

## DISCRIMINATION OF MENTAL WORKLOAD LEVELS IN HUMAN SUBJECTS WITH FUNCTIONAL NEAR-INFRARED SPECTROSCOPY

ANGELO SASSAROLI<sup>\*,‡</sup>, FENG ZHENG<sup>\*</sup>, LEANNE M. HIRSHFIELD<sup>†</sup>,  
AUDREY GIROUARD<sup>†</sup>, ERIN TREACY SOLOVEY<sup>†</sup>,  
ROBERT J. K. JACOB<sup>†</sup> and SERGIO FANTINI<sup>\*</sup>

*\*Department of Biomedical Engineering  
Tufts University, Medford, MA 02155, USA*

*†Department of Computer Science  
Tufts University, Medford, MA 02155, USA*

*‡angelo.sassaroli@tufts.edu*

We have applied functional near-infrared spectroscopy (fNIRS) to the human forehead to distinguish different levels of mental workload on the basis of hemodynamic changes occurring in the prefrontal cortex. We report data on 3 subjects from a protocol involving 3 mental workload levels based on working memory tasks. To quantify the potential of fNIRS for mental workload discrimination, we have applied a 3-nearest neighbor classification algorithm based on the amplitude of oxyhemoglobin (HbO<sub>2</sub>) and deoxyhemoglobin (HbR) concentration changes associated with the working memory tasks. We have found classification success rates in the range of 44%–72%, which are significantly higher than the corresponding chance level (for random data) of 19.1%. This work shows the potential of fNIRS for mental workload classification, especially when more parameters (rather than just the amplitude of concentration changes used here) and more sophisticated classification algorithms (rather than the simple 3-nearest neighbor algorithm used here) are considered and optimized for this application.

*Keywords:* Diffuse optical imaging; near-infrared spectroscopy; functional brain imaging; working memory; mental workload.

### 1. Introduction

Functional near infrared spectroscopy (fNIRS) is a non-invasive optical technique that can be developed into a portable or wearable monitor of blood flow and tissue oxygenation. It is sensitive to both oxyhemoglobin (HbO<sub>2</sub>) and deoxyhemoglobin (HbR) species with high temporal resolution (a few milliseconds) and a spatial resolution of  $\sim 5$  mm. The main limitation of fNIRS for brain studies in adult human subjects is the lack of sensitivity to brain structures deeper than the cerebral cortex. However, most of the higher-order brain functions take place in the cortex, and therefore, it is an important target for research studies. Functional near infrared spectroscopy has been used in a number of brain studies, from motor area activation,<sup>1–4</sup> to visual<sup>3,5,6</sup> and cognitive<sup>7–12</sup> stimulations.

In this work, we show the potential of fNIRS to detect and discriminate different levels of mental workload that involve the use of working memory. We have collected continuous wave (CW) optical data on the forehead of three subjects, and after translating the CW signals into HbO<sub>2</sub> and HbR concentration changes, we have applied a 3-nearest neighbor classification algorithm based on the maximum amplitude of concentration changes with respect to the initial, rest value. The reasoning behind this choice of classification algorithm is that changes in the cerebral blood volume and blood flow, which are reflected in the changes of HbO<sub>2</sub> and HbR concentrations in brain tissue, are expected to scale according to different levels of mental workloads. Our results show that even this simple classification algorithm, based on one single parameter, can lead to the discrimination of different levels of mental workload. These promising results are prompting further research aimed at developing a system for reliable and real-time assessment of mental workload. One specific application that we envision is a new class of dynamical human-computer interfaces that make use of continuous feedback based on the user's mental workload.

## 2. Materials and Methods

### 2.1. *Experimental setup*

The optical instrument was a near-infrared optical spectrometer from ISS, Inc., Champaign, IL (OxiplexTS) comprising two detector channels (A and B) and a number of laser diodes (in this study, we have used a total of 8 laser diodes, 4 emitting at a wavelength of 690 nm, and 4 at 830 nm), all coupled to optical fibers. The source and detector fibers were arranged over two planar and somewhat flexible optical probes to adapt to the curved surface of the subjects' forehead. Each probe received one detector fiber (from channel A for the right forehead side, and from channel B for the left side) and two pairs of individual source fibers (each pair consisting of one fiber for 690 nm and one fiber for 830 nm). Three prisms in each probe were used to reflect the light from the illumination fibers to the forehead and from the forehead to the collection fiber. The illumination and collection fibers were arranged along one line (linear array), and the source-detector distances used in this study were 1.5 cm and 3 cm. The optical probes and the optical fibers connecting them to the light sources and optical detectors are schematically shown in Fig. 1. The optical probes were firmly held on the subject's forehead by means of a black elastic band. The acquisition rate of the optical system was set to 6.25 Hz, so that we collected one data point every 160 milliseconds.

