Sensing Cognitive Multitasking for a Brain-Based Adaptive User Interface

Erin Treacy Solovey[1], Francine Lalooses[1], Krysta Chauncey[1], Douglas Weaver[1],
Margarita Parasi[1], Matthias Scheutz[1], Angelo Sassaroli[2], Sergio Fantini[2], Paul Schermerhorn[3],
Audrey Girouard[4], Robert J.K. Jacob[1]

[1]Tufts University	[2]Tufts University	[3]Indiana University	[4]Queen's University
Computer Science	Biomedical Engineering	Cognitive Science	School of Computing
161 College Ave.,	4 Colby St., Medford,	Bloomington, IN	25 Union St., Kingston,
Medford, MA 02155, USA	MA 02155, USA	47406, USA	ON K7L 3N6, Canada

{erin.solovey, francine.lalooses, krysta.chauncey, douglas.weaver, margarita.parasi, matthias.scheutz,
angelo.sassaroli, sergio.fantini}@tufts.edu, pscherme@indiana.edu, audrey@cs.queensu.ca, jacob@cs.tufts.edu

ABSTRACT

Multitasking has become an integral part of work environments, even though people are not well-equipped cognitively to handle numerous concurrent tasks effectively. Systems that support such multitasking may produce better performance and less frustration. However, without understanding the user's internal processes, it is difficult to determine optimal strategies for adapting interfaces, since all multitasking activity is not identical. We describe two experiments leading toward a system that detects cognitive multitasking processes and uses this information as input to an adaptive interface. Using functional near-infrared spectroscopy sensors, we differentiate four cognitive multitasking processes. These states cannot readily be distinguished using behavioral measures such as response time, accuracy, keystrokes or screen contents. We then present our human-robot system as a proof-of-concept that uses real-time cognitive state information as input and adapts in response. This prototype system serves as a platform to study interfaces that enable better task switching, interruption management, and multitasking.

Author Keywords

fNIRS, near-infrared spectroscopy, multitasking, interruption, brain computer interface, human-robot interaction

ACM Classification Keywords

H5.2 [Information interfaces and presentation]: User Interfaces. - Graphical user interfaces.

General Terms

Human Factors

INTRODUCTION

Multitasking has become an integral part of work environments, even though people are not well-equipped to effectively handle more than one task at a time [26]. While mul-

titasking has been shown to be detrimental to performance in individual tasks [26], it can also be beneficial when a secondary task provides additional information for completing the primary task, such as allowing people to integrate information from multiple sources.

Multiple windows, multiple monitors and large displays make it possible for the interface to handle multitasking, and many researchers have investigated how best to support the user who is balancing multiple tasks. Because multitasking can elicit several different cognitive states, the user's needs during multitasking may change over time. However, it is difficult to determine the best way to support the user without understanding the internal cognitive processes occurring during task performance.

In this paper, we describe a preliminary study and two experiments using neural data in which we identified four mental processes that may occur during multitasking and have direct relevance to many HCI scenarios. These processes are almost indistinguishable by examining overt behavior or task performance alone. However, using our non-invasive brain-sensing system (Figure 1) with functional near-infrared spectroscopy (fNIRS), we can automatically distinguish these four states. By detecting specific cognitive states that occur when multitasking, we can build user interfaces that better support task switching, interruption management and multitasking. We show an example of this with a proof-of-concept adaptive human-robot system that can change behavior based on brain signals received. This prototype system serves as a platform to provide the basis for designing and evaluating future brain-based adaptive user interfaces, with broader applications beyond human-robot team tasks.

Figure 1. Functional near-infrared spectroscopy is a portable non-invasive tool for detecting brain activity.
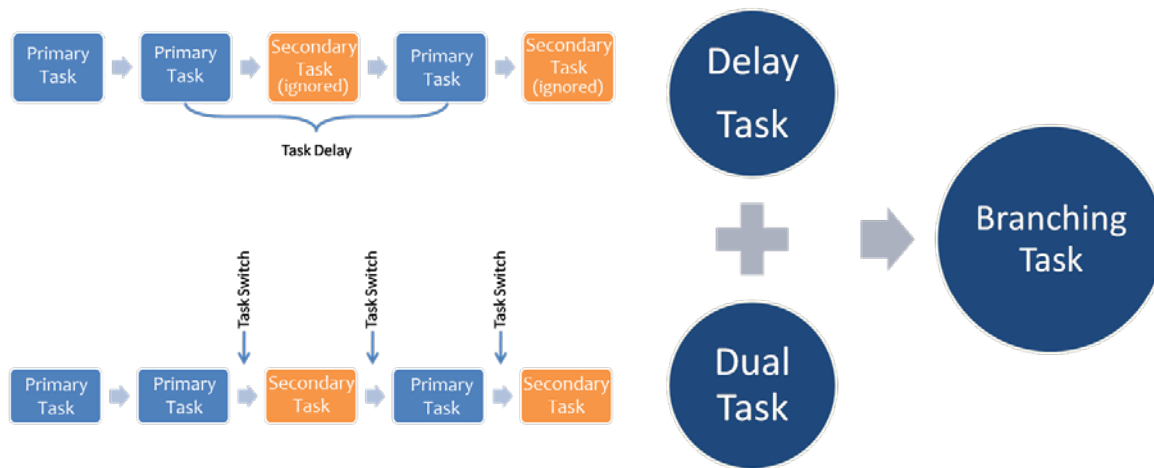
Figure 2. In the Delay scenario, the secondary task requires little attention, but the primary task goal is held in working memory. In the Dual-Task scenario, both primary and secondary tasks require attentional resources to be allocated for each task switch, but goals are not held in working memory. Branching has characteristics of both Delay and Dual-Task scenarios (Figure 3).

