Load-dependent brain activation assessed by time-domain functional near-infrared spectroscopy during a working memory task with graded levels of difficulty

Erika Molteni
Davide Contini
Matteo Caffini
Giuseppe Baselli
Lorenzo Spinelli
Rinaldo Cubeddu
Sergio Cerutti
Anna Maria Bianchi
Alessandro Torricelli
Load-dependent brain activation assessed by time-domain functional near-infrared spectroscopy during a working memory task with graded levels of difficulty

Erika Molteni, Davide Contini, Matteo Caffini, Giuseppe Baselli, Lorenzo Spinelli, Rinaldo Cubeddu, Sergio Cerutti, Anna Maria Bianchi, and Alessandro Torricelli

Abstract. We evaluated frontal brain activation during a mixed attentional/working memory task with graded levels of difficulty in a group of 19 healthy subjects, by means of time-domain functional near-infrared spectroscopy (fNIRS). Brain activation was assessed, and load-related oxy- and deoxy-hemoglobin changes were studied. Generalized linear model (GLM) was applied to the data to explore the metabolic processes occurring during the mental effort and, possibly, their involvement in short-term memorization. GLM was applied to the data twice: for modeling the task as a whole and for specifically investigating brain activation at each cognitive load. This twofold employment of GLM allowed (1) the extraction and isolation of different information from the same signals, obtained through the modeling of different cognitive categories (sustained attention and working memory), and (2) the evaluation of model fitness, by inspection and comparison of residuals (i.e., unmodeled part of the signal) obtained in the two different cases. Results attest to the presence of a persistent attentional-related metabolic activity, superimposed to a task-related mnemonic contribution. Some hemispherical differences have also been highlighted frontally: deoxy-hemoglobin changes manifested a strong right lateralization, whereas modifications in oxy- and total hemoglobin showed a medial localization. The present work successfully explored the capability of fNIRS to detect the two neurophysiological categories under investigation and distinguish their activation patterns. © 2012 Society of Photo-Optical Instrumentation Engineers (SPIE). [DOI: 10.1117/1.JBO.17.5.056005]

Keywords: near-infrared spectroscopy; optical imaging; time-resolved imaging; working memory; sustained attention; cognitive load.

Paper 11583 received Oct. 6, 2011; revised manuscript received Mar. 8, 2012; accepted for publication Mar. 27, 2012; published online May 4, 2012.

1 Introduction

Working memory (WM) is the cognitive function resulting from the active short-term memory system that guides and controls humans’ behavior. It allows the temporary storage and the further recollecting of information.1 Importantly, it is a limited-capacity resource, which thus can easily saturate under extremely stressful conditions, with excessively rapid streams of information, and in disproportionately complex environmental settings.

Psychophysiology has demonstrated sustained activity in the frontal and parietal brain areas while information is held in mind.2–4 Visual WM, specifically, involves the concerted activity of a distributed neural system, including anterior areas in prefrontal cortex (PFC) and the visual cortex posteriorly.5 Within the visual cortex, ventral stream areas are selectively involved in object vision, whereas dorsal stream areas are involved in spatial vision.6 This domain specificity appears to extend forward into PFC, with ventrolateral areas involved mainly in WM for objects and dorsolateral areas involved mainly in WM for spatial locations.2,4,7

Moreover, interactions between the inferior temporal neocortex, PFC, and parietal cortex likely support WM maintenance via an attention-based mechanism, enhancing the relevant information among irrelevant background noise.8,9 Overlap of the cerebral networks of WM and attention has been demonstrated in targeted comparisons.10,11 Specifically, superimposition of activation was observed in the frontal midline and in a part of the right and inferior frontal gyri in functional magnetic resonance imaging (fMRI) studies.12 As a result, competition for processing resources that are shared by the WM and attention systems can lead to a severe limitation of neural processing capabilities. One major bottleneck of information processing, indeed, arises from the common demands on neural resources shared between visual WM and attention during the encoding stage.

Some regions in the left PFC and in the inferior temporal cortex have shown selective responsiveness to cognitive load during a WM task.11 Areas selectively responsive to high attentional demand have been highlighted by Mayer et al.11 within the right prefrontal and bilateral occipital cortex. Consequently, they inferred a hemispheric specialization, with left PFC selectively responsive to WM load and right PFC selectively responsive to attentional demand. On the other hand, some evidence suggests that the medial temporal lobe (MTL) and the PFC serve complementary roles for WM processes: whereas the PFC appears to be predominantly important for WM maintenance of familiar items such as words or numbers, the MTL may

Address all correspondence to: Erika Molteni, Dipartimento di Bioingegneria, Politecnico di Milano, Piazza Leonardo da Vinci 32, I-20133 Milan, Italy. Tel: 390223993322; E-mail: erika.molteni@polimi.it

