Distinguishing Difficulty Levels with Non-invasive Brain Activity Measurements

Audrey Girouard[1], Erin Treacy Solovey[1], Leanne M. Hirshfield[1],
Krysta Chauncey[1], Angelo Sassaroli[2], Sergio Fantini[2], and Robert J.K. Jacob[1]

[1] Computer Science Department
[2] Biomedical Engineering Department Tufts University
Medford, MA 02155, USA
{audrey.girouard,erin.solovey,leanne.hirshfield,krysta.chauncey,
angelo.sassaroli,sergio.fantini,robert.jacob}@tufts.edu

Abstract. Passive brain-computer interfaces are designed to use brain activity as an additional input, allowing the adaptation of the interface in real time according to the user's mental state. The goal of the present study is to distinguish between different levels of game difficulty using non-invasive brain activity measurement with functional near-infrared spectroscopy (fNIRS). The study is designed to lead to adaptive interfaces that respond to the user's brain activity in real time. Nine subjects played two levels of the game Pacman while their brain activity was measured using fNIRS. Statistical analysis and machine learning classification results show that we can discriminate well between subjects playing or resting, and distinguish between the two levels of difficulty with some success. In contrast to most previous fNIRS studies which only distinguish brain activity from rest, we attempt to tell apart two levels of brain activity, and our results show potential for using fNIRS in an adaptive game or user interface.

Keywords: Brain-computer interface, human cognition, functional near-infrared spectroscopy, fNIRS, task classification, game, difficulty level.

1 Introduction

A brain-computer interface (BCI) can be loosely defined as an interface controlled directly or indirectly by brain activity of the user. While most BCI research is designed for direct use with disabled users, we instead focus on passive BCIs for healthy users. Passive BCIs are interfaces that use brain measurements as an additional input, in conjunction with standard devices such as keyboards and mice [1].

Unlike much BCI work which uses electroencephalography (EEG) [2], this research uses functional near-infrared spectroscopy (fNIRS), which is non-invasive, portable, and relatively impervious to user movement (Figure 1). It is also uniquely sensitive to changes in blood oxygenation, which can be used to extrapolate levels of brain activation. This tool has been used in the contexts of biomedical research and experimental psychology, but little has been done to take advantage of it in a human-computer interaction (HCI) context. Researchers have used fNIRS to investigate brain

T. Gross et al. (Eds.): INTERACT 2009, Part I, LNCS 5726, pp. 440–452, 2009.

patterns related to particular mental activities, such as motor imagery [3, 4], mental workload [5], deception [6], or emotions [7]. However, most of these studies concentrate on differentiating between no activity and one activity, while this experiment attempts to differentiate two levels of activity from each other, as well as each level of activity from a resting baseline.

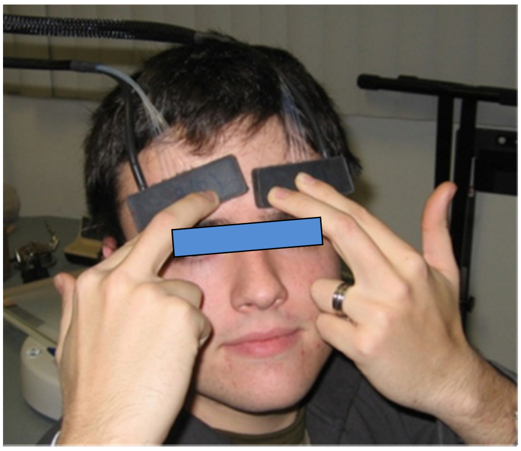

Fig. 1. A picture of a subject with the two probes (usually held by a headband)

The goal of the present study is to distinguish between different levels of game difficulty using fNIRS data collected while subjects played a computer game. The study is designed to ultimately lead to adaptive games and other interactive interfaces that respond to the user's brain activity in real time. Our results show that we can distinguish between the user playing Pacman or being at rest, as well as between two difficulty levels of Pacman.

