Radboud University Nijmegen

Master thesis

Processing and analyzing functional near-infrared spectroscopy data

Hue Dang

Supervisors

Dr. Marian Dekker (Philips Research)

Dr. Jason Farquhar (Radboud University)

Prof. Tom Heskes (Radboud University)

September 2016

Abstract

Functional near-infrared spectroscopy (fNIRS) is a neuroimaging technique which has potential applications in monitoring brain activation and brain disorder treatment. In this thesis, we explore the technique details of fNIRS and the issues in processing and analyzing fNIRS data through the data of four experiments investigating the sensitivity of fNIRS with brain activation in visual cortex, motor cortex and prefrontal cortex which replicated experiments in literature. While different approaches in literature were replicated, an approach for processing and analyzing fNIRS data was proposed which gave comparable statistical results to that in the literature and smoother visualization. Additionally, a solution to find the locations of strongest signal was introduced with results corresponding to locations reported in literature.

Contents

1. Introduction

Brain imaging (or neuroimaging) refers to the usage of imaging techniques to enhance human understanding of brain functions, by generating images of the structure and function of the nervous system, the method is extensively used to support diagnostic, monitoring and treatment of brain diseases. (Bunge & Kahn, 2009; Irani, Platek, Bunce, Ruoccoz, & Chute, 2007).

Functional near-infrared spectroscopy (fNIRS) is the use of near-infrared spectroscopy (NIRS) for the purpose of functional neuroimaging with wavelength between 650 and 1000 nm, to measure the optical absorption changes of tissues. fNIRS works based on the relative transparency to near-infrared light of brain tissues and bone, and the light absorption characteristic of hemoglobin. The amount of hemoglobin in the tissue affects the amount of absorbed light thus using the modified Beer-Lambert law, we could calculate the relative changes of hemoglobin through the changes of attenuation light (Delpy et al., 1988; Villringe & Chance, 1997; Boas et al, 2001). During a measurement, a subject wears a cap on the head. Probes are attached and contacted to the skin through holes on the cap. Each probe has a source or detector optode inside wherein source optode emits light and detector optode quantifies the received light. Each pair of a source and detector optodes defines a measured channel.

Different from other neuroimaging techniques such as electro-encephalography (EEG) and magnetoencephalography (MEG), which directly measure the neuronal activities through the electrical impulse generated by neurons during brain activation, fNIRS and other techniques such as fMRI and PET indirectly access the neuronal activities through the increase of local cerebral blood flow which is the result of metabolic activities associated with brain activation (Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001; Cutini, Moro, & Bisconti, 2012; Irani et al., 2007; Bunge & Kahn, 2009).

fNIRS is a non-invasive neuroimaging technique. Other common non-invasive techniques are fMRI and EEG. While fMRI could measure activities in any interested area of the whole brain at high spatial resolution (mm^3), fNIRS could measure activities at the surface of the brain only and at low spatial resolution (cm^2). However, fNIRS has several potential advantages. fNIRS is less sensitive to motion than fMRI thus more comfortable for many groups of participants such as children, more suitable for long measurement and task which requires movements. Furthermore, fNIRS system and operating costs are much cheaper than fMRI (Irani et al., 2007; Bunge & Kahn, 2009).

Compared to EEG, fNIRS is easier to use in practice as it does not require gel to attach the sensors on the head of the subjects. The spatial resolution of fNIRS is better than that of EEG however if considering the hemodynamic response to stimuli in fNIRS to compare with the neuronal response to stimuli in EEG, the temporal resolution of EEG is much higher than that of fNIRS. EEG and fNIRS complement each other and many researches have been conducted using the two techniques at the same time (Logothetis et al., 2001; Cutini et al., 2012; Irani et al., 2007; Bunge & Kahn, 2009).

Despite the advantages and more than 20 years of development since the coming of the first commercial fNIRS device, the application of fNIRS in clinic is still not widespread (Obrig, 2014; Smith, 2011).

In the scope of this thesis, we will review the technical details of fNIRS neuroimaging and explore the application of fNIRS via four experiments conducted at Philips Research which replicated the early works in literature as follows. The Visual stimulation experiment replicated the third experiment in Wijeakumar, Shahani, Simpson, and McCulloch (2012) wherein the brain activation in visual cortex was examined during the presence of flickering checkerboard. Similarly, the Finger tapping experiment, which replicated

the experiment in Kuboyama, Nabetani, Shibuya, Machida, and Ogaki (2004), investigated the brain activation in motor cortex during different finger tapping frequencies. The Continuous Performance Test, which replicated the experiment in Fallgatter and Strik (1997), and Verbal-Fluency Test, which replicated that in Herrmann, Ehlis, and Fallgatter (2003), studied the brain activation in prefrontal cortex during cognitive tasks.

1.1. Research questions

Based on the four experiments conducted at Philips Research we would like to investigate the reproducibility of fNIRS and if the reported methods in literature could be improved which leads to the first two research questions: ***Are the results from literature with similar experiments reproducible? And to what extend could we improve and automate the fNIRS data processing and analysis for the purpose of brain activity imaging?*** It is also of interest to study the spatial relationship between locations to take the advantage of the measurements in multi-channel data. This leads to the third question: ***Do the locations of strongest signal correspond with the locations reported in literature at subject and group level?*** In addition with the spatial analysis, we would also like to study the difference between left and right hemisphere via the final research question for the Finger tapping experiment: ***Can we find a significant difference between left and right hemisphere per frequency at subject level?***

The thesis is organized as follows. Chapter 2 describes more technical details on fNIRS and provides intensive review over the noise removal methods for fNIRS data. The first question is tackled in Chapter 3 where description of experiments, replication steps, and the corresponding results are discussed. Chapter 4 describes a proposed approach for preprocessing and analyzing data of the four experiments which answers question 2. Chapter 5 introduces a method to answer the questions related to spatial analysis. Finally, chapter 6 concludes the thesis with findings and recommendations.