This paper makes the following contributions: (1)We show that specific cognitive multitasking states, previously studied with fMRI, can be detected automatically with fNIRS which is more practical in HCI. (2)We moved from a simple letter-based task in previous work to actual HCI-related tasks that elicit similar states. (3)We show a working proof-of-concept human-robot platform that supports adaptive behavior based on the cognitive states detected with fNIRS.

Related Work: Multitasking Support
Although computers are capable of handling multiple processes simultaneously, people have a difficult time due to high mental workload from increased working memory demands and the overhead of switching context between multiple tasks. Repeated task switching during an activity may lead to completion of the primary task with lower accuracy and longer duration, in addition to increased anxiety and perceived difficulty of the task [1]. The challenge is to devise an effective way to measure workload and attention-shifting in a dynamic environment, as well as to identify optimal support for multitasking.

Measuring Mental Workload and Other Cognitive States
Managing mental workload has long been an active topic in HCI research and high mental workload has been identified as a cause of potential errors [2]. Researchers have shown that different types of subtasks lead to different mental workload levels [15]. As a measure for mental workload, researchers have proposed pupil dilation [17] in combination with subjective ratings as this is non-invasive, and allows the user to perform the tasks as the data is processed in real time. Other physiological measures, including skin conductance, respiration, facial muscle tension and blood volume pressure, have also been used to detect cognitive or emotional states to improve machine intelligence [6, 23, 29]. While adaptive user interfaces may be designed to reduce mental workload, any automation may also result in reduced situation awareness, increased user complacency and skill degradation, and these human performance areas should be evaluated in the system [28].

Task Switching and Measuring Interruptibility
When managing multiple tasks, interruptions are unavoidable. To address this, researchers have developed systems that try to identify the cost associated with interruption based on different inputs, such as desktop activity, environment context [7, 13, 35], eye tracking [12], or other physiological measures such as heart rate variability and electromyogram [4] and handle interruptions accordingly. They have found interruptions to be less disruptive during lower mental workload [16, 30]. Other studies tried placing interruptions near the beginning, middle or end of a task [5], at task boundaries [25], or between repetitive tasks which were considered as more interruptible [27]. It was also shown that interruptions relevant to the main task tend to be less disruptive for the users than irrelevant interruptions [5].

Various interruption schemes may affect performance in different ways; however, there is no universally optimal interruption scheme. Interrupting the user as soon as the need arises, for example, emphasizes task completeness over accuracy, while allowing the user to defer interruptions indefinitely does the opposite [31]. McFarlane [24] discusses four distinct methods for coordinating interruption—immediate, negotiated (user selects when to be interrupted), mediated (an intelligent agent selects when to interrupt), and scheduled (interruptions appear at fixed times)—and found that no optimal method existed across users and tasks. Thus, it is crucial that the style of interruption adapts to the task. Systems have been developed that quantify the optimal time to interrupt a user by weighing the value against the cost of interruption [15]. In addition to determining the optimal time for switching tasks, researchers have tried to determine the best method for reminding users of pending background tasks. Miyata and Norman [26] note that important alerts specifically designed for someone who is deeply engaged in another task would most likely be inappropriate and may even be disruptive in other situations.
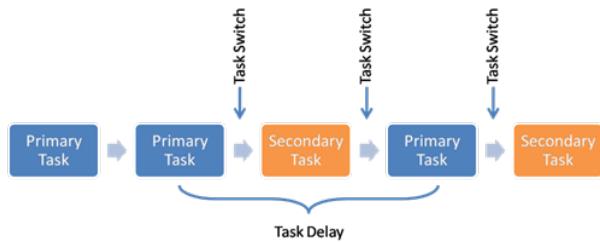
Figure 3. Branching: Primary and secondary task both require attentional resources to be allocated, and the primary task goal must be kept in mind over time.

Figure 4. Conditions from [19]. Stimuli were either uppercase (red) or lowercase (blue) letters from the word "tablet" and the response varied depending on the case. Check indicates a match and x indicates a non-match stimulus.

Brain Sensing for HCI

Real-time cognitive state information can inform the tradeoffs to create intelligent user interfaces. Design of user interfaces that employ real-time cognitive state information has become an emerging topic in HCI recently (for an overview, see [18]). Much of this work has used brain sensing as explicit input to the system to make selections or control the interface [21, 36], although there have been examples of passive brain sensing to be used either as implicit input or for evaluation of user interfaces [9, 11, 22]. Our work focuses on using fNIRS sensors to detect signals users implicitly produce while interacting naturally with a system. These sensors detect changes in oxygenated and deoxygenated blood in a region of the brain by using optical wires to emit near-infrared light [3]. The sensors are easy to use, have a short set-up time and are portable, all characteristics which make fNIRS suitable for use in realistic HCI settings [34]. However, because it is a novel technique for brain sensing, there have been few studies showing specific measurements with fNIRS and their appropriate use in HCI.

Multitasking Scenarios: Branching, Dual Task, Delay

Multitasking behavior involves several high-level brain processes, which vary depending on the types of tasks and the interaction between the tasks. Koechlin et al. [19] described three distinct, but related multitasking scenarios, which they refer to as *branching*, *dual-task*, and *delay*. These are the foundation for the studies described here.

Branching (Figures 2,3) is illustrated by the following scenario: *A user is tackling a complex programming task but is interrupted by an incoming email from her boss that is time sensitive*. Thus, the user must "hold in mind goals while exploring and processing secondary goals" [19]. Branching processes are triggered frequently in multitasking environments and pose a challenge to users.