2 Background and Related Work

2.1 Functional Near-Infrared Spectroscopy

fNIRS measures changes in hemoglobin concentrations [8]. At the near-infrared range, light can pass through most tissues, allowing them to be probed for depths up to 1-3 cm. By measuring the light sent at two wavelengths, we can calculate oxygenated and deoxygenated hemoglobin concentration. The slow hemodynamic changes measured by fNIRS occur in a time span of 6-8 sec [9]. fNIRS provides high temporal resolution (data points measured in the order of tenths of ms), and a spatial resolution of approximately 5mm. However, it can only measure the cortical surface of the brain. In comparison, fMRI has a low temporal resolution but allows whole-brain imaging, including both cortical and subcortical structures. EEG can gather information from electrodes placed all over the scalp, with a high temporal resolution. While there are many brain imaging techniques, each with advantages and disadvantages [2], we believe fNIRS to be a suitable brain sensing technology for HCI research because it is safe, non-invasive, easy to use, and relatively impervious to user movement, as compared to other brain techniques.

2.2 Psychophysiological Related Work

Game play has been measured using psychophysiological signals. For instance, Chen et. al used two physiological measures (heart rate variability and electromyogram) to measure the interruptibility of subjects in different tasks, including a game, and found a high correlation between those measures and the self-report of interruptibility [10]. Other researchers have measured the brain during game play using EEG and demonstrated the ability to distinguish the user resting, exploring the game environment or playing the video game [2]. Based on these results, we wanted to explore the fNIRS blood oxygenation response during different levels of video game play.

Task load and blood oxygenation have been shown to be correlated in a number of non-game environments [11] as well as in more directly relevant game-playing environments. Several fNIRS studies reported a significant variation in hemoglobin concentration in the prefrontal cortex in comparison to resting while playing an arcade game [12], a shooting game, a rhythm action game, a block puzzle and a dice puzzle [13]. Another study showed that one could differentiate between playing and not playing a computer game using functional magnetic resonance imagery (fMRI), by comparing three video games: Space Invaders, Othello and Tetris [14]. These studies all compare rest versus play, but never more than one level of difficulty.

These research papers show a prefrontal cortex response to video game playing, which lead us to believe that the video game Pacman could produce similar activations. However, note that most of the fNIRS studies measure a larger brain region, with probes that are much different than ours, although our current probe format has the advantage of a simple and comfortable setup. The present study applies fNIRS to the human forehead, measuring the anterior prefrontal cortex, a subset of the prefrontal cortex. The choice of Pacman was motivated by the fact that Pacman offers different difficulty levels that keep all other aspects identical, such as the scene and the characters' behavior. It was also desired to study an untested arcade video game with fNIRS, which we believe can be translated to other games of similar mental demand.

3 Experimental Protocol

The goal of this study was to measure brain activity using fNIRS during game play, and to differentiate the brain signal between different levels of a computer game. The arcade game Pacman was selected because of its customizable environment. We implemented a homemade computer version of the game, originally released by Namco (Japan). The user directs Pacman through a maze by pressing arrow keys, with the goal of eating as many fruits and enemies as possible, without being killed. Two levels of difficulty, differentiated by pace and quantity of enemies, were selected through pilot testing.

Participants were hypothesized to be able to distinguish these difficulty levels, so it was also hypothesized that brain measurements would show distinguishable differences in addition to observed differences in performance.

Nine subjects (4 females) participated in this study (mean age of 24.2 years; std 4.15). All were right-handed, with normal or corrected vision and no history of major head injury. Informed consent was obtained, and participants were compensated for their time. All knew of the game, and all but one had previously played it. Participants practiced the game for about one minute to familiarize themselves with our version.

3.1 Design and Procedure

Participants completed ten sets of two trials (one in each difficulty level) over a twenty minute period. In each trial, participants played the game for a period of thirty seconds, and rested for thirty seconds to allow their brain to return to baseline. Conditions within each set were randomized for each subject. The experimental protocol of alternating 30s-long windows of activation and rest was designed to take into account the slow hemodynamic changes that occur in a time span of 6-8 sec [9] as well as a short game cycle that nonetheless allowed performance to level off.

In addition to fNIRS data, we collected performance data—number of times Pacman is killed, as well as number of fruits and enemies eaten. At the end of the experiment, subjects were asked to rate the overall mental workload of each game level with the NASA Task Load Index (NASA-TLX) [15], a widely used measure of subjective mental workload used here as a manipulation check. NASA-TLX provides a ground truth measurement, a benchmark for comparing and validating fNIRS results. It is a collection of questions relating to the task's mental, physical, and temporal demands on the user, their performance, effort and frustration level when executing the task. The NASA-TLX for each level was administered using a paper version (two in total).