2. Background

This chapter describes technique details over fNIRS and provides a review over the noise removal methods for fNIRS data.

2.1. NIRS devices

There are three types of NIRS devices corresponding to three different NIRS techniques, which are continuous-wave (CW), frequency-domain (FD) and time-domain (TD). CW-NIRS measures light attenuation only therefore it could not provide absolute hemoglobin concentration. FD-NIRS measures light attenuation and phase delay, and TD-NIRS measures the shape of the pulse of light as well. The latter two could provide absolute hemoglobin concentration however the devices are much more complex, bulky and expensive than that of CW-NIRS (Delpy & Cope, 1997; Ferrari & Quaresima, 2012; Scholkmann et al., 2014). CW-NIRS devices are simpler, cheaper and portable therefore they are more popular in use than the others.

In the scope of this thesis, as CW-NIRS device was used to conduct experiments, the discussion is applicable for CW fNIRS only. Concentration changes were calculated based on the modified Beer-Lambert law, see Appendix B for more details.

2.2. Accuracy and reliability of fNIRS

For CW-NIRS, concentration changes are calculated using equation (4) and (5) which require to have the value of extinction coefficients and DPF at each wavelength. There are several sources which provide molar extinction coefficient of Hb and HbO at different wavelengths (Cope, 1991; Takatani and Graham, 1979; Moaveni, 1970; Schmitt, 1986). The value of DPF could be measured using FD or TD-NIRS devices. DPFs for various tissues at some wavelengths were provided in Duncan et al. (1995) and Essenpreis et al. (1993), however these values vary across subjects, over time and type of tissues (Duncan et al., 1996). As discussed in Scholkmann et al. (2014), the Modified Beer-Lambert law works based on the assumption of constant change in hemoglobin concentration in tissue which are not true in practice, thus it does not give precise values in quantification. Correction could be done through DPFs however when quantification is not important, which is the case in brain activation monitoring, this step could be skipped as the trend of the signal is still the same. Comparisons of concentration changes calculated using different combinations of extinction coefficients and DPFs are described in Appendix C.

The reliability and reproducibility of fNIRS regarding to visual cortex, motor cortex and prefrontal cortex have been investigated in several studies such as Plichta et al. (2006), Plichta et al. (2007), Kono et al. (2007), Schecklmann, Ehlis, Plichta, and Fallgatter (2008), and Biallas, Trajkovic, Haensse, Marcar, and Wolf (2012). In their experiments, each subject participated the same experiment two or three times with the interval between times varied from 1 day to 1 week, 3 weeks, or 52 weeks. Their results showed that the results were highly reproducible at group level however at single subject level, the reproducibility was low.

2.3. Noise removal

fNIRS signal contains much noise due to the way it is measured and noise should be partly or fully removed for subsequent analysis. Three sources of noise in fNIRS signal, instrumental noise, experimental errors and physiological noise, will be reviewed according to the categorization in Huppert, Diamond,

Franceschini, and Boas (2009) in the following subsections. We will review method based on their effectiveness and automation feasibility.

2.3.1. Instrumental noise

Instrumental noise is often considered as random or white noise and the techniques such as block averaging, smoothing and low-pass filter are often used (O'Haver, 2016).

Smoothing is a simple technique based on the value of adjacent points to create a smoother signal. It uses sliding-average of adjacent points. With smoothing, points' amplitudes which are higher than that of the neighbors will become lower and vice versa. The result signal will be smoother for a larger smooth width, however the original shape could be distorted when smooth width becomes too large. It is difficult to determine the optimum smooth width and depending on the type of smooth technique, optimizing parameters may be needed.

Low-pass filtering works based the transformation of the signal into frequency domain and high-frequency components will be cut off with the assumption that the noise is at high frequency and true signal is at low frequency. The signal is then transformed back to time domain. This technique minimizes distorting true signal if the instrument noise is fully modelled by high-frequency components and an optimum cut-off frequency can be determined. There are several type of low-pass filter such as Butterworth filter and Chebyshev filter; depending on the type of filter, more parameters may be needed. In fNIRS signal, often a complex noise frequency is presented and hence an optimal cutoff for low-pass filter is difficult to determine.

For both smoothing and low-pass filtering methods, once optimal parameters are obtained, the techniques can be automated. Noise reduction methods improves visualization however its effectiveness should be confirmed by a modelling result such as final statistical tests.

2.3.2. Motion artifacts

When the optodes are not well secured on the head of the subject, head movement during the measurement often leads to spikes and usually couples with baseline shift in the measured intensities, which are called motion artifacts (MA) (Huppert et al., 2009; Fekete, Rubin, Carlson, & Mujica-Parodi, 2011; Brigadoi et al., 2014). While spikes could be easily spotted, baseline shift is difficult recognized and often confused as hemodynamic response (Figure 1). Therefore the common attitude is rejecting data which contain MAs (Bauernfeind, Wriessnegger, Daly, & Müller-Putz, 2014; Brigadoi et al., 2014). However when the dataset is small or MAs are unavoidable in some group of subjects such as children, an approach to remove or reduce MAs would be better than rejecting data.