However, some situations may involve frequent task switching without the need to maintain information about the previous task (e.g. *A user is monitoring and responding to high priority software support issues that are logged by clients as well as responding to important emails, and regularly switches between the two tasks*). These tasks are referred to as *dual-task* because there are two tasks that require attentional resources (Figure 2). These situations could also utilize adaptive support in the user interface, but
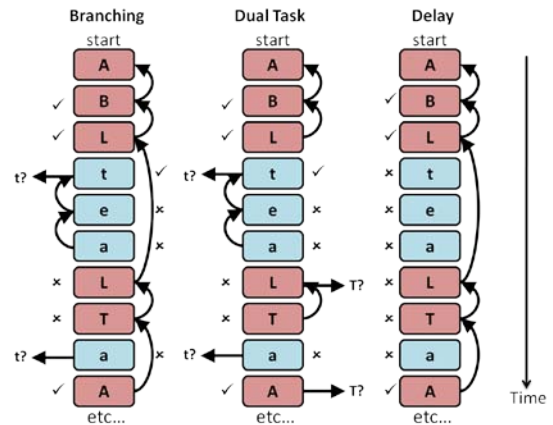
the adaptive behavior would be distinct from that of *branching*.

The third multitasking paradigm is illustrated with the following scenario: *A user is tackling a complex programming assignment and at the same time gets instant messages which the user notices, but ignores*. Here, the secondary task is ignored and therefore requires little attentional resources. They refer to this as *delay* because the secondary task mainly delays response to the primary task (Figure 2).

In their experiment, Koechlin et al. demonstrated using functional Magnetic Resonance Imaging (fMRI) that these three multitasking processes have different activation profiles in the prefrontal cortex of the brain, particularly in Brodmann's Areas 8, 9 and 10. Their task involved processing rules based on letters appearing on the screen. Each stimulus was either an uppercase or lowercase letter from the word "tablet." The expected response from the user was different depending on the case of the letter, so switching between uppercase and lowercase letters would be similar to balancing two tasks. There were four conditions in their experiment, each with different rules for responding, designed to trigger specific multitasking behavior (Figure 4):

1) *Delay*: Are two consecutive *uppercase* stimuli in immediate succession in the word "TABLET"? Ignore *lowercase*.
2) *Dual-Task:* Are two consecutive stimuli *of the same case* in immediate succession in the word tablet? When the case changes, is the first letter in the series a 'T' or 't'?
3) *Branching*: For *uppercase* stimuli, respond as in *Delay*. If the letter is *lowercase*, respond as in *Dual Task*.
4) *Control*: Are two consecutive stimuli in immediate succession in "TABLET"? All stimuli were uppercase.

Koechlin et al. [20] later showed that even during branching, there were distinct activation profiles that varied depending on whether the participant could predict when task switching would occur or whether it was random. The ex-
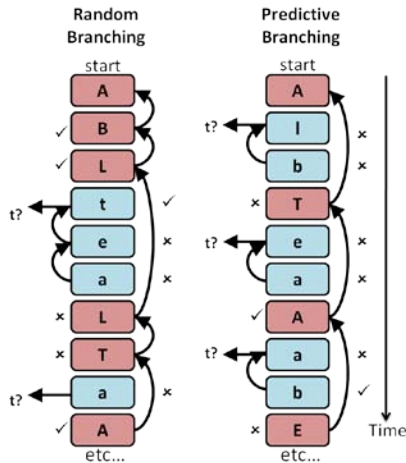
Figure 5. Experimental conditions from Koechlin et al. [20].

perimental setup was almost identical to the earlier study, except that in all conditions, the *branching* paradigm was used. There were two experimental *branching* conditions (Figure 5) and a control:

1) *Random Branching:* Upper- and lower-case letters were presented pseudorandomly.
2) *Predictive Branching:* Uppercase letters were presented every 3 stimuli.
3) *Control Branching:* The same six-letter sequence (A e t a B t) was shown repeatedly.

The significance of these two experiments lies in the fact that all experimental conditions had the same stimuli and the same possible user responses, so the conditions could not be easily distinguished from one another by simply observing the participant. Using fMRI, however, it became possible to distinguish the conditions based on the distinct mental processes (and thus, distinct blood flow patterns) elicited by each task.

In addition, the cognitive states identified in these experiments have direct relevance to many HCI scenarios, particularly when a user is multitasking. Automatically recognizing that the user is experiencing one of these states provides an opportunity to build adaptive systems that support multitasking. For example, by recognizing that most interruptions are quickly ignored, as in the *delay* condition, the system could limit these types of interruptions or reduce their salience as appropriate. Further, if a user is currently experiencing a *branching* situation, the interface could better support maintaining the context of the primary task, whereas during *dual-task* scenarios this would be unnecessary. Finally, distinguishing between *predictive* and *random* scenarios could trigger the system to increase support when the user's tasks become unpredictable.

This paper builds from their experiments with the goal of designing interfaces that recognize these states and behave in appropriate ways to support multitasking. We conducted a preliminary study to reproduce the results of Koechlin et

al. [19] using fNIRS which is practical for HCI settings unlike fMRI in which slight movement can create motion artifacts and corrupt the image [34]. We then followed with two experiments that look at distinguishing the cognitive multitasking states in other scenarios besides the "tablet" task to investigate whether these are generic cognitive processes, and not simply tied to the particular task used in the earlier study. Finally, we designed and built a proof-of-concept platform that recognizes and classifies the fNIRS signal and uses it as input to drive an adaptive human-robot system.

PRELIMINARY STUDY

Our preliminary experiment extends Koechlin et al.'s work [19] to more realistic HCI settings. We wanted to determine whether we could distinguish between *branching*, *dual-task* and *delay* situations. These states were successfully distinguished using fMRI [19], but fMRI is not practical in HCI settings. Our hypothesis was that the same could be achieved using fNIRS. Since the sensors are placed on the forehead, they are particularly sensitive to changes in the anterior prefrontal cortex, where Koechlin et al. [19] showed distinct activation profiles during *delay*, *dual* and *branching* tasks.

Three participants wore fNIRS sensors as they performed the experimental tasks. To trigger the three cognitive states, we used the same experimental paradigm used in [19].