Journal of Biomedical Optics 056005-1

May 2012 • Vol. 17(5)
play a role for maintenance of unfamiliar novel items.\textsuperscript{12} In addition, it has been shown that multi-item memory processing depends on the organization of item representations by the same MTL.\textsuperscript{13} Rissman et al.\textsuperscript{14} have put forward that two neural circuits may dynamically trade off to accommodate the particular mnemonic demands of the task: some results credit the establishment of interactions between frontal gyrus and fusiform face area, mediating maintenance of information at lower cognitive loads, and the involvement of hippocampus, supporting the retention when a higher effort is needed. In addition, Jaeggi et al.\textsuperscript{14} have specifically investigated the modulation effect induced by the mnemonic demand of the task on the activation in the PFC. By showing that activation increases with memory load in the PFC, even when the capacity of the WM's central executive system is exceeded, they definitively uncoupled neural activation from behavioral accuracy measures and performance.

Areas of high neuronal activation show increased oxygen consumption and enhanced blood supply to ensure provision of oxygenated hemoglobin.\textsuperscript{8,15} As a consequence, neural activation can be detected by measuring changes of regional hemoglobin in cerebral blood. Since its development in 1977,\textsuperscript{16} functional near-infrared spectroscopy (fNIRS) has been increasingly employed as a neuroimaging technique: fNIRS allows for the noninvasive monitoring of temporal and spatial changes in oxygenated and deoxygenated hemoglobin concentrations (\textit{O}_{2}Hb and HHb, respectively),\textsuperscript{17-21} thus providing indexes of brain activation. In particular, multichannel fNIRS can measure cortical activation during cognitive processing without interfering with task procedure; indeed, it has been successfully employed in protocols evaluating psychological and emotional pressure\textsuperscript{22} and in pediatric studies.\textsuperscript{23}

Schreppel et al.\textsuperscript{24} investigated the effects of interference resolution processes on activation of the PFC in a WM task by means of fNIRS. Shibuya-Tayoshi et al.\textsuperscript{25} and Nakahachi et al.\textsuperscript{26} employed fNIRS devices to obtain a mapping of the activation of the PFC during the visuospatial WM Trail Making Test. They reported bilateral activation, primarily in the PFC. Lee et al.\textsuperscript{27} examined qualitatively distinct types of error trials during the execution of a delayed WM task to further specify the behavioral and neural difference between errors that arise from a loss of mental representation versus those that arise from the encoding of an incorrect stimulus; they applied both fNIRS and fMRI for a comparison of activations. To our knowledge, though, no study has investigated the capabilities of fNIRS devices to detect hemodynamic changes due to variation in cognitive load. Although fMRI and positron emission tomographic (PET) studies have fathomed the matter,\textsuperscript{14,28-31} the applicability of easy-to-use, noninvasive optical instrumentation could open the investigation to much broader cohorts.

Last, the fNIRS technique has been employed for the investigation of WM abnormalities in a number of psychiatric cohorts, ranging from patients with major depressive disorder,\textsuperscript{32} to those with schizophrenia\textsuperscript{33} and children showing attention deficit/hyperactivity disorder.\textsuperscript{34}

The present study mainly introduces two items of novelty with respect to the literature mentioned above: (1) the employment of NIRS devices in the investigation of the modulatory effect arisen in the frontal hemodynamic pattern by means of stimulation with differential cognitive load and (2) the use of time-domain fNIRS (based on high-repetition-rate light sources and picosecond detection of photon time-of-flight) to improve depth sensitivity to cortical hemodynamic changes.\textsuperscript{35-38}

The aims of this work were (1) to determine whether multichannel time-domain fNIRS is able to discern differences in frontal patterns of activation due to pure attentional and WM tasks and (2) to assess time-domain fNIRS sensitivity in detecting modulation effects induced by the mnemonic demand of a cognitive task with graded levels of difficulty. This was done by focusing on frontal cortical activation during an n-back task with different stages of memory requirement.

## 2 Materials and Methods

### 2.1 Time-Domain fNIRS

Two pulsed diode lasers, operating at frequencies of 690 and 820nm, with 80 MHz repetition rate and 1 mW average power (PDL, Picoquant GmbH, Germany), were used as light sources. The laser heads were connected to multimode graded index fibers (50/125 μm) by means of a homemade coupler that combines a neutral density attenuator (554-082, Edmund OptiK GmbH, Germany), with variable attenuation in the range 0 to 80 dB, and a standard FC fiber optics coupler. Multimode graded index optical fibers (50/125 μm) with different lengths and a 2-by-2 fused fiber optic splitter (VISNIR5050, OZ Optics, Canada) were used to time multiplexing of the laser pulses at the different wavelengths and to create two independent channels. In each channel, a 1-by-9 fiber optic switch (F-0SM9, PiezoJena GmbH, Germany) created up to nine independent sources or injection points; therefore 18 sources were available.