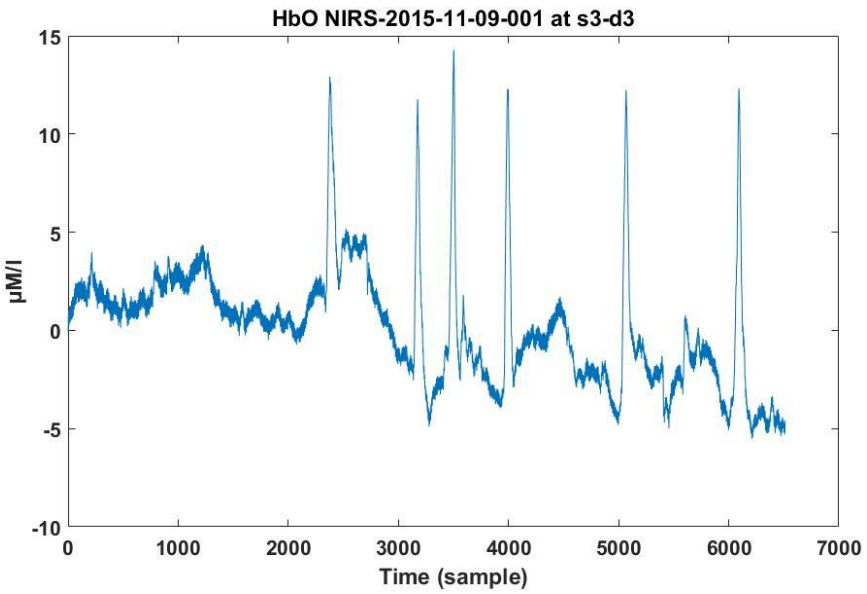


Figure 1. Spikes and baseline shift.

Several methods have been introduced for MAs removal. In the first group, extra measure is collected concurrently with fNIRS data for example by using pairs of source and detector optodes which are very close to each other such that the light does not reach the brain surface (reference channels). This extra data could be used to support the noise removal from the signal in the active area (Zhang, Strangman, & Ganis, 2009). This approach is promising since motion artifacts are expected to present in both signals with similar pattern. Unfortunately, the extra measurement was not available in this study. In the second group, where no extra data available, some prominent methods are described as follows with their limitations.

The Principle Component Analysis (PCA) method (Zhang, Brooks, Franceschini, & Boas, 2005; Wilcox, Bortfeld, Woods, Wruck, & Boas, 2005), based on the observation that when MA occurs, similar peaks appear at the same time points in different channels with different levels of amplitude (Figure 2). The authors argued that these variances were caused by MAs and they were correlated. Thus applying PCA on the segments corresponding to the MA in all channels, the systematic variance caused by MA could be extracted by a number of principle components, and this systematic variance is subsequently subtracted from the MA segments in the signal. This method requires a highly correlated peaks and a good estimation of the number of components to extract the systematic variance. A small number of components may not remove the MA fully and a large number may remove not only MA but also the true signal. There is no rule to determine that number of components precisely. Some subjective strategies are used in practice such as a threshold for the total variances, for example 80% (Brigadoi el al., 2014) or 85% (Wilcox et al., 2005) for the first few ranked components however this step still requires manually checking the result.

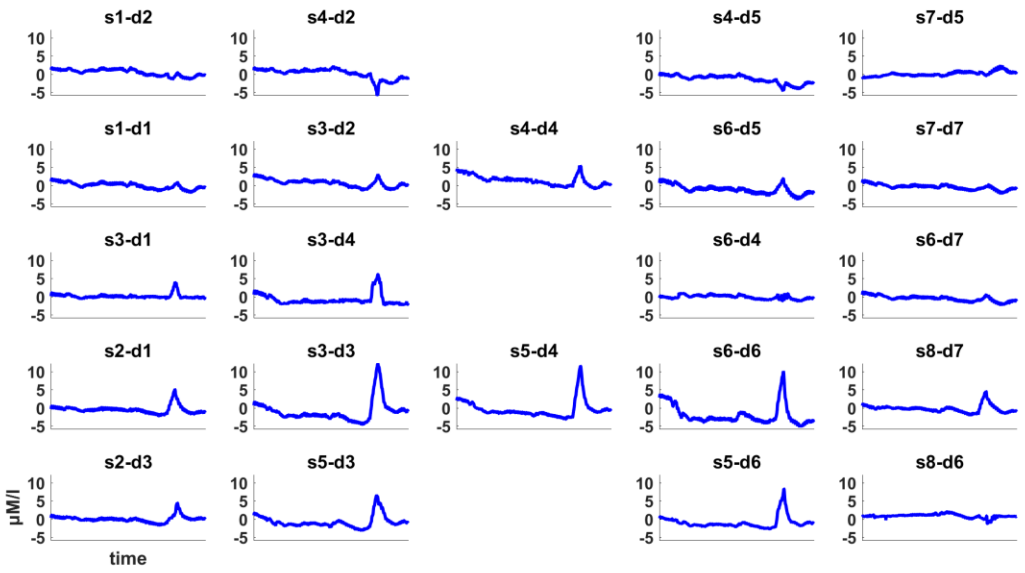


Figure 2. Data of a subject with motion artifact, each graph shows HbO changes in a corresponding channel.

Another method is the work of Cui, Bray, and Reiss (2010) based on their observation with simulation and empirical data that when MA occurs, the HbO and Hb changes are positively correlated while normally they are negatively correlated. The authors proposed that with three assumptions, HbO and Hb changes without MA and noise should be close to perfectly negatively correlated, the non-random noise and the true concentration changes should be not correlated, and a constant ratio between HbO and Hb changes during the time course. In this case when all the assumptions are met, the HbO and Hb changes after MA corrected could be calculated based on the ratio of the standard deviation of the noisy HbO and Hb changes. This method is simple, fast, and can be fully automated. However the assumptions are not realistic which may lead to unreliable result.

Spline interpolation method proposed by Scholkmann, Spichtig, Muehlemann, and Wolf (2010) corrects each MA in the signal per channel by subtracting the interpolated MA segment from the MA segment and subsequently shifting the corrected segment to match it with the baseline of the segment preceding the MA segment. This method requires the start and end point of the MA being identified and although algorithm for this purpose is provided, this process still needs manually checking as MAs vary across time, subjects and experiments.

Due to time constraint, we have not checked all available methods, but in general, spikes could be identified and treated however the biggest challenge is the correction of baseline shift because the shift can be confused with hemodynamic response.