To determine whether these tasks could be distinguished, we performed leave-one-out cross validation in Weka [10] to classify the fNIRS sensor data. In MATLAB, the fNIRS signal was detrended by fitting a polynomial of degree 3 and then a low-pass elliptical filter was used to remove noise in the data. Using support vector machines, we achieved reasonably high accuracy classifying the tasks across the three participants (68.4% mean across three pairwise classifications, and 52.9% accuracy for three-way classification). This was a small sample of users, and we hope to achieve higher accuracy, but found the results encouraging enough continue in this research direction.

MULTITASKING EXPERIMENTS

From the promising results of the preliminary study, we investigated whether we could detect these three states in other tasks and domains that are more relevant to interactive user interfaces. Our hypothesis was that the cognitive functions elicited in the "tablet" tasks were generic processes that occur during multitasking. Numerous HCI scenarios involve multitasking, and we chose a human-robot team scenario to further explore the detection of cognitive multitasking in user interfaces.

Multitasking in Human Robot Interaction

Human-robot team tasks inherently involve multitasking, as the user is both performing his or her part of the task, while monitoring the state of the robot(s). Thus, these tasks provide an appropriate example for studying adaptive multitasking support, and may see improved performance with
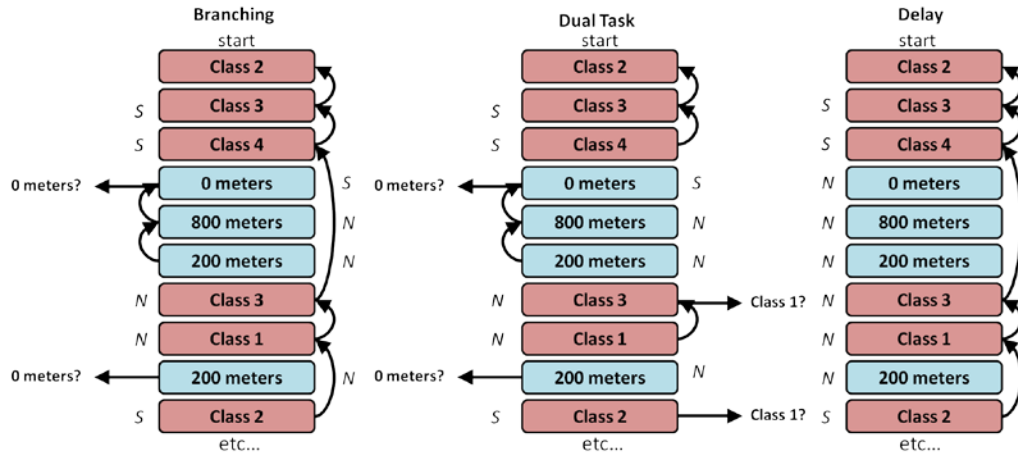
Figure 6. Stimuli and responses for conditions in Experiment 1. These conditions are analogous to those in [19]. (See Figure 4).

brain-based adaptive interfaces. Thus, the simple word-related task was replaced by a human-robot interaction (HRI) task that has similar properties.

Experimental Tasks
We conducted two separate experiments which built from the human-robot team task described by Schermerhorn and Scheutz [32] and adjusted it to include tasks that would induce *delay*, *dual-task* and *branching*, similar to our preliminary study. The tasks involved a human-robot team performing a complex task that could not be accomplished by the human nor the robot alone. The robot and the human had to exchange information in order to accomplish the task. The robot continually updated the human operator with status updates to which the human responded.

In the two separate studies, the participant worked with a robot to investigate rock types on the surface of Mars and had to perform two tasks. The robot presented the participant with status updates, either about a newly found rock or a new location to which it moved. Each rock classification update informed the user of the newly discovered rock's class, which was based on size and ranged from Class 1 to Class 5. Each location update alerted the user of the robot's current location. The spacecraft to which the robot was transmitting could detect the robot's location to the nearest kilometer and assumed the robot was moving in a straight line. Thus, the location updates presented to the user ranged from 0 to 800 meters, in 200 meter increments.

The participant's primary task was to sort rocks, and the secondary task was to monitor the location of the robot. Each time the participant received a status update from the robot (in the form of a pop-up on the screen), s/he had two possible responses: either respond with the left hand by typing "S" to signify *same* or the right hand by typing "N" to signify *new*. After a rock classification, "S" instructed the robot to store the rock in the *same* bin, while "N" instructed the robot to store the rock in a *new* bin. After a location update, "S" instructed the robot to maintain the *same* transmission, while "N" instructed the robot to begin

a *new* transmission. The correct response after a particular update varied among the conditions.

Experiment 1: Delay, Dual-Task & Branching
The first experiment contained three conditions, analogous to those in [19], each with its own rules for the user response (Figure 6):

Delay: Do two successive rock classification messages follow in immediate consecutive order? If so, put it in the *same* bin. If not, select a *new* bin. For all location updates, begin a *new* transmission.

Dual-Task: Do two successive messages *of the same type* follow in immediate consecutive order? If so, select the *same* rock bin or maintain the *same* transmission. If the update is of a different type (switch task between rock and location), is the message either a Class 1 rock or a location of 0 meters? If so, select the *same* rock bin or maintain the *same* transmission. In all other cases, place the rock in a *new* bin or begin a *new* transmission.

Branching: For *rock classification* messages, respond as in *Delay*. If the update is a *location*, respond as in *Dual Task*.

Participants
This study included 12 healthy volunteers (10 male), between the ages of 18 and 34. Four additional volunteers had participated in the study, but are not included in this analysis because their performance in the tasks was below 70% in more than two trials per condition, indicating that they were not correctly performing the tasks. In addition, data from another participant is not included due to technical problems with the fNIRS system. All participants were right-handed, had English as their primary language, had no history of brain injury and had normal or corrected-to-normal vision. Informed consent was obtained for all participants. This experiment was approved by our institutional review board.