3.2 fNIRS Equipment

We collected fNIRS data using an OxiplexTS, from ISS, Inc. (Champaign, IL). Our setup is comprised of two probes (see Figure 2). Each source emits two wavelengths (690 and 830nm), with a sampling rate of 6.25Hz. The probes were placed in the middle of the forehead. We chose to use the data from the two last sources of each probe only (with source-detector distances of 2.5 and 3cm), because they reach deeper into the cortex. The shallower source-detector axes are thought to pick up primarily systemic responses happening in or on the skin.

Movement artifacts picked up by the fNIRS probes can include both general limb movement, and specific skin movements (e.g. frowning). The user was seated at ease, with their right hand positioned to reach the arrow keys of a standard keyboard comfortably, with the fingers resting on the keys, minimizing all movement of the arm and hand. We asked the users not to move their limbs, or to frown, but they were not constrained in any way. We did not measure their eye blinks or frowning, but we did visually observe their behavior. We did not find a visual correlation between such small movements and the preprocessed data. A pilot test indicated that small finger movements show up only minimally in our data, and this noise is mostly removed with filtering.

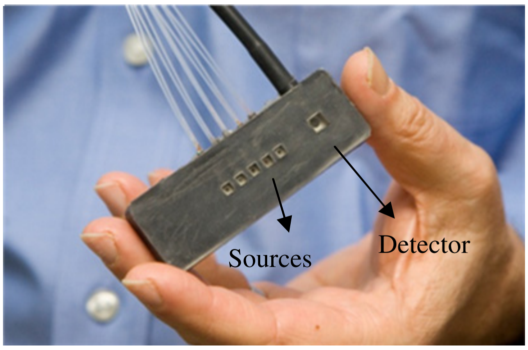

Fig. 2. A picture of the right probe. A probe includes a detector (larger square) and four light sources (smaller squares). While the probe has five possible light sources, only four sources can used at once because of hardware constraints. Moreover, we decided to only use data from two sources, the two furthest from the detector. The picture shows the side that will be on the forehead.

4 Analysis Techniques and Results

4.1 Behavioral Results and Performance Data

In this section, we performed an analysis on the non-brain data collected, such as the NASA-TLX results and the game performance statistics.

NASA-TLX. We analyzed results from the NASA-TLX data to confirm that users perceived the two difficulty levels as different. Results indicated an average mental workload index of 26 (std 12.9) for the easy level, and 69 (std 7.9) for the hard level, on a 100 point scale. This difference was significant according to a two sided t-test ($p<0.01$), and confirm our manipulation.

Performance Data. We also examined the performance data. Every data source collected showed a significant difference between the two difficulty levels ($p<0.05$). Figure 3 displays the average value of the data collected.

4.2 Brain Data

fNIRS is still a new methodology, and as such it lacks well-established preprocessing and analysis methods [16]. Each researcher is currently left to his or her better judgment to find a method that works best. Some researchers choose to do a visual inspection of the data to determine patterns [17], while most use some sort of statistical analysis of the data, with no real consensus on how to perform this analysis. Many perform paired t-test on averaged concentration change for each trial [18], while others average all the trials at each time point and performs t-test to compare each point with a baseline point [5, 12]. Additionally, a small number of researchers perform machine learning classification and clustering on fNIRS data [4, 5].

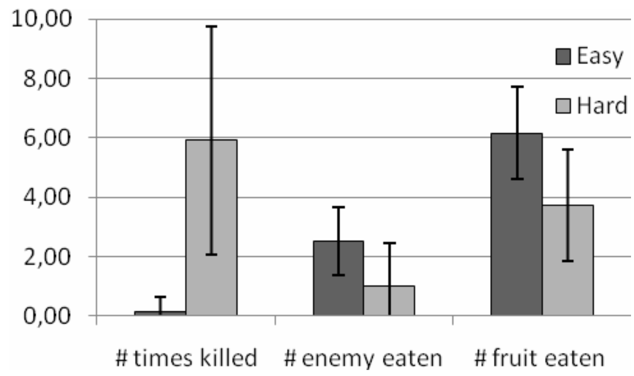

Fig. 3. Graph of data collected, with standard deviation, for each difficulty level, averaged over trials and subjects. The difference between each level is significant for each data type.