2.3.3. Physiological noise

fNIRS signal is considered as a mixture of multiple components which are caused by neuronal or systemic activities wherein neuronal activities may be evoked by the stimulus (functional brain activities) or not (spontaneous brain activities). Similarly, systemic activities may be evoked by the stimulus (such as changes in blood pressure, blood flow in the cerebral or in the superficial layers) or not (heart rate, respiration, Mayer wave, very low frequency oscillations, in the cerebral or in the superficial layers). Any

components caused by non-functional brain activity are considered as physiological noise (Scholkmann et al., 2014).

High-frequency components could be removed by using low-pass filter. Several cutoffs have been published such as 0.5 Hz (Keles, Barbour, & Omurtag, 2016; Yücel, Selb, Boas, Cash, & Cooper, 2014; Cui, Bray, Bryant, Glover, & Reiss, 2011), and 0.4 Hz (Spichtig, Scholkmann, Chin, Lehmann, & Wolf, 2012) based on the slow change of the typical hemodynamic response (several seconds). Similarly, very low frequency components could be removed by using high-pass filter with cutoff 0.01 Hz (components last more than one minute). However components in the frequency range 0.01 – 0.5 Hz such as respiration (~0.3 Hz) and Mayer wave (~0.1 Hz) have frequencies overlapped with that of the typical hemodynamic response, which could not be removed using low- or high-pass filter. It should be noted that, some lower cutoffs were also published in literature however they are more applicable for other purposes such as classification which does not require to have the full range of true hemodynamic response while for the purpose of brain activation monitoring that is crucial.

Several methods have been proposed. Among them, the most common technique is block averaging, also called conventional averaging, which is very common in neuroimaging field by repeating the task several times and considering the average per time-points of all repetitions (Wolf et al., 2002; Scarpa, Cutini, Scatturin, Dell'Acqua, and Sparacino, 2010; Scholkmann et al., 2014).

The method proposed by Zhang et al. (2005) assumes that the physiological noise such as respiration and Mayer wave dominates spatial pattern in the baseline condition data and these artifacts also exist in task condition data. In this case, similar to the Principle Component Analysis method in removing MAs, PCA can be used to model the structure of the artifacts in the baseline data from multiple channels and that can be subsequently used to remove the artifacts from the task condition data. It is assumed that the temporal covariance in the task condition is similar to that in the baseline condition. This condition is hard to meet unless the experiment should be designed for the temporal alignment (Huppert et al., 2009). As the noise artifacts are modeled from the baseline condition, the method could not remove the noise which is evoked by the stimulus in the task condition data.

A prominent approach in reducing physiological noise is using reference channels (also called short separation channels) wherein besides the standard channels whose source-detector optode distance is around 3 cm, additional channels are set whose source-detector optode distance is less than 1 cm such that the channels are sensitive only to signal in the superficial layers. The signals in these reference channels could be used to reduce noise in the signals in the standard channels (Gagnon et al., 2011; Zhang et al., 2010). Another approach is using additional measure of physiological noise such as respiration, heart rate and blood pressure measured concurrently with fNIRS and applying wavelet coherence analysis (Kirilina et al., 2013) or independent component analysis (Patel, Katura, Maki, & Tachtsidis, 2011) to remove noise. These advanced methods require time and data to investigate therefore we do not consider them in the scope of this thesis.

3. Experiments and replication

This chapter describes the experiments which were conducted at Philips Research and the replication of processing and analyzing fNIRS data.

The first experiment, Visual stimulation, related to the visual cortex, was similar to that in Wijeakumar et al. (2012). The second experiment, Finger tapping, related to the motor cortex, was similar to that in Kuboyama et al. (2004). The last experiment related to the prefrontal cortex was a combination of two experiments, Continuous performance test (CPT) similar to that in Fallgatter & Strik (1997) and Verbal-fluency test (VFT) in Herrmann et al. (2003). There are four sections in this chapter corresponding to the four aforementioned experiments. Each section describes one experiment, its corresponding results in the literature, and subsequently replication steps and the corresponding replication results.

3.1. Visual stimulation experiment

3.1.1. Literature

3.1.1.1. Method

Wijeakumar et al. in 2012 investigated the effect of high-contrast checkerboard stimuli on the primary visual cortex and posterior parietal cortex. Ten subjects, age 15-50 years, were asked to look at three modes of stimuli, two patterns reversal at a frequency of 7.5 Hz (15 reversals per second), on/off presentation, and a static checkboard. In each mode, 10 repeated trials were performed where each trial contained 30-seconds presentation of stimulus following by a 30-seconds control gray screen.

Figure 3. Two conditions in one trial in visual stimulation experiment in Wijeakumar et al. (2012).

Three experiments were conducted. In the first one, concentration changes of two subjects at position O_1 and O_2 were examined. The results showed that the HbO changes in reversing checkerboard mode (the most aggressive mode) were larger than in the other two modes, which might be the main reason that the authors used the data corresponding to this mode in the subsequence experiments. In the second experiment, concentration changes in reversing checkerboard mode of two other subjects at 15 locations over the occipital and posterior parietal cortices were examined, the results showed that concentration changes were differentiable between two conditions at locations overlying V1 (O_1 and O_2 positions). In the third experiment, data of all subjects corresponding to reversing checkerboard mode at five locations O_1, Oz, O_2, 5% above O_1 and 5% above O_2, showed increasing HbO changes after 10 seconds from onset. All the mentioned positions were based on the International 10-20 system for EEG electrode placement. The measurements were conducted with a two-channel two-wavelength Frequency Domain Multi-Distance oximeter, OxiplexTS ISS Inc. The third experiment was the one which was replicated at Philips Research, therefore all the subsequent sections are related to that experiment only.

