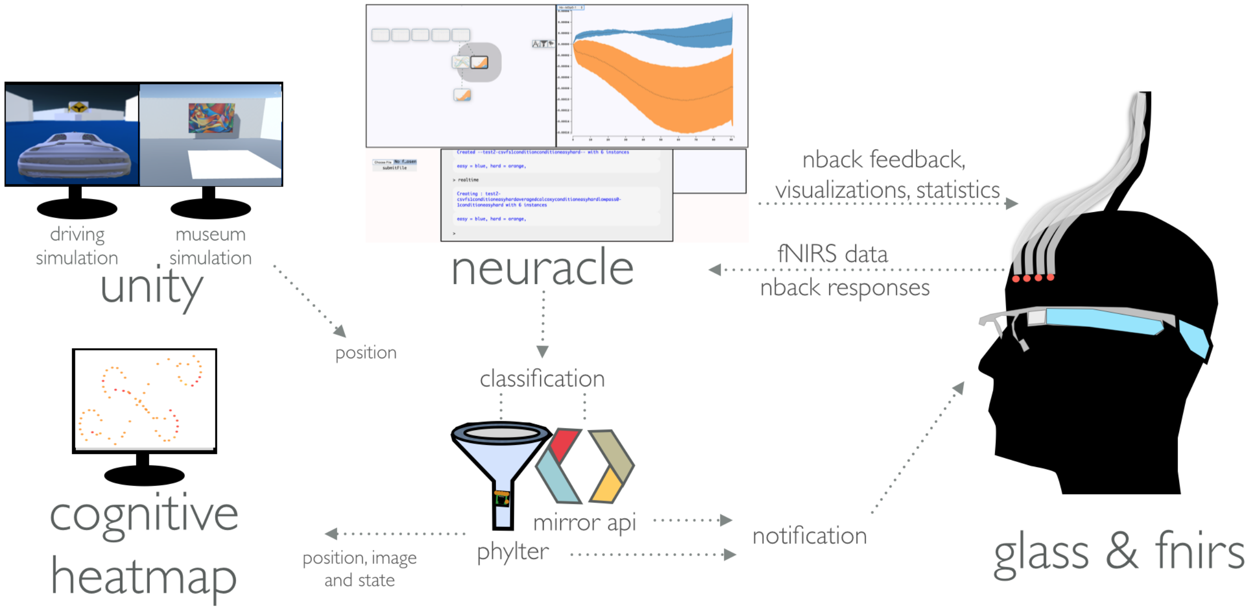

Figure 3. System architecture for Neuracle and implemented Glass prototypes. fNIRS data communicates to Neuracle server via an intermediate database. Here, the user calibrates a machine learning algorithm on their own data by solving an n-back in the console. Realtime classifications can then be redirected to a local port, (e.g. to Phylter) or to a web application communicating with Glass via the Mirror API.

and (c) the console, which communicates system output and recognizes over fifty commands for data manipulation, self-calibration, and more.

Participants

Five participants (3 female), ranging from ages 20 to 25, partook in self-calibration. The same participants partook in the *route recommendation* experiment (section 4a), and two of the five completed the *cognitive heatmap* (section 4b).

Method

In the experiment, tasks A (data synchronization), C (machine learning), and D (real-time classification) are completed by the 'trainer' (in this case, the experimenter), and task B (self-calibration) is completed by the one whose brain is being calibrated (in this case, the subject). We tested this workflow in an experiment, using fNIRS 16-channel ISS OxiplexTS: half sensing activation changes in the left DLPFC and half in the right DLPFC. The two probes are identical, each including source-detector pairings at 2cm, 2.5cm, 3cm, and 3.5 cm. Raw data values were converted to (de)oxygenation measurements using Boxy software, and relayed in realtime to Neuracle.

A: Data synchronization

Neuracle is designed to interact with changing values in a MySQL database. Once data is confirmed to be streaming into this database, the connection can be opened from Neuracle's console, which places a visual representation of the data in the data-view of the screen. Double clicking on this object shows a streaming view of the data in the visualization view.

B: Self calibration

To associate data with estimated mental states, the user must then initialize a set of trials where task is known by specifying a pattern to describe the condition, length, and quantity of trials. Neuracle supports built-in calibration for cognitive workload using the n-back. For the n-back, participants listen to an audio recording of numbers, with a 2.5s interstimulus interval, and enter these into the system. An affirming smile appears following correct responses, and a red frown follows incorrect response; classification accuracy is displayed to the user (and also saved along with the data, so that trials with many errors can be eliminated from analysis). In between trials, the user reports their experience of difficulty and focus. For inductions not yet built into Neuracle, the system merely displays trial transitions, playing helpful sound alerts if the user is performing the task on another screen.