Four parallel detection chains accomplished acquisition of time-resolved reflectance curves. The parallel use of the four detection and acquisition lines enabled a total of 16 independent detectors. Fiber optic bundles, 1.5 m length (Loprok GmbH, Germany), were used for light collection from human tissue.

The system was controlled by a personal computer (Pentium IV 3.5 GHz, 2 Gb RAM), which hosted the acquisition boards and stored data. Personalized software, written in C language, was used to control the instrument. The software was interfaced to a microcontroller unit (dsPIC30F6014, Microchip Technology, AZ), which was used for the hardware control of the instrument. The microcontroller unit generated trigger signals for the synchronization of data acquisition by the TCSPC boards and of the sequential activation of the sources. The instrument response function (IRF) obtained by filling all the propagating modes of the bundle had a FWHM of approximately 500 ps. A detailed description and characterization of the system can be found in Contini et al.\textsuperscript{19}

The system also interfaced with dedicated software (Presentation, Neurobehavioral Systems, Albany, CA) for precise stimulus delivery and experimental control program.

For each wavelength \(\lambda\), a reference time-domain fNIRS curve \(R_0(\tau, \lambda)\) was derived by averaging the tracks recorded during the initial baseline period. Fitting of \(R_0(\tau, \lambda)\) yielded the reference absorption value \(\mu_\text{at}(\lambda)\). Then, at each recording time \(\tilde{\tau}\) during the experiment, changes in the absorption coefficient were derived\textsuperscript{39} as

\[
\Delta \mu_\text{at}(\lambda, \tilde{\tau}) = -\frac{1}{\nu \sqrt{\tau}} \ln \left( \frac{R(\tau, \lambda, \tilde{\tau})}{R_0(\tau, \lambda)} \right),
\]

where \(\nu\) is the speed of light in the medium, \(\tau\) is the arrival time of photons, and \(R(\tau, \lambda, \tilde{\tau})\) is the time-domain fNIRS curve at the recording time \(\tilde{\tau}\). To enhance the contribution from deep layers and remove possible disturbances caused by...
superficial layers, a correction method based on the use of late time windows (τ = 1750 to 2500 ps) was also applied.41,42 It is well known that depth information in time-domain fNIRS is encoded in the time-of-flight of photons:37,38,43 early photons probe the superficial layers of the head, whereas late photons have a higher probability to visit deeper layers. Finally, the absorption coefficient was then derived from corrected late gate intensities as

\[ \mu_a(\lambda, \overline{\tau}) = \mu_{a0}(\lambda) + \Delta \mu_a(\lambda, \overline{\tau}). \]  

Taking the assumption that, in the wavelengths applied here, O₂Hb, HHb, and H₂O are the main chromophores contributing to absorption, and also assuming that H₂O concentration is unlikely to be subject to relevant modifications and thus can be considered constant, O₂Hb and HHb concentration changes were derived by Lambert-Beer law. Then, changes in total hemoglobin content (rHb = HHb + O₂Hb) were calculated.

Finally, low-frequency and high-frequency components were extracted from the hemodynamic data by use of a digital filter (ninth-order digital Chebyshev filter, Matlab) with the stop-band 30 dB down. Transition frequency was selected as the inverse of the period of the task modulation.

2.2 Optodes Localization

The optical probe was placed over the head to cover the underlying prefrontal cortex, and light sources were centered (according to the international 10 to 20 system for the EEG electrode placement) at the Fp1 and Fp2 for left and right sides, respectively (Fig. 1). The source-detector distance was 2 cm, in accordance with previous experiments.44 It should also be considered that, in time-resolved fNIRS, the capability of photons to penetrate brain tissue does not depend on the inter-fiber distance.43 No downward sliding of the fiber holders was observed at the end of the measurements in any subject. Custom-made fiber holders were designed to keep fibers on the forehead with black rubber pads. Holders were kept in place on the head with two-sided tape and velcro bands. The minimum distance between the light source located on the left (right) forehead and detectors placed over the right (left) forehead was 5 cm. This distance was chosen after a trial, performed with both phantoms and in vivo measurements, aiming to rule out cross-talk between optical signals detected simultaneously in the two areas.

Fig. 1 Position of fNIRS probes (injection sources in red, collection bundles in green) on the subject’s head.

2.3 Statistics

To examine whether there were statistically significant changes in fNIRS signals, amplitude analysis on the O₂Hb, HHb, and tHb measures was performed. A t-test was conducted for assessing differences between fNIRS data acquired over the two hemispheres (right versus left). Then, repeated measures ANOVAs were conducted, with two within-group factors: condition (rest and loads) and channel. Post-hoc analyses were performed with pairwise ANOVAs. A P value of 0.05 was considered to be statistically significant. The Greenhouse-Geisser procedure was used to correct for sphericity violations, where necessary. t-Tests were also conducted on behavioral data, to assess differences in performance between conditions. All the statistical analyses were performed with NCSS Statistical Software (NCSS Statistical & Power Analysis Software, http://www.ncss.com).