For visualization, the last 15 seconds of the control condition in each trial of a subject were considered as baseline in that trial to perform the baseline correction.

For statistical tests, concentration changes of ten subjects at five aforementioned locations in the reversing checkerboard mode during the last 15-second of two conditions (stimulus presentation and control) were considered. RM ANOVA was used to test the effect of location, condition and the interaction between location and condition, 5 locations x 2 conditions for each type of concentration changes (Hb, HbO and tHb).

3.1.1.2. Results

The results in Wijeakumar et al. (2012) showed that:

a) The effect of location, condition and the interaction between location and condition were significant for HbO, Hb and tHb changes.

b) HbO, Hb and tHb changes at Oz were significantly lower than that at O1 and O2.

c) HbO, Hb and tHb changes at Oz were not significantly different from that at 5% above O1 and 5% above O2.

d) The increase observed in HbO changes was greater than the decrease observed in Hb changes.

e) The trend of tHb was similar to that of HbO.

3.1.2. Experiment at Philips Research

3.1.2.1. Procedure and materials

The visual stimulation experiment was performed at Philips Research to replicate only the third experiment in Wijeakumar et al. (2012). Twelve subjects, mean age 33.17 (SD=8.42), 5 females, participated the experiment after signing a written informed consent. Subjects were asked to sit still for 2 minutes to measure baseline, after that 10 trials were performed. In each trial, subjects were asked to look at the screen where a picture of a checkerboard inverting in color at a rate of 15 Hz displayed for 30 seconds (stimulus presentation condition) followed by a luminance-matched gray screen displayed for 30 seconds (control condition). One subject performed only 9 trials.

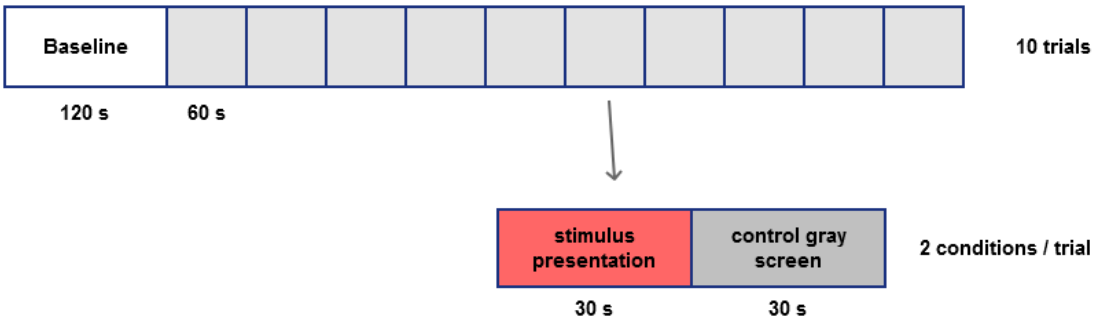


Figure 4. Visual stimulation experiment performed at Philips Research.

The measurements were conducted with CW NIRSport device from NIRx Medical Technologies LLC., which measured in two wavelengths, 760 nm and 850 nm. The sampling rate was set at 7.8125 Hz with 22 channels setting. The inter-optode distance was 30 mm. The output data were light intensities received at detector optodes which were saved in ASCII format in two files corresponding to two wavelengths. Additional files were also created which contained other information necessary for the preprocessing and

analyzing, see Appendix A. The positions of probes and channels are shown in Figure 5, the presentation of channels is illustrated in Figure 6.

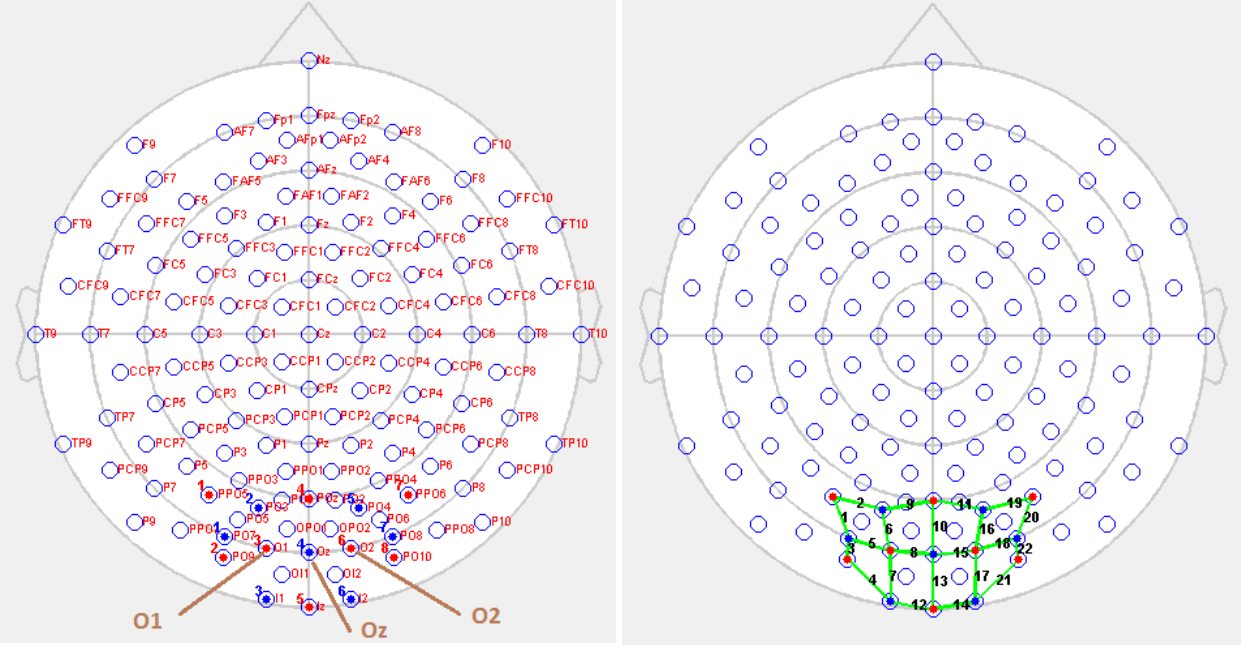


Figure 5. Topology of probes and channels in the Visual stimulation experiment conducted at Philips Research. Viewpoint is from the top of the head with the nose side at the top of the figures. Red and blue points with numbers present source and detector optodes. Left: Position of probes regarding to the EEG locations. Brown lines show the EEG position O_1, Oz and O_2. Right: Topology of the recorded channels. Green lines with number present the channels. The software nirsLAB (version 2016.01, NIRx Medical Technologies, LLC.) was used to generate these figures.