C: Machine Learning Design and Evaluation

Neuracle provides a suite of tools for involving the human in the machine learning loop. When training data has been collected, the experimenter can examine condition-partitioned views of the data (see figure 4). These visualizations provide a preliminary outlook into the quality of the data, suggest outliers that ought to be excluded, as well as inform filtering and choice of features. Next, the experimenter can apply bandpass filters or other manipulation techniques (e.g. baseline subtraction or zscoring); in practice, such manipulation introduces a new dataset to the view, so that the choice can easily be reverted or compared to. (For states with consistent inter-participant physiological signatures, macros can be defined to implement a decided sequence of manipulations.)

The trainer can then define the feature-set; in Neuracle, a feature is specified by three values (a) descriptive statistic (b) channel (singletons or averaged), and (c) time (what subset of the data to examine). There are boilerplate descriptive statistics (e.g. mean , slope, and the first and second derivative) as well as more advanced features (e.g. the sax-distance [30] predictive granger-causality estimates [16] between two channels (useful for connectivity analyses), and the signal power at a specific frequency. These features have been selected for processing fNIRS in the time and frequency domain as well as for connectivity analysis, but apply generally to any physiological sensor. When a new feature-set is introduced, Neuracle automatically ranks each feature's information gain [24], which the experimenter can view by double clicking its visual representation.

Finally, the experimenter chooses among thirteen Weka machine-learning algorithms by evaluating them with cross-fold validation; this command shows the internal leave-one-out classification accuracy for the method, as well as confidence-thresholded estimates. A prudent experimenter does not wield this specificity to overfit nor does she flaunt inflated classification scores resulting from exhaustive trial-and-error. If classification accuracies appear abysmal, Neuracle supports tools for loading old datasets from other subjects, and either using this in lieu of or in conjunction to the subjects' data.

D: Realtime classification

When the experimenter has trained the machine-learning algorithm, she can instruct it to classify the streaming data, and request a classification on the last n datapoints, where n corresponds to the number of points in the trials it was trained on. This command can then be set to occur at timed intervals with the repeat command, and classifications streaming from the console can be redirected to a different address (e.g. a local port in the software in charge of wearable adaptation).

Results and Discussion

Participants reported appreciating the presentation of feedback as a green or red smiley, and also viewing time series that represented changes of their own brain activity. Figure 4 shows that the workload induction task elicited a comparable high workload signal to the literature [19].

Data manipulation was kept at a minimum to support realtime processing. For each trial, the slope, standard deviation, and mean were used as input to Weka's Logistical Model Tree [27]. No filtering was applied, in part since short term oscillations reflect heart rate and breathing patterns, which also correlate with workload [17]. In leave-one-instance out evaluation, classification accuracies ranged from 50 to 85% (m =65%, s=17%); and when classifications with an associated confidence measure above 75% were considered (47 of 70 samples), accuracies ranged from 50 to 88% (m =70%, s =17%).

After the experiment, all subject data was appended, and we evaluated the internal leave-one-out-accuracy (training data included data from the individual and others). Remarkably, the LMT correctly identified whether a trial was a two or zero back 87.5% of the time (90% when confident) even though individual subjects had classification accuracies as low as 50%. This underscores the importance of using many trials for machine learning, and suggests that future experiments ought to merge current calibration data with historic data. The most informative feature in this analysis as measured by information gain [24] was the mean deoxygenated hemoglobin value in the second half of the trial of a probe sensing the right DLPFC. These results validate Neuracle's capability to induce distinguishable cognitive states in its user; in the next section, we show that these datasets can be deployed into machine learning algorithms which probabilistically estimate the user's mental state in real time.

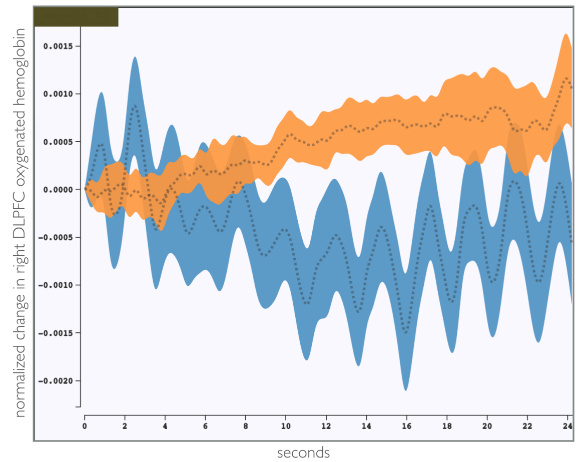

Figure 4. Average change in right-DLPFC oxygenation for all 14 trials of the five participants (35 0-backs in blue and 35 2-backs in orange). The dotted line represents the mean and thickness of the area chart represents one standard deviation. Left hemisphere measurements look nearly identical

3b. Adaptive Filtering

Motion artifacts constitute the major bottleneck for brain monitoring in an fNIRS-based wearable computer setting. Slight movements can cause a slight decoupling between the sensor and the skin. Adaptive filtering, introduced in [54], is an effective technique for filtering systemic trends in local fNIRS data. Bandpass filtering successfully removes breathing and heart rate, but it leaves in tact spontaneous low frequency oscillations that do not have neurological origin. These oscillations would be present in both a shallow and deep source-detector pairing, and can be eliminated from the deep source-detector pairing with knowledge about what frequencies were common between them.