We performed two analyses of the brain data to confirm the presence of differences in hemoglobin concentrations for the different conditions: a classic statistical analysis to establish the differences between conditions, and a more novel task classification that will show the possibility of using this data in a real-time adaptive system.

Data Preprocessing. We preprocessed the raw data to remove artifacts and transform it into concentration of oxygenated and deoxygenated hemoglobin. To remove motion artifacts, and optical changes due to breathing and heart beat, we applied a folding average filter using a non-recursive time-domain band pass filter, keeping frequencies between 0.01Hz and 0.5Hz. The filtered raw data was then transformed into oxygenated hemoglobin and deoxygenated hemoglobin concentrations (respectively [HbO] and [Hb]), using the modified Beer-Lambert law [8].

Given the assumption that the brain returns to a baseline state during each rest period following the stimuli, even though it may not be the same baseline state in each rest period, we shift each trial so that the initial value is zero to control for differences in initial state. Finally, we separate each trial according to *Activeness*—whether the user was playing or resting. Figure 4 illustrates trials of data for a particular stimulus.

Statistical Analysis. For the statistical analysis, we average each trial of each condition to get a mean value of [HbO] and [Hb], for each difficulty level, activeness, hemisphere and channel. We then apply a factorial repeated measures analysis of variance (ANOVA) on *Difficulty level* (2) x *Activeness* (2) x *Hemoglobin Type* (2) x *Hemisphere* (2) x *Channel* (2) x *Subject* (9). This factorial ANOVA will observe differences within each participant, and determine if they are significant across participants. If the end result is to construct a system that can respond to different individuals with a minimum of training, we need to know how different we should expect individuals to be—hence including subjects as a factor in the analysis. Given the novelty of the fNIRS method, and the lack of well established analysis methods in previous work in this area, the cortical distribution of the(combination of channel and hemoglobin type effects cannot yet be predicted beforehand. In addition to the statistical significance, we report the effect size of the interaction (ω^2), which is the magnitude of the observed interaction,

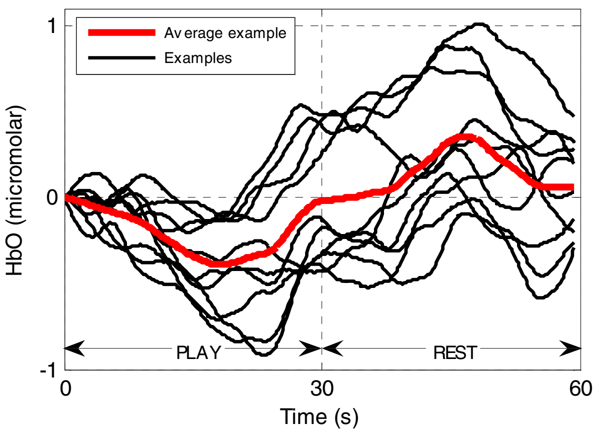

Fig. 4. Example of fNIRS data. The data displayed is subject 2's [HbO], from source 3 of the right probe, filtered. The red, ticker line indicates the mean of all trials. The left half of the data was taken when the user was playing the easy Pacman, and the right half was the rest period following.

and indicates practical significance. An omega-squared measure of 0.1 indicates a small effect, 0.3 a medium effect and 0.5 a large effect [19].

We found that the *Hemoglobin Type* was a significant factor, with a medium effect ($F(1, 8)=6.819$, $p<0.05$, $\omega^2=0.39$). This was expected, because [Hb] and [HbO] are present in different concentrations in the blood. The interaction of *Channel* x *Hemoglobin Type* is also significant, with a medium effect ($F(1, 8)=5.468$, $p<0.05$, $\omega^2= 0.33$), indicating that [Hb] and [HbO] are not the same at a given channel.

Game-playing compared to resting are significantly different as an interaction with channel with a large effect size (*Activeness* x *Channel*, $F(1, 8)=27.767$, $p<0.001$, $\omega^2= 0.75$), showing that there is a difference between playing Pacman and resting, and that this difference varies as a function of the cortical depth of the measurement (that is, the source-detector distance). We also observed that the interaction of *Activeness* x *Channel* x *Hemoglobin Type* is significant, with a medium effect ($F(1, 8)=5.412$, $p<0.05$, $\omega^2= 0.32$).