Design and Procedure
Before the experiment, each participant was given the opportunity to become familiar with each of the three tasks
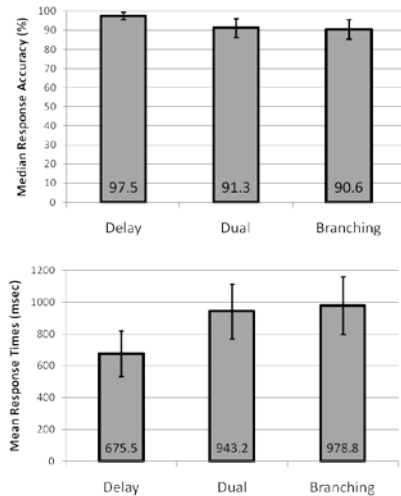
Figure 7. Behavioral results for Experiment 1: median accuracy & standard deviation (top); mean response time and standard deviation (bottom).

during a practice session without the fNIRS sensors. The conditions were presented in counterbalanced pseudo-random order. Each task was repeated until the participant achieved greater than 80% accuracy in the task. After this accuracy was achieved for all three conditions, the fNIRS sensors were placed on the participant's forehead. The participant was presented with an initial rest screen, which was used to collect a baseline measure of the brain activity at rest. After that, the user had to complete ten 40-second trials for each of the three conditions, which were presented randomly. Between each task, the user was presented with the instructions for the next task, followed by a rest screen.

Equipment
We used a multichannel frequency domain OxiplexTS from ISS Inc. (Champaign, IL) for data acquisition. Two probes were placed on the forehead to measure the two hemispheres of the anterior prefrontal cortex (Figure 1). The source-detector distances were 1.5, 2, 2.5, and 3cm. Each distance measures a different depth in the cortex. Each source emits two light wavelengths (690nm and 830nm) to detect and differentiate between oxygenated and deoxygenated hemoglobin. The sampling rate was 6.25Hz.

Results
To examine the differences between the three task conditions, we looked at behavioral data collected during the experiment as well as the fNIRS sensor data. In both experiments, any trials where the participant achieved less than 70% accuracy in the task performance were removed in the analysis, since this would indicate that the subject was not actually performing the task correctly.

Behavioral Results: In the three conditions, the stimuli were essentially the same, as were the possible responses. Thus, it would be difficult for an observer to detect any difference from the screen contents or the subject's behavior alone.

Like the sensor data, response time and accuracy measurements can be obtained automatically without interfering with the task so we investigated whether they would vary depending on the condition.

Statistical analysis was performed utilizing the InStat statistical package by GraphPad Inc. All variables were tested for normal distribution with the Kolmogorov-Smirnov test. For normal distributions, the repeated measurements one-way analysis of variance (ANOVA) with the Tukey post-hoc test for multiple comparisons was used. For non-Gaussian distributions, we used the Friedman (non parametric repeated measurements ANOVA) test. The level of statistical significance was set at 0.05 (Figure 7).

Since *dual task* and *branching* behavioral results are similar, the factor was not significant overall, but is in pairwise comparisons. We found statistical significance in response time between *delay* and *dual* ($p < 0.001$), *delay* and *branching* ($p < 0.001$), but not between *dual* and *branching* ($p > 0.05$). Similarly, we found statistical significance in accuracies between *delay* and *dual* ($p < 0.05$), *delay* and *branching* ($p < 0.05$), but not *dual* and *branching* ($p > 0.05$). Also, correlations between accuracy and response time for each task were not statistically significant. We also looked at learning effects based on response time and learning effects based on accuracies as users progressed through the experiment. We did not find a learning effect.

Statistical Analysis of Signal: We wanted to determine whether the hemodynamic response measured by fNIRS has a different signature between the three conditions. For each of the two probes, we selected the fNIRS measurement channels with the greatest source-detector distances (3cm), as these channels are expected to probe deepest in the brain tissue, while the closer channels are more likely to pick up systemic effects and noise. From each of these channels, we calculated both the change in oxygenated hemoglobin and deoxygenated hemoglobin using the modified Beer-Lambert law [3] after removing noise with a band pass filter. Thus, we used four channels corresponding with
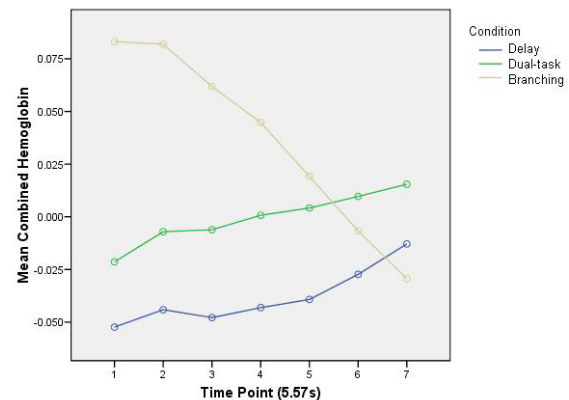

Figure 8. Combined oxygenated and deoxygenated hemoglobin by condition for Experiment 1.

changes in oxygenated and deoxygenated hemoglobin on the left and right hemispheres. Since the hemodynamic changes occur over a 5-7 second period, we simplified the signal for analysis by dividing the time series measurement for each trial into seven segments (~5.57 second each) and took the mean over these segments for the four channels.

In order to confirm that there were differences in brain activity during the three conditions, we did an ANOVA comparing condition means within subjects. Since there were multiple sensors, factors for the distribution of sensors were included (left/right hemisphere), as well as a factor for hemoglobin type (oxygenated or deoxygenated) and the time point. We used the Greenhouse-Geisser ANOVA values to correct for violations in sphericity. We found a main effect of condition (F(2,22)=4.353, p=0.029), in which total hemoglobin measures were overall higher in the branching condition than in the dual-task or delay condition (Figure 8). There were no other significant effects in this analysis.