Neuracle comes equipped with the software necessary for adaptive filtering. Specifically, it uses the Widrow-Hoff Least Mean Squared (LMS) algorithm to remove common inter-channel oscillations (as detailed in [54]). The following analysis shows one of this paper's author's brain activity, and therefore needs to be interpreted with caution. It nevertheless illustrates the potential effectiveness of the adaptive filtering algorithm. In this self experiment, the author alternated between four trials of sixty seconds in a passive mind wandering eyes-open state, separating 5 three-minute trials of eyes-closed meditation, in which the author attempted to align conscious experience with breath while recognizing and

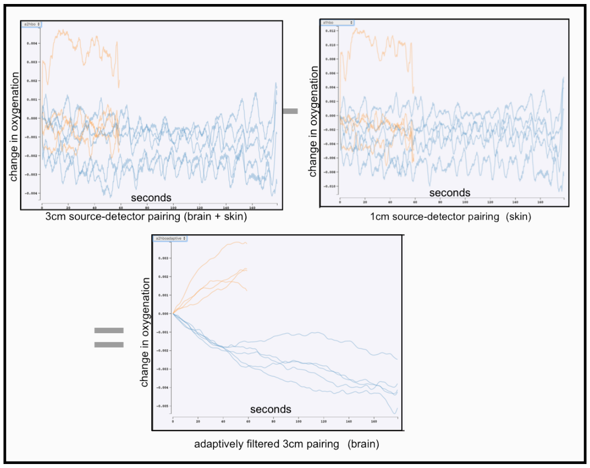

Figure 5. The effect of adaptive filtering on left dlPFC oxygenation data. Orange trials show 60 seconds of controlled eyes-open rest. Blue trials show 180 seconds of eyes closed meditation

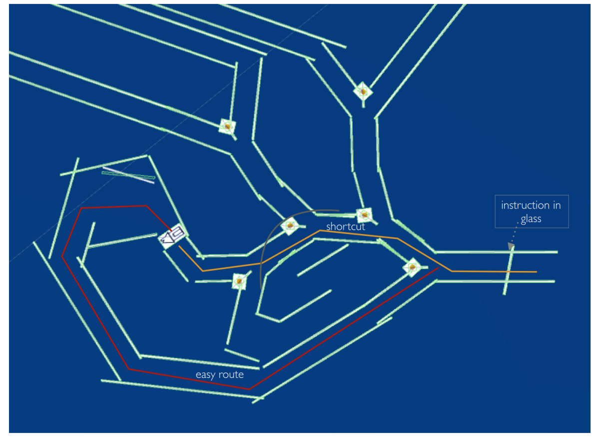

Figure 6. Unity map design shows two paths to reach a destination: one straightforward route and one which requires the participant to remember four simultaneous directions

mindfully discarding any incoming thoughts or feelings. The purpose of this exercise was to obtain two conditions which maximized brain-exclusive differences in state in order to test an adaptive filter algorithm under pristine conditions. These conditions differ by degree of deliberate engagement and volume of incoming sensory input - features of mind likely to implicate signal only in brain and not skin.

Figure 5 illustrates the presence of a signal invisible without adaptive filtering. The bottom graph shows a dataset in which the frequencies present in a 1 cm source-detector pairing have been subtracted from the frequencies present in a channel showing a source 3 centimeters apart from the same detector. Since the near channel only detects oxygenation oscillations from skin and the far channel detects oscillation from both skin and brain, an adaptive filter enables the investigation of an exclusively brain-based signal. In Figure 5, the adaptively filtered data shows that meditation involves a reduction in oxygenation content in the left dlPFC, before it starts to increase in the resting condition. This filter works both offline and real-time in Neuracle, thereby contributing the enabling software for more effective wearable neuroimaging as well as an exploratory self discovery approach to analyzing the meaning of the fNIRS signal. This discovery period can be leveraged both to rapidly prototype feasible states and to intelligently search the forehead for the most effective probe placement.