2.4 Generalized Linear Model

Both low- and high-frequency components of O₂Hb, HHb, and tHb time courses (time series) were analyzed with the generalized linear model (GLM), which is a well-known statistical method suitable for analyzing time series data.45 The GLM approach has been extensively applied in fMRI data analysis46 and has been introduced for analysis of fNIRS data.50

A time series for a given channel was modeled as a linear combination of L regressors (known functions) plus an error term:

\[ Y = X \beta + \epsilon, \]

where \( X \) (\( r \times L \) matrix) is called the design matrix (DM), each column is a regressor, and \( r \) is the number of time points. Each column \( L \) of matrix \( X \) contains the predicted hemodynamic response for one experimental block over time \( \tau \). \( \beta \) is the vector of the unknown parameters, one for each regressor, weighting (i.e., quantifying) the contribution of each predictor (i.e., regressor) for modeling the functional time series (i.e., columns of matrix \( Y \)), and it serves as the parameter set for subsequent hypothesis testing. Errors \( \epsilon \) are assumed to have normal distribution with zero mean and covariance \( \sigma^2 V \), where \( \sigma^2 \) represents variances and \( V \) is the temporal correlation matrix (identical at all voxels).

Origin for correlation on error terms can be found in cardiac, respiratory, and vasomotor effects. The ordinary least squares estimation of the parameters is

\[ \hat{\beta} = (X^T X)^{-1} X^T Y. \]

Inference is then performed by testing the predictor variables through the \( t \)-statistic:

\[ t = \frac{\hat{X}(\tau)}{\sqrt{C_\hat{\beta}(\tau)}}. \]

where \( \hat{X}(\tau) \) is the product of the transposed contrast vector with the response signal strength and \( C_\hat{\beta}(\tau) \) is the error variance.

The interest of this work was the analysis of the mean level of brain activation, deductible from levels of O₂Hb and HHb, and of their time course. The task was modeled as a series of
consecutive boxcar regressors, each one marking with a unit value a single condition (e.g., two activation periods with the same cognitive load) and with a zero value all the remaining part of the task. The timing of the regressors was chosen according to the different activities performed during the experiment. The GLM was applied separately for $O_2Hb$, $HHb$, and $Hb$ data. The hemodynamic response function was the same for $O_2Hb$, $HHb$, and $Hb$ data. First-level analysis (i.e., on single subjects) was performed first. Activated areas were extracted by setting uncorrected $P$ value of the $t$-statistics at 0.05.

Inferential procedures at a second level (e.g., about activated channels over all subjects) were then performed, by setting uncorrected $P$ value of the $t$-statistics at 0.001, and the DM was accordingly modified.

Contrast matrix was designed for a direct comparison of (1) activation during the two task periods and rest condition with respect to the reference (whole test) [see Fig. 2(a)], (2) activation during each cognitive load and rest condition with respect to the load/rest immediately harder/easier [see Fig. 2(b)], and (3) activation during each cognitive load and rest condition with respect to the reference (whole test) [see Fig. 2(b)].

Cerebral activation was indicated by positive $t$ values in the $O_2Hb$ and $tHb$ time series, and by negative $t$ values for $HHb$. The results from these second-level analyses were plotted as statistical parametric maps in NIRS_SPM, which illustrate the brain regions where increased or decreased $O_2Hb$, $HHb$, and $tHb$ correlate with the stimulation protocol over time.

As shown in Butti et al., Pearson product moment correlation coefficient at lag 1 of basin fNIRS signals is not significantly different from zero for both $O_2Hb$ and $HHb$ series, if the $P$ value is set at 0.05. Thus, no corrections for serial correlations were applied in the GLM.

### 2.5 Activation Task

Each participant was given a computerized variant of the $n$-back task, using letters of the alphabet as memoranda. The $n$-back task is a verbal memory test that provides four levels of memory load (0-, 1-, 2-, and 3-back), presented in a factorial design. It yields a total of eight task blocks (four conditions, repeated twice). For each difficulty modality, computerized lists of 30 stimuli were constructed and then presented in a pseudorandom order, making up a total of 240 intermingled letters shown to the subject. A block of a single condition lasted 67 s, consisting of 5-s presentation of task instructions (e.g., “2-back letter task”), followed by a 2-s pause and then 30 stimuli (250-ms presentation of a stimulus, followed by a 1750-ms blank-screen interstimulus interval [ISI]). Each block of a single condition was presented separately, and between the blocks a rest of 30 s was given. A 60-s baseline and 60-s recovery were acquired at the start and at the end of the test, for a total acquisition time of 16 min.