Finally, we observed a significant interaction of *Difficulty Level* x *Activeness* x *Channel* x *Hemoglobin Type*, with a small effect size ($F(1, 8)= 7.645$, $p<0.05$, $\omega^2= 0.18$). This interaction shows that we can significantly distinguish between the activeness of the participant, and the degree of difficulty of the current game when we take into account the channel and the hemoglobin type. This confirms our initial hypothesis.

Machine Learning Classification. Statistical analysis confirmed our hypothesis that the brain signals in the different conditions were significantly different. We then wanted to determine whether this signal could be used in an adaptive user interface. To do this, we used machine learning to train a classifier.

We chose sequence classification [20] because of its simple nature. Sequence classification applies a label to an entire sequence of data, and uses each data point as a feature. In our case, a sequence is one trial, containing 180 points. We used the same preprocessing as for the statistical analysis, but we use non-zeroed data, as it is more similar to data we would have in a real time brain-computer interface.

Because we have multivariate data (8 recordings for each time point: 2 probes x 2 channels x 2 hemoglobin types), we classify each channel individually first. To combine the results of all these classifications, each classifier votes for the label of the example. We used a weighted voting technique that takes into account the probability distribution of each example by each classifier.

The classification algorithm used is k-nearest-neighbors (kNN), with k=3. kNN uses the label of the three most similar examples (the closest neighbors) to the example to classify, and assigns a label based on the weighted average of their labels. We used a random 10-fold cross-validation in all classifications. We trained the classifier on part of one subject's data, and then tested for this specific subject with the left out data. This procedure was repeated for each subject.

We attempted three types of classification: (a) *Activeness* (Play versus Rest), (b) *Difficulty level* (Easy versus Hard), and (c) *Two difficulty levels and rest* (Easy versus Hard versus Rest). To accomplish each classification, we selected and/or grouped the trials differently. For *Activeness*, we combined all playing trials into one class, and all resting trials into another to form two classes. For *Difficulty Level*, we compared the easy and hard levels using the play trials only. Finally, in *Two difficulty levels and rest*, we compared three conditions: the play period of the easy level, the play period of the hard level, and all rest periods. Figure 5 shows the average accuracy of each type of classification (accuracy averaged over subjects).

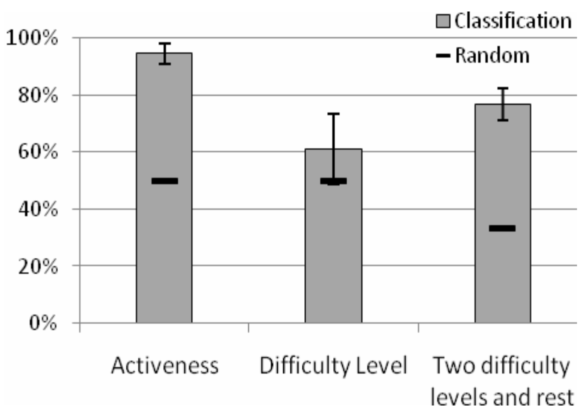

Fig. 5. Average accuracy for different classifications, with the standard variation and the random classification accuracy. *Activeness* compares the playing trials to the resting trials; *Difficulty Level* compares the easy and hard levels using the play trials only; *Two difficulty levels and rest* compares the easy playing trials versus the hard playing trials versus the resting trials.

5 Discussion

Our analyses show that we can distinguish between subjects being active and passive in their mental state (*Activeness*), as well as between different levels of game complexity (*Difficulty Level*) in this particular task when combined with the activeness of the participant, the channel and hemoglobin type measured. The classic statistical analysis confirmed that these conditions produced different patterns in blood oxygenation level, and the machine-learning analysis confirms that these patterns can be distinguished by the classifiers used.

While some might argue that performance data is sufficient to classify the difficulty level of a game and can be obtained without interference, the goal of this study is to investigate the use of the brain measurements with fNIRS as a new input device. In a more complex problem, performance and brain data coming from fNIRS might not be as related, e.g. if the user is working hard yet performing poorly at some point. In addition, distractions may also produce workload increases that would not obvious from monitoring game settings and performance, and thus may necessitate brain measurements. That is, a participant playing a simple game while answering difficult questions might also show brain activity relating to increased workload that would be incomprehensible based only on performance data (e.g. [21]). In real, non gaming situations, we might not have performance data like in the present case, as we don't always know what to measure— how hard is an air traffic controller working, or a person creating a budget on a spreadsheet? The use of the brain signal as an auxiliary input could provide better results in these situations.