4. USER INTERFACE EXPERIMENTS

We ran two additional experiments in order to validate Neuracle's realtime capability and to demonstrate novel implicit interactions with head-mounted computers. In these experiments, the same five participants completed tasks in a unity game simulating the real world, as the machine-learning algorithms described above transmitted classifications that altered some hidden aspect of the interaction.

4a. Adaptive Route Recommendation

Google, Apple and other technology companies have dedicated considerable resources towards accurately mapping the planet's roads and collecting detailed information about their convenience and traffic patterns. Research in human spatial navigation ability lays the foundation for the user interface of these systems, informing parameters such as the the timing and frequency of turn notifications as well as the level of detail most amenable to a balance between comprehensible and information-rich content. But in many cases, optimal settings for these parameters is not fixed across time and user, but highly dynamic, a function of the user's mood or workload. There is usually more than one route to reach a particular destination, and often routes vary both in speed and complexity. Invariably, turn-by-turn navigation software tends to suggest the swiftest, shortest route, sometimes guiding users down difficult paths that lead them astray. If one path is slightly slower but significantly easier to process and navigate than an alternative route, then an up-to-date model of the user's cognition can aid the choice of what path to suggest.

Architecture

We simulate a wearable turn-by-turn navigation system by integrating Google Glass with an Imagent fNIRS imaging device. Imagent broadcasts raw data to Boxy software, which transmits calculations of neural oxygenation levels to Neuracle. Neuracle rebroadcasts predictions of the user's cognitive workload to Phylter [23], which intercepts user locations from Unity software, and decides whether or not to recommend the shortcut. We built four roads (e.g Figure 6) in Unity: each supports one quick but challenging route, one long but easy route, and a variety of other routes which lead the user astray. The quick route requires the user to read, rehearse, and execute four directions.

Method

The user began by solving either an oral 0-back (low cognitive workload) or a 1-back (high cognitive workload). After twenty seconds, the driving simulation began (as the user continued to solve the n-back), and within ten seconds, the suggested route appeared on Glass. In the static condition, the user always received the shortcut: four simultaneous instructions. In the adaptive condition, the user received the shortcut when they had a low predicted cognitive workload

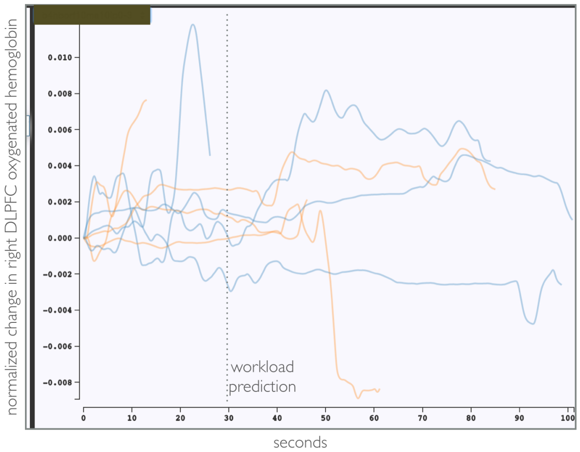

Figure 7. Concurrent fNIRS measurement in right DLPFC during driving simulation for two subjects. In blue trials, participant completed a 0-back (easy), and in orange trials, participants completed a 1-back (hard) while driving. The system delivered the direction after 30 seconds, and trials varied in length depending on how long it took to reach the destination. In the few samples shown, easy trials tended to provoke more variable data than hard trials, which matches the effect observed in offline calibration analysis (see Figure 4).

and the longer route when they had a high predicted cognitive workload.

Results and Discussion

In the adaptive condition, the system correctly displayed the appropriate route in Glass in each of the ten instances (2 roads for 5 participants), depending on its present calculation of the user's workload. However, users did not arrive at the destination faster nor suffer fewer collisions during the adaptive condition. They completed the simulation approximately as fast when completing a 0-back (m=99s, s =50s) as when they completed a 1-back (m=105, s=44). Figure 7 (which shows concurrent brain measures) shows the raw data on which Neuracle based its prediction of user workload. This validates the bidirectional communication capability of Neuracle; fNIRS data can be tagged from external software, which enables it to learn new user classifications post-calibration.

The high variability of results indicates that adaptive route recommendation requires careful settings to several subject-specific parameters, which may be difficult to extract in a laboratory setting. A proper cloud-based service like Google Maps would infer difficulty and skill parameters of route recommendations and users based on how frequently the suggested turn failed to coerce the correct future GPS coordinates. Concurrent measurements of the user's mental state could then identify the degree to which taking the wrong turn was sensitive to their mental state, and suggest easier routes for the state-sensitive users when in a high workload or otherwise compromised state.