At the end of the first four blocks (0-, 2-, 1-, and 3-back), in the middle of the test, subjects saw the word REST for approximately 90 s before the same four blocks, in different order (2-, 1-, 3-, and 0-back) began, providing a supplementary rest period. On each trial, subjects observed stimuli presented in the center of a computer screen (Presentation software, Neurobehavioral Systems). Letters were presented in 24-point Helvetica. Subjects responded to each stimulus presentation by pressing the left or right button on the mouse in their right hand. To respond to a stimulus as a target, subjects pressed the button under their index finger (left button); to respond to a stimulus as a nontarget, they pressed the button under their middle finger (right button).

In the 0-back condition, the target letter was any letter that matched the one specified during the instructions given at the beginning of the block (in this test, letter “A”). In the 1-back condition, the target was any stimulus identical to the immediately preceding stimulus. In the 2-back and 3-back conditions, the target was any stimulus identical to the stimulus presented two or three trials prior, respectively (Fig. 3). Stimuli were targets on 20% of trials, at all levels of load (every condition).

Subject performance during recording was monitored in terms of reaction time and accuracy (number of target letters identified correctly). Any correct reaction to the stimulus (both target and nontarget) was a commission; any incorrect answer to the stimulus was a commission error. Any missing response (to both target and nontarget stimuli) was an omission error.

---

**Fig. 2** Design matrix DM1, modeling baseline, task, and rest (a); time course proceeds from top to bottom, and baseline, task, rest, and reference functions are aligned from left to right. Design matrix DM2, modeling baseline, the four cognitive loads, and rest (b): time course proceeds from top to bottom, and baseline, 0-, 1-, 2-, and 3-back, rest, and reference functions are aligned from left to right.
Before beginning the experimental task, all participants were trained on the task and were given a pretesting session to ensure that they had satisfactory understanding of the instructions. It is well established that the n-back task utilizes executive functions, and that these functions are located in prefrontal cortex.

2.6 Subjects

Nineteen healthy volunteers, comparable for age (mean 27.15 years, SD 2.01, range 23 to 30) and school attendance (mean 19.9 years, SD 1.2, range 17 to 21), participated in the experiment. Seventeen subjects were right-handed. All volunteers were native Italian speakers and were not paid to participate. The neuropsychological status of each participant was assessed by means of an unpublished adapted Italian version of the Neuropsychological History Questionnaire. They all had normal vision and had no history of psychiatric disorders. None of them showed any neuropsychological illness; cognitive level and attentive capability were normal; none had any first-degree relatives with a psychiatric illness. The study was approved by the Institutional Review Board. Written informed consent was obtained from all volunteers after the examination and test procedure had been explained.

3 Results

3.1 Behavioral Results

The 19 subjects committed an average of 1.58 errors (SD 0.69) in the n-back task, thus falling in the normal range. On average, they committed 0.77 errors during the 0-back condition, 1.28 errors during the 1-back, 2.14 during the 2-back, and 2.15 during the 3-back. They answered correctly to the target in 511 ms (SD 36). More specifically, reaction times on targets were 464 ms for the 0-back condition, 508 for the 1-back, 540 ms for the 2-back, and 535 ms for the 3-back.

A statistically significant difference was found for errors committed at the four different cognitive loads (P < 0.05), as well as for reaction times, grouped for the same four loads (P < 0.05).

3.2 fNIRS Results

Results emerging from time-domain fNIRS acquisitions show that variations of O2Hb remained positive in the PFC, over both hemispheres and during the entire test (Fig. 4). Activation level increased during the first half of the test, reaching a peak in the central part of it; then, activation remained stable during the second part of the task, with a final and further increase in the last recovery period. In addition to the large, persistent changes described hitherto, transient and less prominent changes were observed time-locked to the active blocks presentation.

Variations of HHb globally stayed in a range of negative values throughout the test. Right channels showed an initial decrease and an ensuing increase in value, which moved toward the restoration of the rest value; left channels exhibited an initial decrease of HHb as well, with a delayed recovery.

Globally, larger variations of O2Hb and HHb were found over the right frontal area, with respect to the contralateral hemisphere. This hemispheric lateralization, visually discerned on raw data, was later confirmed by statistical analysis: t-test confirmed the statistical difference between the average signals from right and left channels for both O2Hb and HHb. The prevalence of activation on the right side was then assessed, and the right PFC showed higher O2Hb increases than the left one.

Moreover, repeated-measures ANOVA conducted on HHb revealed a strong effect of time (P < 0.001) and time-by-channel interaction (P < 0.001), while a main effect of channel was not observed (P > 0.05). The effect of time is paramount; the task generated a strong hemodynamic response throughout its duration, presenting a momentary decrease only in correspondence with the central rest period. Standard deviation increased with time within both parts of the task, thus testifying to a greater uniformity of behavior among subjects during the “early” response to the task (and then a more pronounced intersubject variability for the long-term response). Repeated-measures ANOVA conducted on O2Hb attested to a significant main effect of time and time-by-channel interaction (P < 0.001). Effect of channel was not statistically significant (P > 0.05).