### 2.2. *Experimental protocol and data analysis*

Five healthy subjects completed 24 tasks where they viewed 5 sides of a rotating cube on a computer screen. The side faces of the cube consisted of sections of different colors, as shown in Fig. 1. In the experiment, the cube presented to the subject had either 2, 3, or 4 sections of different colors, and the number of colored

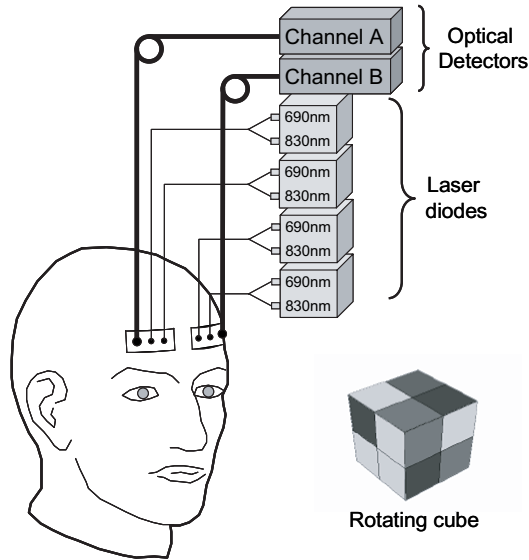


Fig. 1. Experimental setup showing the optical probes placed on the subject's forehead and the fiber optics connections to the 2 optical detectors (channel A: right side; channel B: left side) and the 8 laser diodes (4 at 690 nm and 4 at 830 nm). The rotating cube, displayed on a computer screen in front of the subject, is also shown.

sections defined the level of mental workload. During each task, the subjects were asked to count the total number of sections of each color in the 5 cube sides and their final answer was recorded for verification. The display of the different sides of the cube was performed as follows. One lateral face of the cube was presented to the subjects for 8 seconds; then the cube underwent a  $90^\circ$  rotation (which lasted 3 seconds) and stopped when the next lateral face was facing the subject. The process was repeated 3 times for a total rotation angle of  $270^\circ$  during which 4 lateral faces of the cube were shown to the subjects. Four extra seconds were allowed at the end of the last rotation for the examinations of the colored sections on the top face as well. The top face was always visible during the rotations while the bottom face was hidden. The period of time during which 5 cube sides were presented to the subjects, which lasted 45 seconds, defined a workload period. Each workload was followed by 40 seconds of rest, for a total task/rest period of 85 seconds. Four different workload levels (0, 2, 3, 4), as defined by the number of colors on the cube sides, were presented randomly 6 times to the subjects, for a total measurement duration of about 35 minutes (including 1 minute of initial baseline). In workload level 0, no cube was displayed on the screen and the subjects were asked to clear their minds and not to engage in any particular mental task. As the number of colors increased, it was necessary for the subjects to keep more information in working memory in order to remember the number of sections of each color for the cube sides that they had already viewed.

The CW intensities of the optical data were processed by calculating an average rest value, defined as  $I_0$ , in order to apply the modified Beer–Lambert law<sup>13</sup> (MBBL) to translate intensity changes into absorption changes. The changes of oxyhemoglobin and deoxyhemoglobin concentrations,  $\Delta[\text{HbO}_2]$  and  $\Delta[\text{HbR}]$ , respectively, were calculated from the data at a 3-cm source-detector separation by assuming differential path-length factors (DPF) of 6.5 and 5.8, at the wavelengths of 690 nm and 830 nm, respectively.<sup>14</sup> Afterwards, the temporal trends of  $\Delta[\text{HbO}_2]$  and  $\Delta[\text{HbR}]$  were de-trended and filtered with a moving average band-pass filter in the frequency range of 0.01–0.10 Hz, and, finally, a folding average procedure was applied to the 6 repetitions of each workload. We studied the statistical significance of the points in the folding average by using a modified  $t$  test, namely the Dubey/Armitage–Parmar algorithm, at the level of significance  $\alpha = 5\%$ . This algorithm offers a heuristic method to control the rate of statistical errors of type I (false positives) when the data points are correlated.<sup>15,16</sup>