4b. Cognitive Heatmap

A well-designed wearable computer supports its user's goals even when its services are not explicitly requested [13]. Equipped with a camera, the device could take repeated snapshots of the user's vantage point. As a result, the user would have a log of data that documented every activity engaged, every acquaintance encountered, and every location visited. If images were tagged with meaningful descriptions about the user's state, the user could search time points of interest and also catalog the sorts of events that stimulated particular mental states. For example, if the application operated while its user attended a lecture, she could afterwards revisit a state-indexed timeline, and discover the moments that stimulated the highest levels of cognitive workload, or, conversely, the moments when she wasn't paying attention. In this experiment, we prototype a cognitive heatmap that enables users to select relevant photos based on the associated cognitive state.

Architecture

The architecture resembles the previously described system, except Neuracle's cognitive workload predictions never alter the contents of Glass. Instead, Unity position coordinates, Glass photos, and mental state predictions are all associated in a single file, which D3 visualization software uses to display a set of points, colored on a continuous scale from yellow to red representing the confidence that their workload is high, which the user can click on to retrieve the associated image (see Figure 3).

Method

Participants worked as a museum curator, instructed to count the distribution of paintings in various rooms. Each room was a simulated environment, and the game automatically moved them from one point to another; they controlled the direction of the camera. Some rooms included only one type of painting, and thus the task amounted to counting the number of paintings. Other rooms included two easily distinguishable types of paintings (Pop Art and East Asian), and updating and rehearsing this distribution was meant to stimulate high cognitive workload. These rooms contained signs pointing to the content and time of future exhibits. In a later task, the user had to use the cognitive heatmap to find the image that contained this sign. If the system worked correctly, then these would be by the high cognitive workload (red) points.

Results and Discussion

Participants reported using the indexing of their mental state as a search mechanism. However, they did not find the photo quicker when the map was colored according to their cognitive workload. It was challenging to coerce the subject to fixate their gaze on the relevant image, and so some of the generated maps included only a blurred rendition of the target image. We nevertheless think the system shows promise, especially with well-calibrated models whose confidence values communicate true model uncertainty and in use cases where it need only pluck out a fraction of true positives in order to benefit the user, e.g. if the user only has time to review one lecture out of a longer series.

CONCLUSION

Head-mounted wearables surround the human brain, and are thus well-suited to detect changes in the user's cognitive and emotional state using low-cost portable neuroimaging. Functional neuroimaging literature has identified several cognitive and emotional dimensions with neural correlates in lower, surface brain regions, which invite seamless probing using an fNIRS-integrated wearable computer. Software detecting transitions between these states can improve interaction in a medium requiring inputs that supports its ability to continuously serve the user without demanding too much from him/her. We have identified five categories for how to utilize this input, and elaborated on how to leverage implicit input to control the when, where, and what of notification streams for immediate impact, as well as passively tag information gleaned from the context (e.g. time and space) with associated mental state.

Neuracle provides a cloud-based interface for bidirectional and realtime communication with wearable computers. The wearable broadcasts physiological data that Neuracle intercepts and processes, before returning a probabilistic estimate of the user's current mental state. The interface provides the tools necessary to initiate calibration schemes (coupling live data with known mental states) and explore an optimal series of filtering, data manipulation, feature extraction, and machine learning. In a small validation study, Neuracle was shown to (a) enable the experimenter to make live decisions about how to train a machine learning algorithm, (b) enable the participant to complete a self-calibration protocol without experimenter intervention, and (c) support realtime changes to two wearable applications in Google Glass. Offline evaluation suggests that these estimates had a good chance at corresponding to realtime classification. Interestingly, the results from the current study suggest that merging data across subjects (i.e., using both the current and other participant data) leads to the most accurate model in the absence of an extensive (many-trialed) individual calibration scheme. In addition, an SVM learned to discriminate between an addition and resting task with 69% accuracy, having aggregated data from fifty other subjects but not having ever peeked at the subject supplying the test-case. These results suggest the effectiveness of saving and sharing data to avoid or supplement individual calibration.

The realtime output of Neuracle successfully powered adaptation and passive tagging in two micro experiments. These small experiments yielded no definitive conclusions about the utility of the implemented interfaces, but nevertheless demonstrate novel adaptation methods in a new medium. They highlight the importance of using mental state classifications covertly and with reliable confidence measures in situations where the alternative is system ignorance. By distributing the source code for this architecture, we hope to stimulate further investigation into using mental state state classification for wearable user interfaces.