To tentatively discriminate between the tonic contribution to activation, due to subject’s involvement in the cognitive challenge, and the phasic cerebral activity, which is more specifically linked to the working memory effort, low- and high-frequency components were extracted from the fNIRS signals. The two contributions were separately considered, with reference to the performance of the cognitive task (Fig. 5).
As mentioned in the previous section, high-frequency components of the signals were first analyzed by means of DM1 [Fig. 2(a)], and then statistical τ-maps were calculated for the analyzed group of volunteers using the t-value parameters. These interpolated maps were obtained through the resolution of the complete GLM at group level. Each map represents the distribution associated to a specific contrast between two regressors. Response of the PFC to task presentation is given in Fig. 6(a) for O2Hb, HHb, and tHb. The temporal trend noted above is confirmed by these activity maps. The task originated a strong metabolic response, leading to an increasing difference between parameter values registered during the test and during the baseline period. The test, indeed, was characterized by a strong and steady activation. Unlike the initial baseline period, the interleaved and final recovery periods could not provide any significant deactivation, if compared to the task. Modifications were spread over the PFC, attesting to a general involvement of the frontal lobe in this task. Some hemispheric lateralization to the right was found for HHb, whereas O2Hb and tHb showed their activation foci medially.

High-frequency components were then analyzed by means of DM2 [Fig. 2(b)] for a load-by-load signal analysis. Results seemed to suggest that high-frequency activity underwent a modulation induced by the task presentation sequence, thus following the alternation of task/rest blocks [Figs. 5 and 7(c)]. Owing to this characteristic, this component should be the one that mostly represents the real cortical activity related to the WM effort. In Fig. 6(b), the 0-back (pure attentive) condition is contrasted against baseline. Maps of O2Hb, HHb, and tHb are presented. HHb shows a strongly significant right lateralization. O2Hb and tHb maps attest to a different distribution of cerebral oxygenation over the PFC during the lightest period of the task, showing a peak of statistics medially.

Figures 6(c), 7(a), and 7(b) show the pseudocolor representations of the t-statistic values of O2Hb, HHb, and tHb courses (high-frequency signals) for 1-back, 2-back, and 3-back loaded conditions, respectively, contrasted with the baseline. At any load, HHb shows a strong activation, considerably lateralized over the right hemisphere. O2Hb also shows activation, though more uniform and medially located. Spatial distribution of tHb shows a mild lateralization to the left, more pronounced in the 1-back and 3-back conditions.

The study of neighboring loads provided us with congruent results: contrasts between 0-back (pure attentional) and higher loads (working memory conditions) highlighted significant differences in activation pattern [see Fig. 7(c) for an example of activation pattern resulting from contrast between the pure attentional and the easy WM conditions]. Contrasts between different WM loads (1-back versus 2-back, 1-back versus 3-back, etc.) showed a trend of difference in activation patterns in the left hemisphere, though not significant.

GLM was finally run for low-frequency data (by means of both DM1 and DM2). Maps revealed a faint activation of O2Hb in the 1-back and 2-back conditions. Surprisingly, HHb proved to be significant at every task load, if contrasted with baseline. A poor right lateralization was still detected for HHb; though omnipresent, it showed no statistical significance.

During the employment of GLM, the variance of model residuals was calculated, for both DM1 and DM2. Table 1 shows our findings.

4 Discussion

4.1 Methodological Issues

A limitation of the fNIRS technique is that signals contain contributions from both the brain and the overlying tissues. Specifically, the fNIRS signal can be highly sensitive to extracerebral blood volume changes. In addition, differences in the dynamics of O2Hb and HHb variations often reported in fNIRS studies have been generally interpreted as a potential indicator of differences in physiological mechanisms of neurovascular coupling across cortical areas. Surveys have also shown that this fact may indeed depend partly on hemodynamic changes unrelated to the cortical increase in blood flow. The most reliable hypothesis, then, is that this difference may be caused by systemic changes.

In contrast to other neuroimaging methods and to the continuous-wave fNIRS, time-domain fNIRS can inherently differentiate between truly cortical and systemic changes, by exploiting depth information encoded in photon arrival times: on average, photons detected after long times of flight probe deeper tissue layers than early photons. In doing so, a separation of surface and cortical contributions is made feasible. In the present work, extracerebral signal was...
isolated and treated separately to obtain neater information from the cortex. It was verified that extracerebral contribution gives no information significantly related with the task (data not shown). Further discussion on this topic is provided elsewhere.42

4.2 Discussion of Results

Both pure attentional effort and engagement of executive functions (such as WM) are known to activate the dorsolateral PFC.2–4 In spite of this, attention and higher cognitive faculties differ phenomenologically, and cannot be regarded as the same process. Rather, attention is nowadays understood to serve as an underlying process on which the other cognitive activities can graft and develop. It is then possible that physiological differences existing between attention and higher processes are mirrored in some unlikeness in the dynamics and extension of PFC activation.9 The n-back task decouples attentional and mnemonic contributions to the global cognitive effort, and in addition allows the investigation of cortical activation during incremental cognitive exertion and, eventually, fatigue.