Experiment 2: Random & Predictive Branching

To follow up on the first study, we conducted a second experiment to determine whether we could distinguish specific variations of the branching task. This experiment had two conditions that were analogous to those in [20], in which the participant was always following the *branching* rules described in Experiment 1:

Random Branching: Rock classification and location update messages were presented pseudorandomly.

Predictive Branching: Rock classification messages were presented every three stimuli.

Ideally, when using computer systems, the default scenario for a user would be similar to the *predictive* condition, and therefore the user would be able to plan ahead and handle incoming work appropriately. If we could automatically identify that the user is experiencing *random* or unpredictable behavior, there may be appropriate adaptations that the system could make to better support the user, which we are exploring with the adaptive interface platform described below. This experiment investigates whether we can automatically detect the different scenarios using fNIRS.
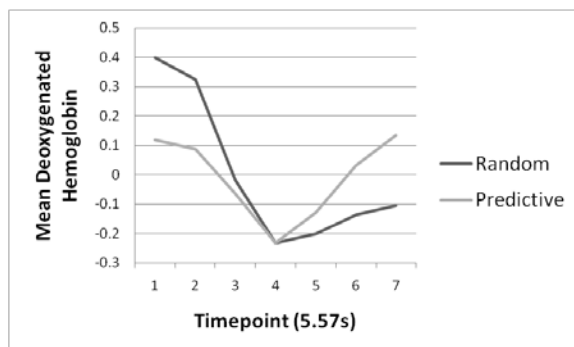

Figure 9. Average of deoxygenated hemoglobin over all subjects by condition for Experiment 2.

Participants
This study included 12 healthy volunteers (5 male), between the ages of 19 and 32. Three additional volunteers had participated, but are not included in this analysis because their performance in the tasks was below 70% in more than two trials per condition, indicating that they were not correctly performing the tasks. In addition, data from another participant was not included due to technical issues with the fNIRS system.

Design, Procedure & Equipment
This experiment used the same procedure and equipment as in Experiment 1. However, in this experiment, there were only two experimental conditions as described above and the participants completed eighteen trials of each condition, which were counterbalanced.

Results
Behavioral Results: As in Experiment 1, we collected response time and accuracy throughout the study to determine whether the conditions elicited different measurements.

Statistical analysis was performed utilizing the InStat statistical package by GraphPad Inc. All variables were tested for normal distribution with the Kolmogorov-Smirnov test. For normal distributions, a paired t-test was used. For non-Gaussian distributions, we used the Wilcoxon matched-pairs signed-ranks test.

There was no statistically significant difference in response time between *random* (*M*=998.67, *SD*=190.02) and *predictive* (*M*=992.81, *SD*=213.34) branching, *t*(215)=0.53 (p>0.05). There also was no statistically significant difference in accuracy between *random* (*M*=93.982, SD=8.144) and *predictive* (M=92.824, SD=8.765) branching (p>0.05). Also, correlation between accuracy and response time for *random* branching was not statistically significant (p>0.05), but there was a statistically significant correlation in the *predictive* branching condition (p<0.0001).

Statistical Analysis of Signal: Our goal was to determine whether the hemodynamic response measured by fNIRS has a different signature between the two conditions. Our analysis was the same as in Experiment 1.

We found an interaction between branching type, timepoint, and hemoglobin type (F(6,66)=3.035, p=0.038). This effect indicates that, although there was no significant change in oxygenated hemoglobin, deoxygenated hemoglobin levels were higher in the random branching type than the predictive branching type for the first half of trials, but reversed for the second half (Figure 9). Therefore, it should be possible to distinguish these two conditions using only the deoxygenated hemoglobin measure from fNIRS, which is in accordance with Huppert et al. who found that the magnetic blood oxygenation level dependent (BOLD) signal used in fMRI is most tightly tied to deoxygenated hemoglobin in fNIRS [14]. There were no other significant effects in this analysis.

PROOF OF CONCEPT BRAIN-BASED ADAPTIVE USER INTERFACE PLATFORM

In the two experiments described above, we verified that there is a significant difference between the cognitive multitasking conditions in the fNIRS signal. Because we can statistically differentiate them, we can apply machine learning techniques to automatically identify these cognitive multitasking states, in real time, in a user interface. This information could then be used to drive the user interface to better support cognitive multitasking.

As a proof-of-concept, we developed a platform for studying brain-based adaptive user interfaces of this type. The system has two main components: the Online fNIRS Analysis and Classification (OFAC) system [8] and the Distributed Integrated Affect, Reflection, Cognition Architecture (DIARC) [33] for human-robot interaction (Figure 10).

The OFAC system receives raw fNIRS signals from Boxy acquisition software (ISS, Inc.), classifies them in real time and sends the classification results to a specialized DIARC component. In addition to the normal *online* mode where signals are classified and sent in real time, OFAC supports *offline* mode which simulates the analysis and classification of previously recorded data, and will be useful in experimenting with various adaptive strategies. As OFAC receives data, it stores the signal in a database and preprocesses the signal (filtering, conversion to oxy- and deoxy- genated hemoglobin values using the modified Beer-Lambert law [3]), and performs classification using Weka [10].

The DIARC [33] is an integrated architecture for working with complex robots that can support various levels of robot autonomy and other adaptive robot behavior. The architecture can interface with physical robots in an environment, but also has a simulation mode that allows for simulated interactions with a robot on a computer screen (Figure 11), along with several different environment configurations.