Although Neuracle was designed to be a web-based platform for intercepting, processing, and responding to fNIRS data from disparate wearable computers such as Google Glass, its key affordance may be in providing a live representation of data, bundled with the necessary filtering tools and machine learning as well as benchmark interfaces for state induction. In distributing this source code, we hope to inspire fellow brain-computer interfacing researchers to adopt a somewhat less conventional but more efficient approach to deciphering the elusive fNIRS signal. It is known that in aggregate fNIRS oxygenation values tend to increase during states of higher workload [19]; it is known that a machine learning algorithm can predict cognitive state better than chance on offline signals [51]. But what is not known is a consistent and predictable profile of the fNIRS signal that occurs reliably on any given subject, and part of this ignorance is due to the difficulty of inducing state externally. We therefore expect worthwhile progress to be made if the researchers developing the necessary filters, machine learning algorithms, calibration protocols, and state induction tasks begin to tinker and explore in conditions in which the ground truth of mental state can be controlled and introspected.

ACKNOWLEDGMENTS
(Omitted for blind reviewing)

REFERENCES
1. Daniel Afergan, Evan M Peck, Erin T Solovey, Andrew Jenkins, Samuel W Hincks, Eli T Brown, Remco Chang, and Robert JK Jacob. 2014a. Dynamic difficulty using brain metrics of workload. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 3797–3806.
2. Daniel Afergan, Tomoki Shibata, Samuel W Hincks, Evan M Peck, Beste F Yuksel, Remco Chang, and Robert JK Jacob. 2014b. Brain-based target expansion. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*. ACM, 583–593.
3. Hasan Ayaz, Patricia A Shewokis, Scott Bunce, Kurtulus Izzetoglu, Ben Willems, and Banu Onaral. 2012. Optical brain monitoring for operator training and mental workload assessment. *Neuroimage* 59, 1 (2012), 36–47.
4. Alan Baddeley. 1992. Working memory. *Science* 255, 5044 (1992), 556–559.
5. Alan D Baddeley and Graham J Hitch. 1974. Working memory. *The psychology of learning and motivation* 8 (1974), 47–89.
6. Felix Bermpohl, Alvaro Pascual-Leone, Amir Amedi, Lotfi B Merabet, Felipe Fregni, Nadine Gaab, David Alsop, Gottfried Schlaug, and Georg Northoff. 2006. Dissociable networks for the expectancy and perception of emotional stimuli in the human brain. *Neuroimage* 30, 2 (2006), 588–600.
7. Mark Billinghurst and Thad Starner. 1999. Wearable devices: new ways to manage information. *Computer* 32, 1 (1999), 57–64.
8. Judson A Brewer, Patrick D Worhunsky, Jeremy R Gray, Yi-Yuan Tang, Jochen Weber, and Hedy Kober. 2011. Meditation experience is associated with differences in default mode network activity and connectivity. *Proceedings of the National Academy of Sciences* 108, 50 (2011), 20254–20259.
9. Randy L Buckner, Jessica R Andrews-Hanna, and Daniel L Schacter. 2008. The brain's default network. *Annals of the New York Academy of Sciences* 1124, 1 (2008), 1–38.

10. Maher Chaouachi, Imène Jraidi, and Claude Frasson. 2011. Modeling mental workload using EEG features for intelligent systems. In *User modeling, adaption and personalization*. Springer, 50–61.

11. Daniel Chen, Jamie Hart, and Roel Vertegaal. 2007. Towards a physiological model of user interruptability. In *Human-Computer Interaction–INTERACT 2007*. Springer, 439–451.

12. Daniel Chen and Roel Vertegaal. 2004. Using mental load for managing interruptions in physiologically attentive user interfaces. In *CHI'04 extended abstracts on Human factors in computing systems*. ACM, 1513–1516.

13. Anind K Dey, Daniel Salber, Gregory D Abowd, and Masayasu Futakawa. 1999. The conference assistant: Combining context-awareness with wearable computing. In *Wearable Computers, 1999. Digest of Papers. The Third International Symposium on*. IEEE, 21–28.

14. Younes Rakhshani Fatmehsari and Zahra Moussavi. 2015. Designing and Implementing a Portable Near-Infrared Imaging System for Monitoring of Human's Brain Functional Activity. *Journal of Medical Devices* 9, 2 (2015), 020921.

15. Evelyn Glotzbach, Andreas Mühlberger, Kathrin Gschwendtner, Andreas J Fallgatter, Paul Pauli, and Martin J Herrmann. 2011. Prefrontal brain activation during emotional processing: a functional near infrared spectroscopy study (fNIRS). *The open neuroimaging journal* 5 (2011), 33.

16. Clive WJ Granger. 1988. Some recent development in a concept of causality. *Journal of econometrics* 39, 1 (1988), 199–211.

17. Thomas C Hankins and Glenn F Wilson. 1998. A comparison of heart rate, eye activity, EEG and subjective measures of pilot mental workload during flight. *Aviation, space, and environmental medicine* 69, 4 (1998), 360–367.