WM errors may arise from failing to encode the target, losing internal representations during maintenance, or incorrect coding of the target held in mind.25 The first case prevents the subject’s brain from entering the maintenance step of memory, potentially causing a lack of PFC activation with respect to correct trials; the third case is not affected by the risk of an absence of activation demand, generating an incorrect recollection; the second case clearly consists of a trade-off between the other two cases. The n-back task requires a constant encoding load: velocity of encoding never changes among conditions. The difficulty of the task is modulated by the maintenance request only. For this reason, if subjects perform correctly during the 0-back condition, a decay of PFC activation due to encoding performance can be excluded and—importantly—any change in PFC activation can be attributed to maintenance only. In our study, all subjects performed very well during the 0-back condition, and they proved to be able to afford even the most difficult (3-back) condition of the task.

fNIRS activation studies have demonstrated that cognitive tasks generally cause increases of \( O_2\text{Hb} \) and \( t\text{Hb} \) associated with decreases of HHb in the prefrontal cortex.51–65 The

Fig. 6 Cortical activation during WM task: high-frequency signals. The activation maps are obtained from \( t\)-statistic values, assessed by a second-level analysis of the GLM model. From left to right: \( O_2\text{Hb}, \) HHb, and \( t\text{Hb} \) maps. (a) Each map depicts the significant difference between task condition and baseline. (b) Significant difference between the 0-back condition and baseline. (c) Significant difference between the 1-back condition and baseline. \( P < 0.001 \) for all maps.
increases of $O_2$Hb and $t$Hb reflect the occurrence of redundant regional cerebral blood flow (rCBF) increases in response to neuronal activation, and the decrease of HHb is caused by evoked-rCBF rises that exceed the increases in oxygen consumption during neuronal activity. Our results agree with these previous findings, as shown in Fig. 4. It has been reported that cortical activation results in an overshoot of regional blood supply that is disproportionately larger than the increase in oxygen demand. This phenomenon is known as the physiological uncoupling of oxygen metabolism and blood supply. Augmented arterial blood flow into the cerebral cortex washes away deoxygenated hemoglobin more rapidly, resulting in a decrease in HHb with an increase in $O_2$Hb. In this study, an explanation by the above mechanism is applicable.

**Table 1** Variances of the GLM residuals. From left to right, design matrices are listed, together with frequency contributions and values of the variance of GLM residuals for oxygenation, deoxygenation, and total hemoglobin signals. All values are obtained in the form of mean value and standard deviation from the whole population. In all cases, DM2 gave a better description of data, thus providing lower residuals.

<table>
<thead>
<tr>
<th>DM</th>
<th>Contribution</th>
<th>$O_2$Hb</th>
<th>HHb</th>
<th>$t$Hb</th>
</tr>
</thead>
<tbody>
<tr>
<td>DM1</td>
<td>High frequency</td>
<td>0.0730 + 0.0193</td>
<td>0.0160 + 0.0067</td>
<td>0.0477 + 0.0127</td>
</tr>
<tr>
<td>DM1</td>
<td>Low frequency</td>
<td>0.0121 + 0.0194</td>
<td>0.0031 + 0.0065</td>
<td>0.0071 + 0.0128</td>
</tr>
<tr>
<td>DM2</td>
<td>High frequency</td>
<td>0.0707 + 0.0079</td>
<td>0.0156 + 0.0027</td>
<td>0.0464 + 0.0045</td>
</tr>
<tr>
<td>DM2</td>
<td>Low frequency</td>
<td>0.0113 + 0.0079</td>
<td>0.0029 + 0.0027</td>
<td>0.0066 + 0.0045</td>
</tr>
</tbody>
</table>
The relative $O_2$Hb concentration in frontal brain tissue increased significantly and bilaterally during the performance of the $n$-back test, compared with a resting condition. This result can be interpreted as an expression of increased cerebral perfusion as a consequence of an activated brain metabolism during the task. Simultaneously, HHb showed a progressive reduction. This latter effect is consistent with an enhanced HHb draining, secondary to hyperperfusion.