5.1 Brain Activation When Playing Pacman: Play versus Rest

Results indicate the presence of a distinct brain signal when playing Pacman, in comparison to the rest periods. The *Activeness* classification in Figure 5 yields an average accuracy of 94.4%. It indicates a noticeable difference between the playing signal, and the resting signal. This corresponds to the results obtained with the statistical analysis, where *Activeness* was a significant factor in multiple interactions. This provides real time measurements that could be used in an adaptive interface. Our results corroborate those of previous studies that showed prefrontal cortex activity related to video games, measured with fNIRS.

5.2 Difficulty Levels: Easy versus Hard

The *Difficulty level* of the game was shown to be a significant factor in this experiment in both types of analyses. This is supported with the fact that users perceived the two levels as being significantly different according to the NASA-TLX. Hence, we can say that there was a significant cognitive difference between the two levels. Previous fNIRS game experiments [12, 13] only analyzed stimuli versus non-stimuli periods (which in this experiment we have called activeness), and not two levels of difficulty, making this result an advance over prior work.

However, the statistically significant interaction that included *Difficulty Level* had a small effect size, and classifying the difficulty of playing periods yields an average accuracy of 61.1%. This relatively low accuracy indicates that it is difficult with this classifier to differentiate between the two levels, which relate to the small effect size found in the statistical analysis. We also observed significant inter-subject variability: four participants scored between 65% and 85%. This indicates that the two difficulty levels might be significantly different with only part of the participants. As everyone's brain varies greatly, this is not a surprising result.

A comparison of three types of conditions (*Two difficulty levels and rest*) indicates an encouraging average accuracy of 76.7%, explained by the low differentiation between the difficulty levels, and the high separation between the activeness of the subjects. We must note that the difference in brain signal measure is not strong. One explanation may be that the difference in mental processes between each level manifests itself in other brain locations besides the anterior prefrontal cortex (location measured), such as in the dorsolateral prefrontal cortex. It could also be that the difference between the two difficulty levels was not big enough to cause strong changes in activation.

Results are consistent with prior work. Distinguishing work from rest was relatively easy, but discriminating different workload levels was harder, with significant inter-subject variability. Similar results have been found over decades of EEG work (e.g. [22, 23]), which may suggest fundamental limitations in making fine discriminations between two similar workload levels.

5.3 Subject Movement

We noted earlier that subjects' motions can sometimes be picked up by fNIRS devices. We believe that by simply asking the subjects to restrain their movement (major limb movements, as well as yawning and frowning), and by applying a filtering algorithm, we can minimize these motion artifacts. The data showed in this paper corroborates this hypothesis. The experiment was located in a quiet work environment, our subjects did use the keyboard, and significant differences between conditions were still obtained. This is good news for the use of fNIRS in HCI, as it shows the feasibility of using such tool in a real setting. We hypothesize that the use of the mouse would also be acceptable because those movements are usually minimal.

Overall, the findings indicate the presence of brain activation in the anterior prefrontal cortex when playing Pacman. Because the activation of the different levels of difficulty is correlated with mental workload (measured with NASA-TLX), we can presume that the difficulty level in this experiment is also correlated with mental workload.

The machine learning results show that fNIRS data can be classified easily, suggesting great potential as an auxiliary input for an adaptive interface. In the long run, our goal is to be able to classify data in real time.

6 Future Work

There is much interesting work to be done with fNIRS that could benefit the HCI community. Next steps include converting an offline classifier into a real-time algorithm that accepts streaming data for use in an adaptive interactive user interface. Additional data analysis could further resolve the temporal dynamics of classification efficacy, such as detecting workload changes within the first 2, 5, or 10 seconds instead of 30. Furthermore, a probe with more sensors, placed differently, could lead to a stronger signal, as it would pick up changes in blood oxygenation in more locations.

Saito et al observed a larger activation cluster in the dorsolateral prefrontal cortex with the games of Othello and Tetris than with Space Invaders [14]. This was justified with the fact that Othello and Tetris require spatial logical thinking (planning and memory of prior moves). The game of Pacman relates more to Space Invaders than to Othello or Tetris, as both are arcade games, and not puzzles, suggesting the possibility of a stronger signal with a different game. In addition, previous work using fNIRS to study video games compare different types of games (e.g. shooter game versus puzzle game), which could be interesting to experiment with, such as contrasting different levels in other types of games. This could verify whether differentiating two levels of video games yield weak results in other game types, or that Pacman's main brain activation is located elsewhere.