The classification algorithm is based on the maximal amplitude of the  $\text{HbO}_2$  or  $\text{HbR}$  concentration changes recorded during each 45-second workload period and the initial 20 seconds of the following rest period (to account for a delay in the recovery to baseline). We have performed the classification analysis on workloads 0, 2, and 4. As a pilot trial, we tried to classify only 3 workloads which in principle, should elicit rather different brain activations. From the filtered intensity data, we picked one trial of one workload (query data) and determined its maximum change of  $\text{HbO}_2$  or  $\text{HbR}$ . Then, we compared the maximum amplitude of this query trial with that of the remaining 17 trials. Here, we considered a 3-nearest neighbor algorithm, where we considered a classification to be successful if at least 2 of the nearest neighbor data points belong to the same workload as the query data point. This procedure is repeated 18 times, for each of the 6 trials of the 3 workloads. In this way, we can quantify the classification success rate as the ratio of number of correct classifications to the total number of classifications. We observed that in the case of a 1-nearest neighbor algorithm, the success rate of classification associated with a random collection of  $k$  groups, each consisting of  $n_i$  random data points is given by  $\sum_{i=1}^k [n_i(n_i - 1)]/[N(N - 1)]$ , where  $N = \sum_{i=1}^k n_i$ . In the case of the 3-nearest neighbor algorithm that we used, the expression for the success rate from a collection of random data is  $\sum_{i=1}^k \{n_i(n_i - 1)(n_i - 2)[3(N - 1) - 2n_i]\}/[N(N - 1)(N - 2)(N - 3)]$ . By applying this latter expression, we can quantify the success rate associated with chance in our case of  $n_1 = n_2 = n_3 = 6$ , which turns out to be 19.1% (and we note that it is significantly lower than the corresponding chance level for the 1-neighbor case of 29.4%).

### 3. Results and Discussion

The optical data on 2 of the 5 subjects examined presented significant motion-related artifacts and could not be analyzed for mental workload classification. Even though fNIRS is much less sensitive to subject's motion than functional magnetic

resonance imaging (fMRI), it may nevertheless be negatively affected by motion-related contributions to the optical data. Despite our efforts at careful and stable placement of the optical probes on the subjects' foreheads, and our instructions to the subjects to limit their head movements during the protocol, 2 of the subjects found it hard to stay motionless and the optical data collected on them were affected by several motion artifacts. Therefore, we report here the data collected on 3 subjects. Subjects I and II were 26- and 22-year-old females, and subject III was a 22-year-old male. Figure 2 displays the relative changes of 690 nm intensities  $\Delta I(t)/I_0$  (where  $\Delta I(t) = I(t) - I_0$ , with  $I(t)$  intensity at time  $t$ ) at the two source-detector distances of 1.5 cm and 3 cm for channel B (left side of the forehead) of subject I. The intensity traces in Fig. 2 are obtained after applying a de-trending procedure to the intensity raw data. Figure 2 shows that the intensity changes are greater at the larger source-detector separation (3 cm), as it should be expected by the common properties of photon migration under diffusion condition. Also, the ratio of changes between the farthest and shortest distance is more than twice, as it should be approximately expected (on the basis of the MBBL) for this pair of source-detector distances if the changes of absorption were uniformly distributed. This feature, together with the fact that the intensity changes are synchronous with the mental tasks (identified by the shaded areas in Fig. 2), is consistent with hemodynamic changes occurring as a result of brain activation in the cortex. Figure 2 shows that mental workload periods (tasks 2, 3, 4) induce measurable intensity changes that are not observed during control conditions (task 0). Figures 3 to 5 show the bilateral folding-average concentration changes (i.e., the average traces over the 6 repetitions of a given task) for workloads 0, 2, and 4, measured on the 3 subjects for  $\Delta[\text{HbO}_2]$  (left panels) and  $\Delta[\text{HbR}]$  (right panels). The folding-average points that are significantly different from the initial rest value, according to the Dubey/Armitage–Parmar algorithm, are indicated by light-gray shading. With the