To receive input from the OFAC system, we have created a DIARC component to which OFAC sends cognitive multitasking state classification results via sockets. DIARC can then use these messages to change the robot's goal structures, allowing the robot to adapt its autonomy and behavior. In addition, we have provided a simulation mode, where cognitive state classifications can be entered manually by inputting classification confidence levels for each of the possible states. This allows for debugging and testing
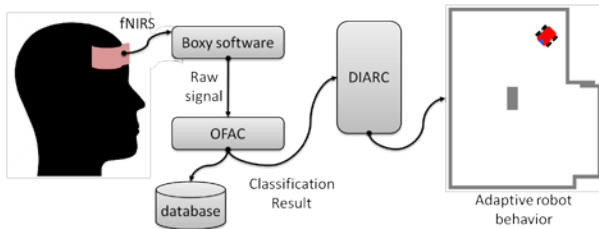

Figure 10. Brain-based adaptive user interface platform.

of the robot adaptive behaviors, without requiring a human to be physically connected to the fNIRS sensing system.

Integrating the OFAC system with DIARC provides the robot with cognitive state information of the human, affording the robot to adapt its behavior to better support and collaborate with the human operator. Driven by fNIRS cognitive multitasking input, DIARC can adapt various aspects of the human-robot interface, such as level of autonomy of the robot, the frequency, length and style of the robot's status updates, as well as the tone of the robot's voice.

As an initial step, we created a task scenario and simple adaptive behavior based on Experiment 1 described above to provide a technical demonstration that the platform functions as intended and can be further expanded to study adaptive multitasking user interfaces. The system was used to complete a human-robot team task where the primary task was to classify rocks and the secondary task was to monitor robot location as in the two experiments above.

When a *branching* state was identified—indicating that the user was tending to multiple tasks and maintaining information about the primary task over time—the robot switched to autonomy mode allowing the user to focus on the primary task of classifying rocks. In autonomy mode, the robot would move autonomously to new locations and find the best location for a new transmission. When a *non-branching* state was recognized, the robot returned to the default behavior, requiring the human to give instructions to the robot about whether it needed to start a new transmission. Staying in autonomous mode throughout the entire task would not be ideal as the human needs to be aware of the robot's location and progress to provide corrective feedback (as the robot's search behavior is not optimal), ensuring that the message is transmitted before time is up.

To validate that *branching* and *non-branching* could be classified in real time and that the robot could receive the signal and adapt its behavior, we did two demonstrations. First, utilizing the *offline* mode of OFAC, we sent previously recorded fNIRS data through the system as if it were being collected on a user in real time. We used the previously recorded data from Experiment 1. The next demonstration of the platform was to show that it can perform in *online* mode while a user was wearing fNIRS sensors.

In both cases, we used the *branching* and *delay* tasks described in Experiment 1 to invoke known *branching* and *non-branching* states. There was a *training* session where the *branching* and *delay* tasks from Experiment 1 were performed (or were simulated using the data collected in Experiment 1), allowing the robot to build a classification model based on the fNIRS sensor data during those tasks. Each condition was completed 10 times in random order to provide training data. Once the training session ended, the system switched to *testing* phase, where the fNIRS signal was

classified using the model that had been built during the *training* session. The entire session took about 30 minutes.

We established that fNIRS data was transmitted from OFAC to DIARC in real time, with fNIRS and robot components running in full operational mode. Robot goals and behaviors adapted based on the fNIRS data. This proof-of-concept confirms that our platform can form the basis of future study on brain-based adaptive user interfaces. We intend to enhance the machine learning techniques to improve the accuracy of the system. We will also analyze the fNIRS response more deeply in our future work, but our initial goal was show that there was a significant difference between the signals we detected for the different conditions. This allows us to discriminate the conditions, and adapt a user interface when each state is detected. In addition, we will expand adaptive behaviors that are supported to study the tradeoffs for making such adaptations.

DISCUSSION AND CONCLUSION

This paper builds a foundation for brain-based adaptive user interfaces. First, in our preliminary study, we brought research on cognitive activity during multitasking to a system that is practical for HCI by showing that fNIRS sensors could detect states previously studied with fMRI (which cannot be used in HCI settings). In our next two experiments, we further extended this research to HCI by showing that the states elicited in the "tablet" task may be generic processes that occur in more realistic HCI tasks, by using a human-robot scenario. We integrated these findings with the brain-based adaptive interface platform we developed, in order to augment user interfaces with this cognitive multitasking state as input.

We demonstrated that the platform can form the basis for brain-based user-interfaces by implementing a simple adaptive scheme based on the experiments. However, to design a successful adaptive user interface many factors must be considered and the automation scheme will be more complex than that illustrated here. Parasuraman et al. [28] propose a framework for supporting user cognition with automation which describes four stages of human information processing (sensory processing, perception, decision making and response selection), each of which can have a different automation level. They also outline criteria for eva-

luating the user interface by examining both human performance measures and also system performance criteria (automation reliability, costs of action outcomes). This framework provides guidelines but does not prescribe specific adaptive behavior for every system. Instead each system must be carefully evaluated and iteratively designed to meet the needs of the users. Our platform will enable us to conduct evaluations of various adaptive behaviors to determine the appropriate strategy for supporting multitasking by utilizing signals coming implicitly from the brain.

Human-robot team tasks provide appropriate scenarios for studying adaptive multitasking support, as they inherently involve multitasking: the user is performing a task, while also monitoring the state of the robot(s). HRI team tasks thus may see improved performance with brain-based adaptive interfaces. There has been much work on adaptive robots that change behavior based on the environment or situation. We plan to develop robots that have a greater understanding of the user's cognitive state during multitasking, and can adapt their behavior to better support the user.

In addition, we believe that similar brain-based user interfaces may support a wide range of contexts that involve multitasking and interruptions, such as air traffic control, management of unmanned aerial vehicles (UAVs), complex data visualization and analytics, and many others. Non-invasive brain sensing provides an additional channel of input to the system without any explicit action by the user. However, any adaptation must be done carefully, to ensure that the user does not feel that he or she has lost control. The platform we have developed will allow us to explore adaptive behavior to find the best strategies for these scenarios. We can now begin to develop complex systems that adapt based on fNIRS brain signals and experimentally evaluate them.