In a larger research context, exploring the use of fNIRS in an adaptive interface would prove interesting for the HCI community. Pacman was chosen in this experiment because of its great potential for passive adaptability: it is easy to change the amount of enemies to maintain interest without overwhelming the user. Results of the comparison of two different levels could be applied to other games of similar mental demand. The correlation between mental workload and difficulty levels in this experiment indicates we could also apply the current results to general applications that respond to such measurements.

There are limitations to using fNIRS in real-time, such as the fact that the metabolic response measured by fNIRS occurs over a few seconds, and the difficulty of filtering out motion artifacts in real time. This suggests that a real time user interface would be hard-pressed to produce an immediate, perfect response. Using fNIRS as a passive supplemental input will avoid some of these issues since the interface would not be dependent on this signal for its interaction. The interface can be adapted in a subtle matter, when we have a high degree of certainty in the user's cognitive state. In the case of an adaptive Pacman, changing the difficulty level should not be clearly noticeable to the user.

7 Conclusion

In this experiment, we have shown that functional near-infrared spectroscopy can distinguish between the brain at rest and the brain activated when playing a video game, both using statistical analysis and machine learning classification. We also demonstrated that we can differentiate two levels of difficulty. The activation of the different levels of difficulty is correlated with mental workload, measured with NASA-TLX. Hence, we can presume that the difficulty level in this experiment is correlated with mental workload. However, our classification accuracy was low when distinguishing easy or hard levels.

We introduced fNIRS as a new input device to the HCI community. It shows potential by its ability to measure different brain signals, such as difficulty level and mental workload, and its ease of use, and quick setup time. This is a step forward, as previous work only studied the activeness of the user during video games using fNIRS. We believe this work to be a stepping stone to using fNIRS in an adaptive user interface, in this case a passive brain-computer interface. In a real time user interface, we could use fNIRS measurement as an additional input on which to adapt the interface. In the case of Pacman, it could be used to modify the game's difficulty level to keep the user in an ideal game level, always challenged without being overwhelmed.

Acknowledgements. The authors would like to thank our colleagues and alumni in the HCI research group at Tufts University; Michel Beaudoin-Lafon, Wendy Mackay, and the In|Situ| research group; and Desney Tan at Microsoft Research for their helpful inputs and encouragement. We thank the National Science Foundation (Grant Nos. IIS-0713506 and IIS-0414389), the US Air Force Research Laboratory, the Natural Sciences and Engineering Research Council of Canada, the US Army Natick Soldier Research, Development and Engineering Center for support of this research. Any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of these organizations.