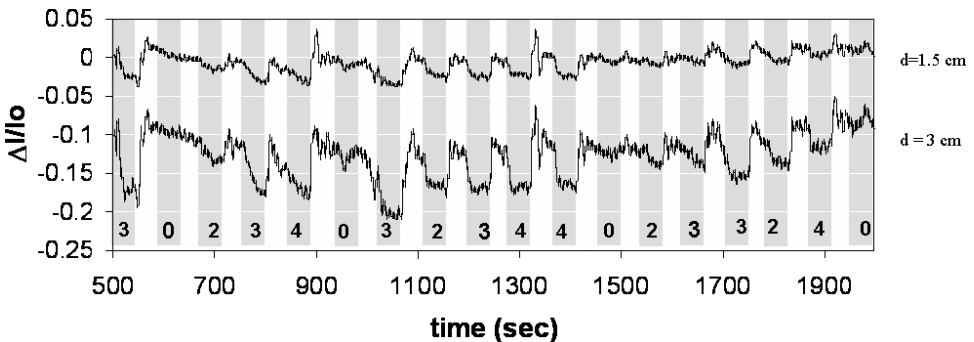


Fig. 2. Time traces of the relative changes of 690 nm optical intensity,  $\Delta I(t)/I_0$ , measured at 2 source-detector distances,  $d$ , of 1.5 cm and 3 cm on the left side of the forehead of subject I. The data at  $d = 3$  cm has been shifted down for clarity. The shaded areas represent the 45-second tasks corresponding to the workload levels (0, 2, 3, or 4) indicated on the bottom of each shaded area.

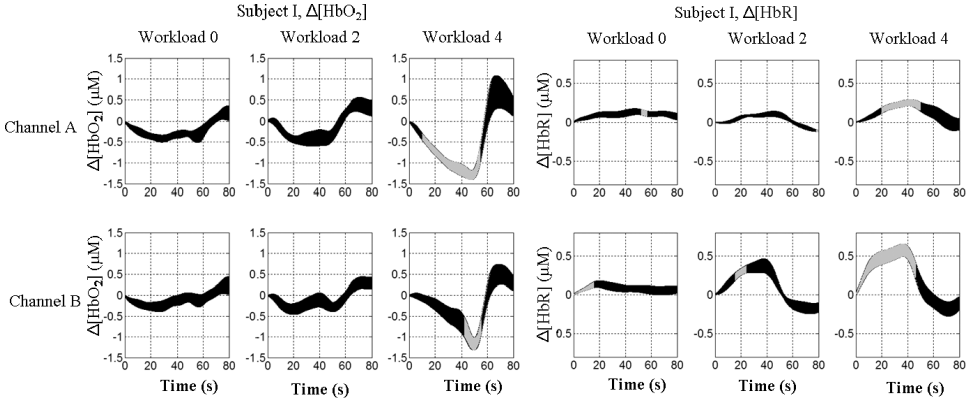


Fig. 3. Folding average of changes in oxyhemoglobin concentration ( $\Delta[\text{HbO}_2]$ , left panels) and of changes in deoxyhemoglobin concentration ( $\Delta[\text{HbR}]$ , right panels) measured for workloads 0, 2, and 4 on subject I. In each panel, the first 45 seconds correspond to the workload period, while the last 35 seconds correspond to rest periods. The concentration changes are obtained from optical data collected at a source-detector distance of 3 cm. Channels A and B refer to the right and left sides of the forehead, respectively. The points that show a statistically significant difference from the initial rest value are identified by light gray shading.

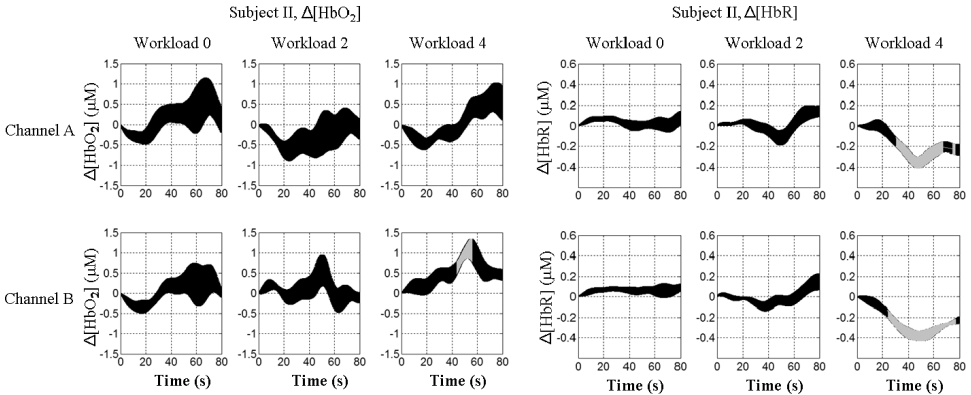


Fig. 4. Same as Fig. 3, for subject II.