Visual Stimulation topology

s1-d2 2	s4-d2 9		s4-d5 11	s7-d5 19
s1-d1 1	s3-d2 6	s4-d4 10	s6-d5 16	s7-d7 20
s3-d1 5	s3-d4 8		s6-d4 15	s6-d7 18
s2-d1 3	s3-d3 7	s5-d4 13	s6-d6 17	s8-d7 22
s2-d3 4	s5-d3 12		s5-d6 14	s8-d6 21

Figure 6. Presentation of channels' location in Visual stimulation experiment, each cell shows the name and index of the corresponding channel. Viewpoint is from the back of the head with the top of the head at the top of the figures.

Our research question

Are there significant differences in Hb, HbO and tHb changes between two conditions: baseline, stimulus presentation including the comparison between five locations?

3.1.2.2. Data analysis

The data were analyzed using Matlab (The MathWorks Inc., MA, USA, version R2015b) and the ezANOVA function of the ez package (Lawrence, 2015, version 4.3) in R (R Core Team, 2016).

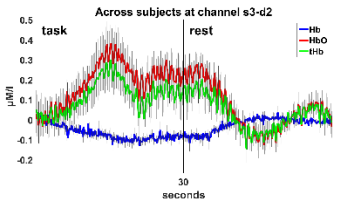
Concentration changes were calculated from the raw intensities data of each subject as described in Appendix B. Baseline was defined as the last 15 seconds of the control condition in the same trial, similar to that in Wijeakumar et al. (2012). Data of one subject were excluded due to motion artifacts which occurred in many trials and in many channels. In total there were data of 11 subjects.

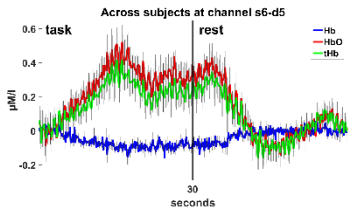
For visualization, the lengths of conditions across trials and across subjects were normalized, wherein samples at the onset of a condition whose condition's length was longer than the minimum condition's length were excluded. After normalization, the lengths of stimulus presentation and control condition were 30.21 and 30.34 seconds respectively. Baseline correction was applied per trial where the mean of the baseline was subtracted from each time point in the trial. After that, average of concentration changes per time point across all repetitions and across all subjects was calculated for each type of concentration changes.

For statistical tests, the data in the normalized length conditions were used, which did not alter the statistical test result as the number of samples which were removed at the onset of each condition was small (up to 2 samples). RM ANOVA was used to test the significance of location (O_1, Oz, O_2, 5% above O_1 and 5% above O_2), condition (stimulus presentation and baseline) and the interaction between location and condition. The data for each RM ANOVA test were the mean of tHb, HbO or Hb changes across trials in each condition per subject, per channel, 5 locations x 2 conditions x 11 subjects. Post-hoc t-test with Bonferroni adjustment on the same data was used for pairwise comparison when any of the effects in the results of RM ANOVA was significant. For each set of locations, there were 3 RM ANOVA tests and up to 9 post-hoc tests, corresponding to 3 types of concentration changes (tHb, HbO and Hb) and 3 effects (location, condition and the interaction between location and condition). All statistical inferences were based on an adjusted alpha level of 5%.

Based on the topology of the experiment (Figure 5), two sets of locations corresponding to channels 8, 10, 15, 6, 16 (s3_d4, s4_d4, s6_d4, s3_d2, s6_d5) and 6, 10, 16, 9, 11 (s3_d2, s4_d4, s6_d5, s4_d2, s4_d5) for the five locations O_1, Oz, O_2, 5% above O_1 and 5% above O_2, by order, were considered. In total, there were 6 RM ANOVA tests and up to 18 post-hoc t-tests.

3.1.2.3. Results





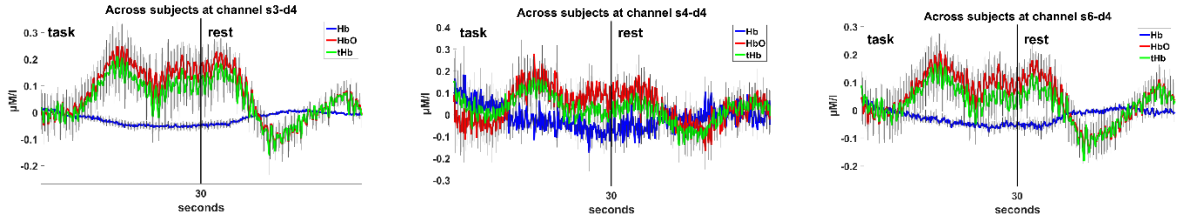


Figure 7. Visual stimulation replication: HbO (red), tHb (green) and Hb (blue) changes at s3_d2, s6_d5 (top, left to right), s3_d4, s4_d4 and s6_d4 (bottom, left to right) corresponding to 5% above O_1, 5% above O_2, O_1, Oz and O_2. Data are presented as mean and standard error of the mean (SEM). Vertical line separates two conditions: stimulus presentation (left) and baseline (right).