References

1. Cutrell, E., Tan, D.S.: BCI for passive input in HCI. In: Proc. ACM CHI 2008 Conference on Human Factors in Computing Systems Workshop on Brain-Computer Interfaces for HCI and Games (2007)
2. Lee, J.C., Tan, D.S.: Using a Low-Cost Electroencephalograph for Task Classification in HCI Research. In: Proc. UIST 2006 (2006)
3. Coyle, S., Ward, T., Markham, C., McDarby, G.: On the Suitability of Near-Infrared Systems for Next Generation Brain Computer Interfaces. In: Proc. World Congress on Medical Physics and Biomedical Engineering, Sydney, Australia (2003)
4. Sitaram, R., Zhang, H., Guan, C., Thulasidas, M., Hoshi, Y., Ishikawa, A., Shimizu, K., Burbaumer, N.: Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain-computer interface. NeuroImage 34, 1416–1427 (2007)
5. Hirshfield, L.M., Solovey, E.T., Girouard, A., Kebinger, J., Sassaroli, A., Tong, Y., Fantini, S., Jacob, R.J.K.: Brain Measurement for Usability Testing and Adaptive Interfaces: An Example of Uncovering Syntactic Workload with Functional Near Infrared Spectroscopy. In: Proceedings of CHI 2009 Conference on Human Factors in Computing Systems (2009)
6. Bunce, S., Devaraj, A., Izzetoglu, M., Onaral, B., Pourrezaei, K.: Detecting deception in the brain: a functional near-infrared spectroscopy study of neural correlates of intentional deception. In: Proceedings of the SPIE International Society for Optical Engineering, vol. 5769, pp. 24–32 (2005)
7. Herrmann, M.J., Ehlis, A.C., Fallgatter, A.J.: Prefrontal activation through task requirements of emotional induction measured with NIRS. Biological Psychology 64(3), 255 (2003)
8. Villringer, A., Chance, B.: Non-Invasive Optical Spectroscopy and Imaging of Human Brain Function. Trends in Neuroscience 20, 435–442 (1997)
9. Bunce, S.C., Izzetoglu, M., Izzetoglu, K., Onaral, B., Pourrezaei, K.: Functional Near Infrared Spectroscopy: An Emerging Neuroimaging Modality. IEEE Engineering in Medicine and Biology Magazine, Special issue on Clinical Neuroengineering 25(4), 54–62 (2006)
10. Chen, D., Hart, J., Vertegaal, R.: Towards a physiological model of user interruptability. In: Baranauskas, C., Palanque, P., Abascal, J., Barbosa, S.D.J. (eds.) INTERACT 2007. LNCS, vol. 4663, pp. 439–451. Springer, Heidelberg (2007)
11. Izzetoglu, K., Bunce, S., Onaral, B., Pourrezaei, K., Chance, B.: Functional Optical Brain Imaging Using Near-Infrared During Cognitive Tasks. International Journal of Human-Computer Interaction 17(2), 211–231 (2004)

12. Nagamitsu, S., Nagano, M., Tamashita, Y., Takashima, S., Matsuishi, T.: Prefrontal cerebral blood volume patterns while playing video games-A near-infrared spectroscopy study. Brain & Development 28, 315–321 (2006)
13. Matsuda, G., Hiraki, K.: Sustained decrease in oxygenated hemoglobin during video games in the dorsal prefrontal cortex: A NIRS study of children. NeuroImage 29(3), 706–711 (2006)
14. Saito, K., Mukawa, N., Saito, M.: Brain Activity Comparison of Different-Genre Video Game Players. In: Proceedings of ICICIC '07 International Conference on Innovative Computing Information and Control, pp. 402–406 (2007)
15. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): Results of empirical and theorical research. In: Hancock, P., Meshkati, N. (eds.) Human Mental Workload, Amsterdam, pp. 139–183 (1988)
16. Butti, M., Caffini, M., Merzagora, A.C., Bianchi, A.M., Baselli, G., Onaral, B., Secchi, P., Cerutti, S.: Non-invasive neuroimaging: Generalized Linear Models for interpreting functional Near Infrared Spectroscopy signals. In: Caffini, M. (ed.) Proceedings of CNE '07 Conference on Neural Engineering, pp. 461–464 (2007)
17. Nishimura, E., Stautzenberger, J.P., Robinson, W., Downs, T.H., Downs, J.H.: A new approach to functional near-infrared technology. IEEE Engineering in Medicine and Biology Magazine 26(4), 25–29 (2007)
18. Matsuda, G., Hiraki, K.: Prefrontal Cortex Deactivation During Video Game Play. In: Shiratori, R., Arai, K., Kato, F. (eds.) Gaming, Simulations, and Society: Research Scope and Perspective, pp. 101–109 (2005)
19. Field, A.P., Hole, G.: How to design and report experiments. Sage publications Ltd., Thousand Oaks (2003)
20. Dietterich, T.G.: Machine learning for sequential data: A review. In: Caelli, T.M., Amin, A., Duin, R.P.W., Kamel, M.S., de Ridder, D. (eds.) SPR 2002 and SSPR 2002. LNCS, vol. 2396, pp. 15–30. Springer, Heidelberg (2002)
21. Nijholt, A., Tan, D., Allison, B., Milan, J.d.R., Graimann, B.: Brain-computer interfaces for hci and games. In: CHI 2008 extended abstracts on Human factors in computing systems. ACM, Florence (2008)
22. Gevins, A., Smith, M.: Neurophysiological measures of cognitive workload during human-computer interaction. Theoretical Issues in Ergonomics Science 4, 113–131 (2003)
23. Allison, B.Z., Polich, J.: Workload assessment of computer gaming using a single-stimulus event-related potential paradigm. Biological Psychology 77(3), 277–283 (2008)