only exception of  $\Delta[\text{HbO}_2]$  for channel A (right side) in subjects II and III, mental workload 4 always results in statistically significant concentration changes. By contrast, workload 2 not always results in statistically significant changes, but when it does (for example, for  $\Delta[\text{HbR}]$  in subject I-channel B, and subject III-channels A and B), it yields concentration changes that have amplitudes lower than those for workload 4 and greater than those for workload 0. For the whole dataset reported in Figs. 3 to 5, including the folding-average traces that do not contain statistical significant points, we observed that the amplitude of the concentration changes tends to increase with the level of mental workload, and we have therefore decided

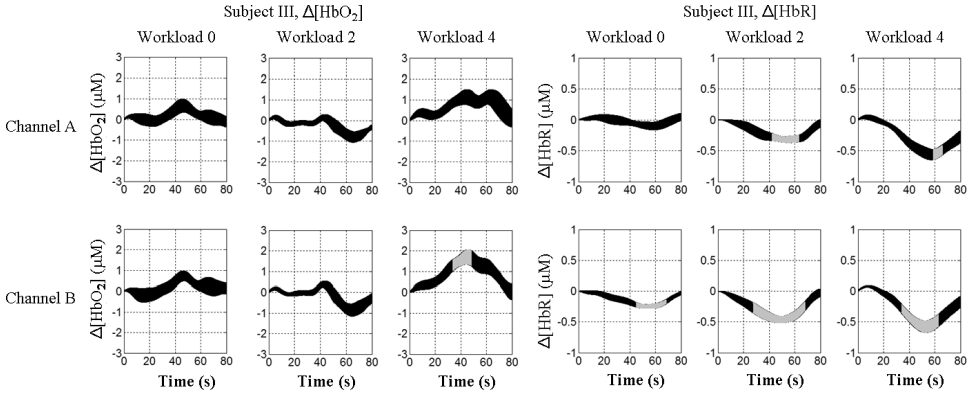


Fig. 5. Same as Fig. 3, for subject III.

to use this parameter (the amplitude of the concentration change) as a starting point in the quest for mental workload classification using fNIRS data. However, because the results from channel A (right side) consistently show either smaller amplitude signal or fewer number of significant data points (the only exception is  $\Delta[\text{HbO}_2]$  for subject I), we have considered only the results from channel B (left side) for the classification analysis. The smaller signal-to-noise ratio for channel A may be explained by a lateralized activation of the prefrontal cortex, but also by a mismatch between the location of the right optical probe and the activated area on the right side of the prefrontal cortex. This indicates the importance of performing spatial mapping, as opposed to point recordings, in functional studies of the brain with near-infrared spectroscopy.

The results for subject I show a decrease in  $\text{HbO}_2$  and an increase in  $\text{HbR}$  in response to the mental tasks, which is contrary to the well-established brain activation pattern of positive  $\Delta[\text{HbO}_2]$  and negative  $\Delta[\text{HbR}]$  that we have observed in subjects II and III. Similar inverted hemodynamic trends (of comparable amplitudes) have already been reported in the literature,<sup>17</sup> and associated with a negative blood oxygen level dependent (BOLD) signal measured in fMRI.<sup>18</sup> In the field of fMRI, some research has been devoted to the study of the possible hemodynamic and electrophysiological origins of the negative BOLD signal.<sup>19–21</sup> To test the reproducibility of our result on subject I, we have repeated the experiment on a separate day by positioning the optical probes on similar areas of the subject's forehead. In the repeated experiment, we again found the same inverted trends of  $\text{HbO}_2$  and  $\text{HbR}$  changes (results not shown).

Figures 6 and 7 report the channel B amplitudes of the changes in  $\text{HbO}_2$  and  $\text{HbR}$ , respectively, associated with all 6 trials for workloads 0, 2, and 4, for all 3 subjects. Figures 6 and 7 also show the average values (black horizontal segments) of the amplitude changes for each task. These average values demonstrate the potential of fNIRS to discriminate the 3 workloads, because, with the only exception of