During a two-channel fNIRS registration of a continuous performance test in healthy subjects, a significant frontal $O_2$Hb increase was found, compared with a baseline at rest, by Fallgatter and Strik.68 This result is generally interpreted as an enhanced perfusion of the frontal lobes consistent with attentive local activation. Lee et al.27 reported an increase for both $O_2$Hb and HHb in a WM task performed by healthy people. Data of the present study seem not to support this evidence, as we found an increase in $O_2$Hb with a concurrent decrease in HHb in the $n$-back task for WM assessment, as in Toichi and Kamio,69 who reported an increase in $O_2$Hb and a concurrent decrease in HHb in tasks of higher cognitive function (i.e., WM). Furthermore, Lee et al.27 also observed an increase of both $O_2$Hb and HHb in tasks of pure attention and a lack of increase in HHb in some subjects during tasks of higher cognitive processing, which we could not observe in our data. They hypothesized that the cerebral metabolic rate of oxygen consumption during tasks of attention may be higher than that during tasks of higher cognitive processing and, consistently with this interpretation, they proposed the change in HHb to be a physiological marker discriminating between attention and higher cognitive processing. The origin of such discrepancies can be found in the differences in the technologies employed: continuous wave NIRS, used in Toichi’s and Lee’s works, could provide data more affected by the superficial systemic contribution.

Our results agree with recent studies by Shibuya-Tayoshi et al.25 and Nakahachi et al.,26 who found a bilateral $O_2$Hb increase prefrontally. Interestingly, a widespread, powerful right HHb activation was constantly present in GLM maps throughout the test. fNIRS literature can provide only partial confirmations of this result.

The analysis of contrasts between neighboring loads has revealed that 0-back (pure attentional) condition and higher load conditions (WM) elicit significantly different activation patterns in the PFC [Fig. 7(c), for example]. This is consistent with the difference between the two cognitive categories at a neural level. Also, contrasts between different WM loads (1-back versus 2-back, 1-back versus 3-back, etc.) showed differences in activation patterns that were maximal in the left hemisphere, though not significant; the left lateralization is consistent with previous results produced by Mayer et al.11 in an fMRI study, while the lack of significance between loads can relate to the scanty statistical power associated with the modulation of PFC activation, which is not comparable with a newly generated activation focus. However, as the attention background is constant throughout the test of our study, the result appears to be specific for the demands of the augmented memory load, and not due to the general effects of attention, perception, and motor response. Linden et al.70 found that PFC responds to WM load beyond the capacity of the parietal-premotor network. Therefore, WM load-selective activation in the PFC might fit within the framework postulating that this brain region subserves extramnemonic processes of top-down control over posterior regions where information is actually stored.71,72 This is in agreement with the continuous increase of PFC activation through loads seen in our study.

Last, comparing Fig. 6(b) with Figs. 6(c), 7(a), and 7(b), it can clearly be seen that the overlap of the cerebral networks of WM and attention observed in the frontal midline by Mayer et al.11 likely finds confirmation in the present study. Figure 6(b), indeed, shows an $O_2$Hb midfrontal pattern due to attention only, which matches the literature. Other confirmations can be found in Hopf et al.3 and Tomasi et al.3

During the use of GLM for the analysis of high- and low-frequency data, the variance of model residuals was calculated for both DM1 and DM2. Table 1 shows our findings. In all cases, DM2 gave a better description of data, thus providing lower residuals. This result confirms that a better modeling of data is obtained if the different conditions administered during the task are separately taken into account inside the linear model.

Last, we point out that, as we considered the presence of the interhemispheric fissure, some spatial undersampling occurred in the medial area, and thus we cannot completely rule out small shifts in localizing the medial activations due to spatial interpolation. This fact, albeit marginal, could be considered one limitation of our study.

## 5 Conclusions

The present investigation by multichannel time-domain fNIRS supports results from fMRI studies, showing enhanced frontal lobe perfusion during performance of the $n$-back task. Overlapping activation for attention-demanding visual search and WM was observed in distributed frontal regions; nevertheless, the combined employment of time-resolved and GLM processing succeeded in identifying distinctive patterns of cortical activation for the two different neurophysiological categories under investigation (attention and short-term memory). Indeed, a subset of regions, in the right prefrontal cortex, showed an additive increase in $O_2$Hb activation passing from attentional demand to WM effort.

In conclusion, the present study (1) successfully explored the capability of fNIRS (and more specifically time-domain fNIRS) to detect and characterize the prefrontal patterns of activation risen by attention and WM, and (2) contributed to highlight the advantages deriving from a multiple application of GLM to fNIRS data. On the other hand, a statistically significant distinction of cognitive load could not be reached in this study.

### Acknowledgments

The authors are grateful to Michele Butti for the contribution given during the earliest part of the study. The research leading to these results has received funding from the European Community’s Seventh Framework Programme under the nEUROPt Project (grant agreement FP7-HEALTH-2007-210706). The authors declare that no conflict of interest exists for the publication of this paper. They declare that this work has never been published previously. Authors all agree with the publication. If accepted, the work will not be published elsewhere in the same form, in English or other languages.

### References