18. John Hannon, Mike Bennett, and Barry Smyth. 2010. Recommending twitter users to follow using content and collaborative filtering approaches. In *Proceedings of the fourth ACM conference on Recommender systems*. ACM, 199–206.

19. Christian Herff, Dominic Heger, Ole Fortmann, Johannes Hennrich, Felix Putze, and Tanja Schultz. 2013. Mental workload during n-back taskquantified in the prefrontal cortex using fNIRS. *Frontiers in human neuroscience* 7 (2013).

20. Takahiro Igarashi, Kaoru Sakatani, Norio Fujiwara, Yoshihiro Murata, Takeshi Suma, Tadashi Shibuya, Teruyasu Hirayama, and Yoichi Katayama. 2013. Monitoring of Hemodynamic Change in Patients with Carotid Artery Stenosis During the Tilt Test Using Wearable Near-Infrared Spectroscopy. In *Oxygen Transport to Tissue XXXV*. Springer, 463–467.

21. Shamsi T Iqbal, Xianjun Sam Zheng, and Brian P Bailey. 2004. Task-evoked pupillary response to mental workload in human-computer interaction. In *CHI'04 extended abstracts on Human factors in computing systems*. ACM, 1477–1480.

22. Kurtulus Izzetoglu, Scott Bunce, Banu Onaral, Kambiz Pourrezaei, and Britton Chance. 2004. Functional optical brain imaging using near-infrared during cognitive tasks. *International Journal of human-computer interaction* 17, 2 (2004), 211–227.

23. Robert JK Jacob. 2015. Phylter: A System for Modulating Notifications in Wearables Using Physiological Sensing. In *Foundations of Augmented Cognition: 9th International Conference, AC 2015, Held as Part of HCI International 2015, Los Angeles, CA, USA, August 2-7, 2015, Proceedings*, Vol. 9183. Springer, 167.

24. John T Kent. 1983. Information gain and a general measure of correlation. *Biometrika* 70, 1 (1983), 163–173.

25. Matthew A Killingsworth and Daniel T Gilbert. 2010. A wandering mind is an unhappy mind. *Science* 330, 6006 (2010), 932–932.

26. Veena Kumari, Steven CR Williams, Jeffrey A Gray, and others. 2004. Personality predicts brain responses to cognitive demands. *The Journal of neuroscience* 24, 47 (2004), 10636–10641.

27. Niels Landwehr, Mark Hall, and Eibe Frank. 2003. Logistic model trees. In *Machine Learning: ECML 2003*. Springer, 241–252.

28. Joseph E LeDoux. 1993. Emotional memory systems in the brain. *Behavioural brain research* 58, 1 (1993), 69–79.

29. Zengyong Li, Ming Zhang, Xiaoyin Zhang, Shixun Dai, Xingxin Yu, and Yan Wang. 2009. Assessment of cerebral oxygenation during prolonged simulated driving using near infrared spectroscopy: its implications for fatigue development. *European journal of applied physiology* 107, 3 (2009), 281–287.

30. Jessica Lin, Eamonn Keogh, Li Wei, and Stefano Lonardi. 2007. Experiencing SAX: a novel symbolic representation of time series. *Data Mining and knowledge discovery* 15, 2 (2007), 107–144.

31. Jinyi Long, Yuanqing Li, Hongtao Wang, Tianyou Yu, Jiahui Pan, and Feng Li. 2012. A hybrid brain computer interface to control the direction and speed of a simulated or real wheelchair. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on* 20, 5 (2012), 720–729.

32. MA Lopez-Gordo, E Fernandez, S Romero, F Pelayo, and Alberto Prieto. 2012. An auditory brain–computer interface evoked by natural speech. *Journal of neural engineering* 9, 3 (2012), 036013.