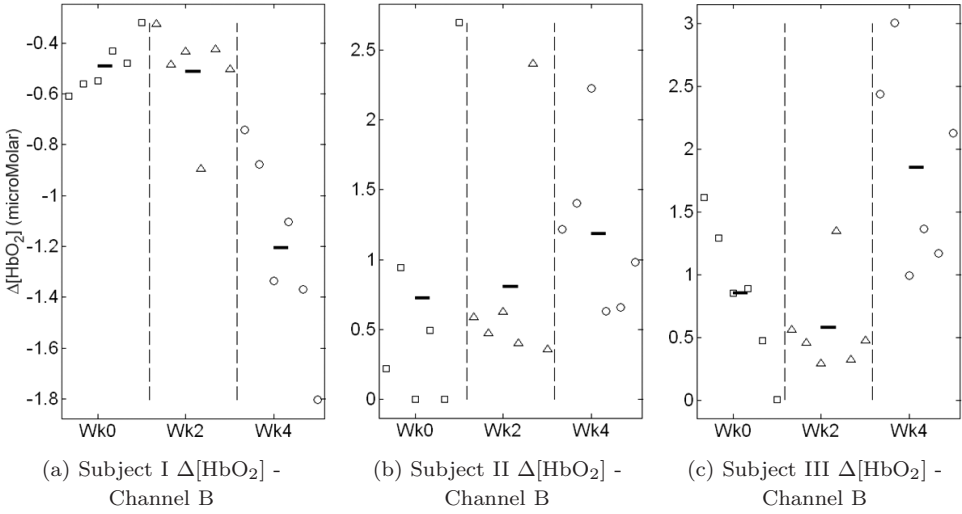


Fig. 6. Amplitude of oxyhemoglobin concentration changes ( $\Delta[\text{HbO}_2]$ ) measured in all trials (6 per each workload level) for mental workloads 0, 2, and 4. (a) Subject I. (b) Subject II. (c) Subject III. The short horizontal segments indicate the average amplitudes for each workload. The respective average values for workloads (0, 2, 4) are  $(-0.49, -0.51, -1.20)$  for subject I,  $(0.73, 0.81, 1.19)$  for subject II, and  $(0.85, 0.57, 1.85)$  for subject III.

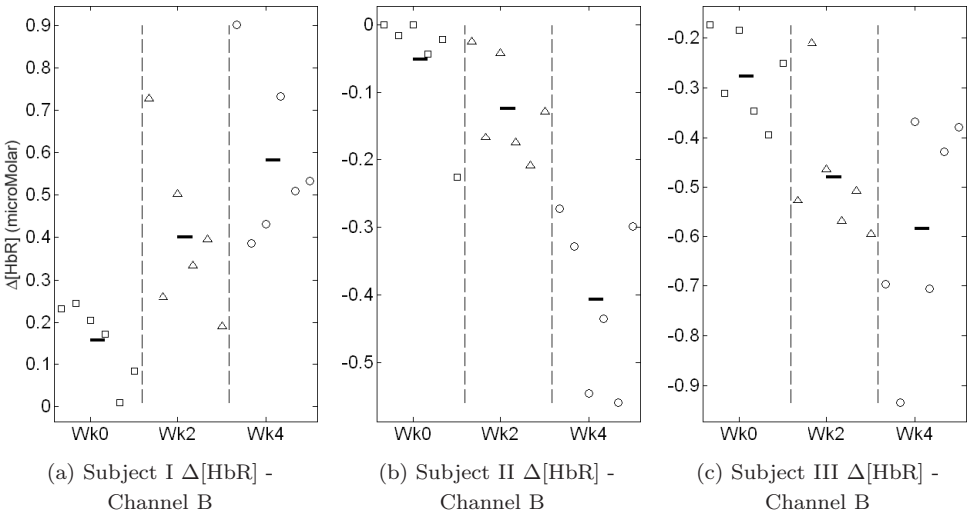


Fig. 7. Amplitude of deoxyhemoglobin concentration changes ( $\Delta[\text{HbR}]$ ) measured in all trials (6 per each workload level) for mental workloads 0, 2, and 4. (a) Subject I. (b) Subject II. (c) Subject III. The short horizontal segments indicate the average amplitudes for each workload. The respective average values for workloads (0, 2, 4) are  $(0.16, 0.40, 0.58)$  for subject I,  $(-0.05, -0.12, -0.41)$  for subject II, and  $(-0.28, -0.48, -0.59)$  for subject III.