REFERENCES
1. Bailey, B.P., et al. The Effects of Interruptions on Task Performance, Annoyance, and Anxiety in the User Interface. *INTERACT'01*, 2001, 593-601.

2. Card, S., Moran, T. and Newell, A. *The Psychology of Human Computer Interaction*. Lawrence Erlbaum Associates, Hillsdale, NJ, 1983.

3. Chance, B., et al. A novel method for fast imaging of brain function, non-invasively, with light. *Optics Express*, 10 (2). 411-423.

Figure 11. The Brain-Adaptive Human-Robot Interaction Platform supports robot navigation (real or simulated), as well as fNIRS cognitive state input (real or simulated).

4. Chen, D., et al. Towards a Physiological Model of User Interruptability. *INTERACT'07*, 2008, 439-451.

5. Czerwinski, M., et al. Instant Messaging: Effects of Relevance and Timing. *HCI*, 2000, 71-76.

6. Fairclough, S.H. Fundamentals of physiological computing. *Interact. Comput.*, 21 (1-2). 133-145.

7. Fogarty, J., Hudson, S.E. and Lai, J. Examining the robustness of sensor-based statistical models of human interruptibility. *CHI'04*, ACM, 2004.

8. Girouard, A., Solovey, E.T. and Jacob, J.K. Designing a Passive Brain Computer Interface using Real Time Classification of Functional Near-Infrared Spectroscopy. *Intl J of Auton. & Adaptive Comm. Sys.* (2010). in press.

9. Grimes, D., et al. Feasibility and pragmatics of classifying working memory load with an electroencephalograph. *CHI'08*, ACM, 2008.

10. Hall, M., et al. The WEKA data mining software: an update. *SIGKDD Explor. Newsl.*, 11 (1). 10-18.

11. Hirshfield, L., et al. Brain Measurement for Usability Testing and Adaptive Interfaces: An Example of Uncovering Syntactic Workload with Functional Near Infrared Spectroscopy. *CHI'09*, 2009.

12. Hornof, A.J.,et al. Knowing where and when to look in a time-critical multimodal dual task. *CHI '10*, 2010.

13. Hudson, S., et al. Predicting human interruptibility with sensors: a Wizard of Oz feasibility study. *CHI '03*, 2003.

14. Huppert, T., et al. A temporal comparison of BOLD, ASL, and NIRS hemodynamic responses to motor stimuli in adult humans. *Neuroimage*, 2006:29(2). 368-382.

15. Iqbal, S.T., et al. Towards an index of opportunity: understanding changes in mental workload during task execution. *CHI '05*, 2005.

16. Iqbal, S.T. and Bailey, B.P. Investigating the effectiveness of mental workload as a predictor of opportune moments for interruption. *CHI extended abstracts*, 2005.

17. Iqbal, S.T., Zheng, X.S. and Bailey, B.P. Task-evoked pupillary response to mental workload in human-computer interaction. *CHI extended abstracts*, 2004.

18. Jackson, M.M. and Mappus, R. Applications for Brain-Computer Interfaces. in Tan, D. and Nijholt, A. eds. *Brain Computer Interfaces: Applying our Minds to Human-Computer Interaction*, Springer, 2010, 89-104.

19. Koechlin, E., et al. The role of the anterior prefrontal cortex in human cognition. *Nature*, 1999, 148-151.

20. Koechlin, E., et al. Dissociating the role of the medial and lateral anterior prefrontal cortex in human planning. *PNAS*, 97. 7651-7656.

21. Kuikkaniemi, K., et al. The influence of implicit and explicit biofeedback in first-person shooter games. *CHI'*10, 2010.

22. Lee, J.C. and Tan, D.S. Using a low-cost electroencephalograph for task classification in HCI research *UIST'06*, ACM Press, 2006.

23. Mandryk, R.L. and Inkpen, K.M. Physiological indicators for the evaluation of co-located collaborative play *CSCW '04*, 2004.

24. McFarlane, D. Comparison of four primary methods for coordinating the interruption of people in human-computer interaction. *Hum.-Comput. Interact.*, 17(1). 63-139.

25. Miyata, Y. and Norman, D. The Control of Multiple Activities. in Norman, D. and Draper, S.W. eds. *User Centered System Design: New Perspectives on Human-Computer Interaction*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1986.

26. Miyata, Y. and Norman, D. Psychological Issues in Support of Multiple Activities. in Norman, D.A. and Draper, S.W. eds. *User Centered System Design: New Perspectives on Human-Computer Interaction*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1986, 265-284.

27. Monk, C., et al. The attentional costs of interrupting task performance at various stages. *HFES Ann. Mtg*, 2002.

28. Parasuraman, R., et al. A model for types and levels of human interaction with automation. *IEEE Trans. on Systems, Man and Cybernetics*, 30 (3). 286-297.

29. Picard, R.W., et al. Toward Machine Emotional Intelligence: Analysis of Affective Physiological State. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23 (10). 1175-1191.

30. Salvucci, D.D. and Bogunovich, P. Multitasking and monotasking: the effects of mental workload on deferred task interruptions. *CHI'*10, 2010.

31. Sasse, A., et al. Coordinating the Interruption of People in Human-Computer Interaction. *INTERACT*, 1999.

32. Schermerhorn, P. and Scheutz, M. Dynamic robot autonomy: investigating the effects of robot decision-making in a human-robot team task. *ICMI*, 2009.

33. Scheutz, M., et al. First steps toward natural human-like HRI. *Auton. Robots*, 22 (4). 411-423.

34. Solovey, E.T., et al. Using fNIRS Brain Sensing in Realistic HCI Settings: Experiments and Guidelines. *UIST'09*, 2009.

35. Starner, T., Schiele, B. and Pentland, A. Visual Contextual Awareness in Wearable Computing. *ISWC*, 1998.

36. Yuksel, B.F., et al. A novel brain-computer interface using a multi-touch surface. *CHI'10*, 2010