33. Kevin Mandrick, Gérard Derosiere, Gérard Dray, Denis Coulon, Jean-Paul Micallef, and Stéphane Perrey. 2013. Prefrontal cortex activity during motor tasks with additional mental load requiring attentional demand: a near-infrared spectroscopy study. *Neuroscience research* 76, 3 (2013), 156–162.
34. Dara S Manoach, Gottfried Schlaug, Bettina Siewert, David G Darby, Benjamin Martin Bly, Andrew Benfield, Robert R Edelman, and Steven Warach. 1997. Prefrontal cortex fMRI signal changes are correlated with working memory load. *Neuroreport* 8, 2 (1997), 545–549.
35. Derek Evan Nee, Joshua W Brown, Mary K Askren, Marc G Berman, Emre Demiralp, Adam Krawitz, and John Jonides. 2013. A meta-analysis of executive components of working memory. *Cerebral Cortex* 23, 2 (2013), 264–282.
36. Adrian M Owen, Kathryn M McMillan, Angela R Laird, and Ed Bullmore. 2005. N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human brain mapping* 25, 1 (2005), 46–59.
37. Sergio Paradiso, Debra L Johnson, Nancy C Andreasen, Daniel S OLeary, G Leonard Watkins, Laura L Boles Ponto, and Richard D Hichwa. 2014. Cerebral blood flow changes associated with attribution of emotional valence to pleasant, unpleasant, and neutral visual stimuli in a PET study of normal subjects. *American Journal of Psychiatry* (2014).
38. Martin P Paulus and Lawrence R Frank. 2003. Ventromedial prefrontal cortex activation is critical for preference judgments. *Neuroreport* 14, 10 (2003), 1311–1315.
39. Evan M Peck. 2014. *Designing Brain-Computer Interfaces for Intelligent Information Delivery Systems*. Ph.D. Dissertation. TUFTS UNIVERSITY.
40. Evan M Peck, Daniel Afergan, and Robert JK Jacob. 2013. Investigation of fNIRS brain sensing as input to information filtering systems. In *Proceedings of the 4th Augmented Human International Conference*. ACM, 142–149.
41. Tue Hvass Petersen, Maria Willerslev-Olsen, Bernard A Conway, and Jens Bo Nielsen. 2012. The motor cortex drives the muscles during walking in human subjects. *The Journal of physiology* 590, 10 (2012), 2443–2452.
42. Yann Renard, Fabien Lotte, Guillaume Gibert, Marco Congedo, Emmanuel Maby, Vincent Delannoy, Olivier Bertrand, and Anatole Lécuyer. 2010. OpenViBE: an open-source software platform to design, test, and use brain-computer interfaces in real and virtual environments. *Presence: teleoperators and virtual environments* 19, 1 (2010), 35–53.
43. Mauricio Rodriguez. 2011. *Development of a Versatile Wireless fNIR System*. Ph.D. Dissertation. Drexel University.
44. Shilad Sen, Jesse Vig, and John Riedl. 2009. Tagommenders: connecting users to items through tags. In *Proceedings of the 18th international conference on World wide web*. ACM, 671–680.
45. Juanning Si, Ruirui Zhao, Yujin Zhang, Nianming Zuo, Xin Zhang, and Tianzi Jiang. 2015. A portable fNIRS system with eight channels. In *SPIE BiOS*. International Society for Optics and Photonics, 93051B–93051B.
46. Edward E Smith and John Jonides. 1999. Storage and executive processes in the frontal lobes. *Science* 283, 5408 (1999), 1657–1661.
47. Erin Solovey, Paul Schermerhorn, Matthias Scheutz, Angelo Sassaroli, Sergio Fantini, and Robert Jacob. 2012. Brainput: enhancing interactive systems with streaming fnirs brain input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2193–2202.
48. Erin Treacy Solovey, Audrey Girouard, Krysta Chauncey, Leanne M Hirshfield, Angelo Sassaroli, Feng Zheng, Sergio Fantini, and Robert JK Jacob. 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*. ACM, 157–166.
49. Gary Strangman, Joseph P Culver, John H Thompson, and David A Boas. 2002. A quantitative comparison of simultaneous BOLD fMRI and NIRS recordings during functional brain activation. *Neuroimage* 17, 2 (2002), 719–731.
50. Charlotte C Teneback, Ziad Nahas, Andrew M Speer, Monica Molloy, Laurie E Stallings, Kenneth M Spicer, S Craig Risch, and Mark S George. 2015. Changes in prefrontal cortex and paralimbic activity in depression following two weeks of daily left prefrontal TMS. *The Journal of neuropsychiatry and clinical neurosciences* (2015).
51. Erin Treacy Solovey, Daniel Afergan, Evan M Peck, Samuel W Hincks, and Robert JK Jacob. 2015. Designing implicit interfaces for physiological computing: Guidelines and lessons learned using fNIRS. *ACM Transactions on Computer-Human Interaction (TOCHI)* 21, 6 (2015), 35.
52. Martijn P Van Den Heuvel and Hilleke E Hulshoff Pol. 2010. Exploring the brain network: a review on resting-state fMRI functional connectivity. *European Neuropsychopharmacology* 20, 8 (2010), 519–534.
53. Thorsten O Zander, Christian Kothe, Sabine Jatzev, and Matti Gaertner. 2010. Enhancing human-computer interaction with input from active and passive brain-computer interfaces. In *Brain-computer interfaces*. Springer, 181–199.
54. Quan Zhang, Gary E Strangman, and Giorgio Ganis. 2009. Adaptive filtering to reduce global interference in non-invasive NIRS measures of brain activation: how well and when does it work? *Neuroimage* 45, 3 (2009), 788–794.