$\Delta[\text{HbO}_2]$  for subject III, the absolute values of the average changes always increase with mental workload. While some cases (most notably  $\Delta[\text{HbR}]$  for subject II) show a limited overlap of the amplitude data for the 3 workloads, overlap of the data groups is typically the case, so that workload classification cannot be simply based on a comparison with fixed threshold values for each workload. The amplitude data points in Figs. 6 and 7 are the basis for the 3-nearest neighbor classification scheme. For each point, which correspond to one specific trial of one specific workload, we look at the 3 closest amplitude values and if at least 2 of them belong to the same workload, we consider that point to be successfully classified.

The results of the 3-nearest neighbor classification are shown in Fig. 8. For this classification algorithm and for our case of 3 groups of 6 data points each, we have derived above (in Sec. 2) the rate of successful classification for a random dataset of 19.1%. This level of chance classification, which is the reference value to evaluate the rate of successful classification of workloads based on the amplitude of  $\text{HbO}_2$  and  $\text{HbR}$  changes, is represented by horizontal dotted lines as shown in Fig. 8. The classification success rates are significantly higher than the chance level of 19.1% for all 3 subjects and for both  $\Delta[\text{HbO}_2]$  and  $\Delta[\text{HbR}]$  amplitudes, demonstrating the potential of fNIRS classification of mental workloads. The workload classification based on  $\Delta[\text{HbR}]$  (range: 55.6%–72.2%) is more accurate than that based on  $\Delta[\text{HbO}_2]$  (range: 44.4%–50.0%), and this is possibly the result of more specific detection of brain activity with  $\text{HbR}$  than  $\text{HbO}_2$  signatures.

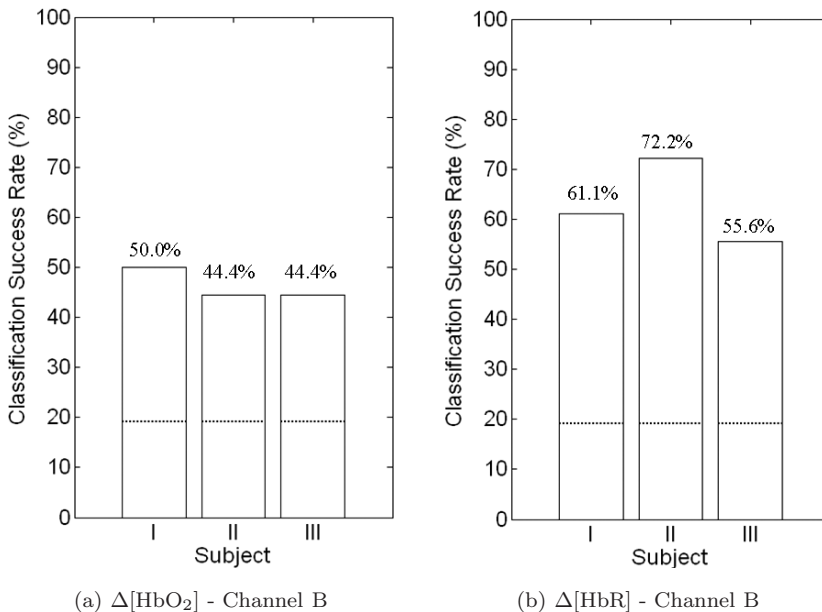


Fig. 8. Classification success rates for discrimination of workloads 0, 2 and 4, based on maximal changes of: (a) oxyhemoglobin or (b) deoxyhemoglobin concentrations measured on the left side of the prefrontal cortex. The horizontal dotted lines represent the chance level of 19.1%.

#### 4. Conclusions and Future Directions

We have demonstrated that fNIRS is sensitive to hemodynamic changes induced by a mental workload task, and that it shows potential to differentiate various levels of involvement of the working memory in human subjects. These results are highly significant because they may lead to the development of fNIRS techniques for the non-invasive, quantitative assessment of brain workload in human subjects under a number of normal working/activity environment conditions. In this study, we have used a 3-nearest neighbor classification algorithm that, despite its simplicity, resulted in successful classification at a rate significantly higher than chance. Future research will build on these results by: (a) developing optical probes that minimize the effect of subject's motion to the optical data; (b) making use of additional parameters for classification, thus taking advantage of the high temporal resolution and multi-channel capability of fNIRS data; (c) using more sophisticated machine learning algorithms for classifications,<sup>22</sup> and (d) using information from optical data collected at short source-detector distances to help discriminate signals originated at the brain cortex. We think the use of more sophisticated machine-learning algorithms is necessary if we want to refine the classification power of the optical data, as, for example, may be required to distinguish the 4 workload levels of the experiment.