Table 1. Visual stimulation replication: Mean across subjects of concentration changes (μM) of the last 15 seconds of the Task condition for set of locations s3_d4, s4_d4, s6_d4, s3_d2 and s6_d5 corresponding to O_1, Oz, O_2, 5% above O_1 and 5% above O_2 (baseline correction applied). SEM stands for standard error of the mean.

		HbO (Mean ± SEM)	Hb (Mean ± SEM)	tHb (Mean ± SEM)
O_1	s3_d4	0.15 ± 0.07	-0.05 ± 0.01	0.10 ± 0.06
Oz	s4_d4	0.10 ± 0.04	-0.05 ± 0.02	0.05 ± 0.06
O_2	s6_d4	0.10 ± 0.05	-0.05 ± 0.02	0.05 ± 0.04
5% above O_1	s3_d2	0.24 ± 0.07	-0.09 ± 0.01	0.15 ± 0.07
5% above O_2	s6_d5	0.34 ± 0.07	-0.09 ± 0.01	0.25 ± 0.06

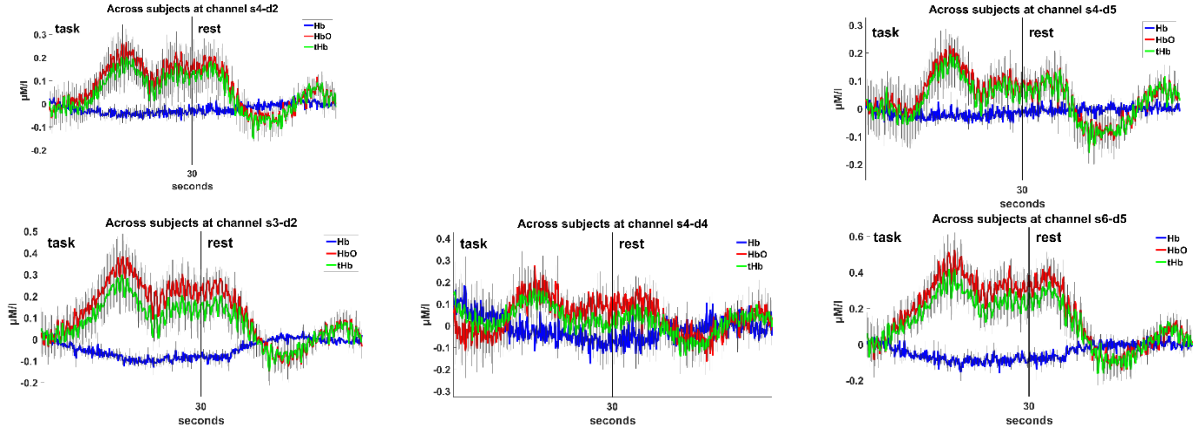


Figure 8. Visual stimulation replication: HbO (red), tHb (green) and Hb (blue) changes at s4_d2, s4_d5 (top, left to right), s3_d2, s4_d4, and s6_d5 (bottom, left to right) corresponding to 5% above O_1, 5% above O_2, O_1, Oz and O_2. Data are presented as mean and standard error of the mean (SEM). Vertical line separates two conditions: stimulus presentation (task) and baseline (rest).

Table 2. Visual stimulation replication: Mean across subjects of concentration changes (μM) of the last 15 seconds of the Task condition for set of locations s3_d2, s4_d4, s6_d5, s4_d2 and s4_d5 corresponding to O_1, Oz, O_2, 5% above O_1 and 5% above O_2 (baseline correction applied). SEM stands for standard error of the mean.

		HbO (Mean ± SEM)	Hb (Mean ± SEM)	tHb (Mean ± SEM)
O_1	s3_d2	0.24 ± 0.07	-0.09 ± 0.01	0.15 ± 0.07
Oz	s4_d4	0.10 ± 0.04	-0.05 ± 0.02	0.05 ± 0.06
O_2	s6_d5	0.34 ± 0.07	-0.09 ± 0.01	0.25 ± 0.06
5% above O_1	s4_d2	0.16 ± 0.07	-0.04 ± 0.04	0.12 ± 0.08

5% above O_2	s4_d5	0.11 ± 0.07	-0.02 ± 0.01	0.09 ± 0.08

The statistical replication results showed that:

a) The effect of condition was significant for HbO, Hb and tHb changes. Same as that in Wijeakumar et al. (2012).

   The effect of location was **not** significant for HbO, Hb and tHb changes, while it was significant for HbO, Hb and tHb in Wijeakumar et al. (2012).

   The effect of the interaction between location and condition was significant for HbO and tHb changes in both sets of locations, and was significant for Hb in one set. This effect was significant for tHb, HbO and Hb in Wijeakumar et al. (2012).

b) HbO, Hb and tHb changes at Oz were not significantly different from that at O_1 and O_2. They were significant in Wijeakumar et al. (2012).

c) HbO, Hb and tHb changes at Oz were not significantly different from that at 5% above O_1 and 5% above O_2. Same as that in Wijeakumar et al. (2012).

d) The increase observed in HbO changes was greater than the decrease observed in Hb changes. Same as that in Wijeakumar et al. (2012).

e) The trend of tHb was similar to that of HbO. Same as that in Wijeakumar et al. (2012).

In visualization, the trends of Hb, HbO and tHb changes across trials and across subjects were similar to that in Wijeakumar et al. (2012).

Table 3. Visual stimulation replication: results of RM ANOVA for set of locations s3_d4, s4_d4, s6_d4, s3_d2 and s6_d5 corresponding to O_1, Oz, O_2, 5% above O_1 and 5% above O_2. Bold p values are the corrected p values based on Greenhouse-Geisser correction, degrees of freedom are not corrected in that case, ns. stands for not significant.