One specific field that can benefit significantly from the demonstrated feasibility of non-invasive, real-time assessment of mental workloads is that of human-computer interaction (HCI). Researchers in HCI face the challenge of collecting quantitative information on human-computer users, which is very important for evaluation purposes as well as for real-time input of dynamic interfaces. Electroencephalography<sup>23</sup> (EEG) and fNIRS<sup>10</sup> have been proposed as practical tools for collecting neurophysiological measures of brain activation, and the results reported here confirm the potential of fNIRS playing a role in the development of innovative schemes of HCI.

#### Acknowledgments

This research is supported by NSF Award IIS-0713506, and NIH Grant DA021817.

#### References

1. Franceschini, M. A., Toronov, V., Filiaci, M. E., Gratton, E. and Fantini, S., *Opt. Express* **6**, 49 (2000).
2. Toronov, V., Walker, S., Gupta, R., Choi, J. H., Gratton, E., Hueber, D. and Webb, A., *Neuroimage* **19**, 1521–1531 (2003).
3. Wolf, M., Wolf, U., Toronov, V., Michalos, A., Paunescu, L. A., Choi, J. H. and Gratton, E., *Neuroimage* **16**, 704–712 (2002).
4. Boas, D. A., Gaudette, T., Strangman, G., Cheng, X., Marota, J. J. A. and Mandeville, J. B., *Neuroimage* **13**, 76 (2001).
5. Rykhlevskaia, E., Fabiani, M. and Gratton, G., *Neuroimage* **30**, 1203 (2006).

6. Csibra, G., Henty, J., Volein, A., Elwell, C., Tucker, L., Meek, J. and Johnson, M. H., *J. Pediatr. Neurol.* **2**, 85 (2004).
7. Hoshi, Y. and Tamura, M., *Neurosci. Lett.* **150**, 5 (1993).
8. Watanabe, E., Maki, A., Kawaguchi, F., Yamashita, Y., Koizumi, H. and Mayanagi, Y., *Neurosci. Lett.* **256**, 49 (1998).
9. Herrmann, M. J., Plichta, M. M., Ehlis, A. and Fallgatter, A. J., *Behav. Brain Res.* **160**, 135 (2005).
10. Izzetoglu, K., Bunce, S., Banaral, O., Pourrezaei, K. and Chance, B., *Int. J. Hum.-Comput. Inter.* **17**, 211 (2004).
11. Akgul, C. B., Akin, A. and Sankur, B., *Med. Biol. Eng. Comput.* **44**, 945 (2006).
12. Chance, B., Nioka, S., Sadi, S. and Li, C., *Adv. Exp. Med. Biol.* **510**, 397 (2003).
13. Sassaroli, A. and Fantini, S., *Phys. Med. Biol.* **49**, N255 (2004).
14. Duncan, A., Meek, J. H., Clemence, M., Elwell, C. E., Tyszczuk, L., Cope, M. and Delpy, D., *Phys. Med. Biol.* **40**, 295 (1995).
15. Sankoh, A. J., Huque, M. F. and Dubey, S. S., *Stat. Med.* **16**, 2529 (1997).
16. Sassaroli, A., Tong, Y., Benes, C. and Fantini, S., *Proc. SPIE* **6850**, 685008 (2008).
17. Hoshi, Y. and Tamura, M., *Neurosci. Lett.* **150**(1), 5 (1993).
18. Lesage, F., Gagnon, L. and Dehaes, M., *Proc. SPIE* **6850**, 68500C (2008).
19. Shmuel, A., Yacoub, E., Pfeuffer, J., Van de Moortele, P. F., Adriany, G., Hu, X. and Ugurbil, K., *Neuron* **36**, 1195 (2002).
20. Schridde, U., Khubchandani, M., Motelow, J. E., Sanganahalli, B. G., Hyder, F. and Blumenfeld, H., *Cerebral Cortex*, doi:10.1093/cercor/bhm208 (2007).
21. Smith, A. T., Williams, A. L. and Singh, K. D., *Hum. Brain Mapp.* **21**, 213 (2004).
22. Fantini, S., Sassaroli, A., Tong, Y., Hirshfield, L. M., Girouard, A., Solovey, T. and Jacob, R. J. K., *OSA Topical Meetings, BIOMED*, BMD14 (2008).
23. Smith, M. E., Gevins, A., Brown, H., Karmik, A. and Du, R., *Hum. Factors* **43**, 366 (2001).