	Effect		**Greenhouse-Geisser epsilon**	**Pairwise comparison**
Hb	Condition	F(1,10)=43.101, *p*<.001*		*p*<.001*
	Location	F(4,40)=0.178, ***p*=0.88**	0.6244431	-
	Interaction	F(4,40)=2.586, ***p*=0.093**	0.5557814	-
HbO	Condition	F(1,10)=12.948, *p*=0.005*		*p*<.001*
	Location	F(4,40)=0.381, ***p*=0.691**	0.5079983	-
	Interaction	F(4,40)=7.479, *p*<.001*		two conditions at s6_d5, *p*=0.041*
tHb	Condition	F(1,10)=5.844, *p*=0.036*		*p*<.001*
	Location	F(4,40)=0.276, ***p*=0.762**	0.5005487	-
	Interaction	F(4,40)=4.775, *p*=0.003*		ns.

* significant, - not applicable

Table 4. Replication: results of RM ANOVA for set of locations s3_d2, s4_d4, s6_d5, s4_d2 and s4_d5 corresponding to O_1, Oz, O_2, 5% above O_1 and 5% above O_2. Bold p values are the corrected p values based on Greenhouse-Geisser correction, degrees of freedom are not corrected in that case, ns. stands for not significant.

Effect	**Greenhouse-Geisser epsilon**	**Pairwise comparison**

Hb	Condition	F(1,10)=25.566, *p*<.001*		*p*<.001*
	Location	F(4,40)=1.221, ***p*=0.312**	0.4121358	-
	Interaction	F(4,40)=6.17, *p*=0.001*		two conditions at s3_d2, *p*=0.028* two conditions at s6_d5, *p*=0.046*
HbO	Condition	F(1,10)=15.803, *p*=0.003*		*p*<.001*
	Location	F(4,40)=0.402, *p*=0.806		-
	Interaction	F(4,40)=6.558, *p*<.001*		two conditions at s6_d5, *p*=0.041*
tHb	Condition	F(1,10)=7.504, *p*=0.021*		*p*<.001*
	Location	F(4,40)=0.75, *p*=0.564		-
	Interaction	F(4,40)=3.95, *p*=0.009*		ns.

* significant, - not applicable

3.1.2.4. Discussion and conclusion

In general the results in Wijeakumar et al. (2012) were reproducible. The trends of tHb, HbO and Hb changes in the replication were similar to that in Wijeakumar et al. (2012). The means of tHb, HbO and Hb changes across subjects in stimulus presentation were significantly different from that in control condition in the five locations O_1, Oz, O_2, 5% above O_1 and 5% above O_2, which was the same as that in Wijeakumar et al. (2012). However statistical results related to location were not reproducible due to overlapping in value ranges of concentration changes between locations in the replicated data. It should be noted that the device used in Wijeakumar et al. (2012) was Frequency Domain Multi-Distance (FDMD), which returned absolute concentration, and the technique supported to remove noise from the superficial layers better (Hueber et al., 2001; Scholkmann & Wolf, 2012) while that in the replication were not absolute changes and still contained noise from the superficial layers.

3.2. Finger tapping experiment

3.2.1. Literature

3.2.1.1. Method

In 2004, Kuboyama et al. studied the effect of finger tapping at different rates on the left primary motor cortex. Nine right-handed and healthy subjects, mean age 27.2 (SD=4.8), 2 females, were asked to press a button with their right index finger at three different frequencies which were determined based on individual's maximum effort (ME), 50% ME and 25% ME. The 25% and 50% ME conditions were randomized and regulated by the sound. A trial included 20-seconds tapping at a specific frequency and 40-seconds rest period. Each trial corresponding to one tapping frequency was repeated 3 times. In total each subject performed 9 trials. The measurements were conducted with 24-channel 2-wavelength Hitachi ETG-100, but the authors selected only one data channel close to the EEG position C3 for the analysis. The sampling rate was 10 Hz. The data returned by their NIRS device were already the changes of hemoglobin concentration (Kawaguchi et al., 2001).

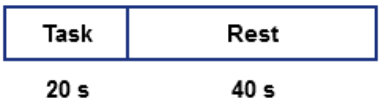


Figure 9. A trial in the Finger tapping experiment in Kuboyama et al. (2004).

It was not clear in the paper how the preprocessing was performed. Correspondence with one of the authors revealed that the rest period preceding each task was used as the baseline for that task, baseline correction was performed and averages of 10-samples blocks (1 second) were used in visualization. RM ANOVA and post-hoc tests with Bonferroni method were used for statistical test. There were no details about the effects which had been tested in the RM ANOVA test.

3.2.1.2. Results

The results in Kuboyama et al. (2004) showed that:

a) Only in the ME condition the increases of Hb, HbO and tHb changes from their baseline are significantly different.

b) The increases of Hb, HbO and tHb changes from their baseline in the ME condition were significantly higher than that in the 50% ME and the 25% ME condition.

c) The increases of Hb, HbO and tHb changes from their baseline in the 50% ME condition were not significantly different from that in the 25% ME condition.

3.2.2. Experiment at Philips Research

3.2.2.1. Procedure and materials

The Finger tapping experiment was performed at Philips Research to replicate the experiment in Kuboyama et al. (2004). Five right-handed subjects, mean age 25.25 (SD=2.99), 2 females, 2 males and 1 unknown, participated in the experiment after signing a written informed consent. Subjects were asked to tap their right index finger on the desk to the sound of the metronome at a specified frequency. Three frequencies, 2Hz, 4Hz and as fast as possible (i.e. maximum effort) were performed. Each tapping period lasted 20 seconds and followed by a 40 seconds of rest (doing nothing). Each trial which contains a tapping period at a specific frequency and a rest period was repeated three times. Before the first trial began, subjects were asked to sit still for 60 seconds to measure their baseline. In total, each subject performed 9 trials at 3 frequencies 2Hz, 4Hz and ME, by order (Figure 10).

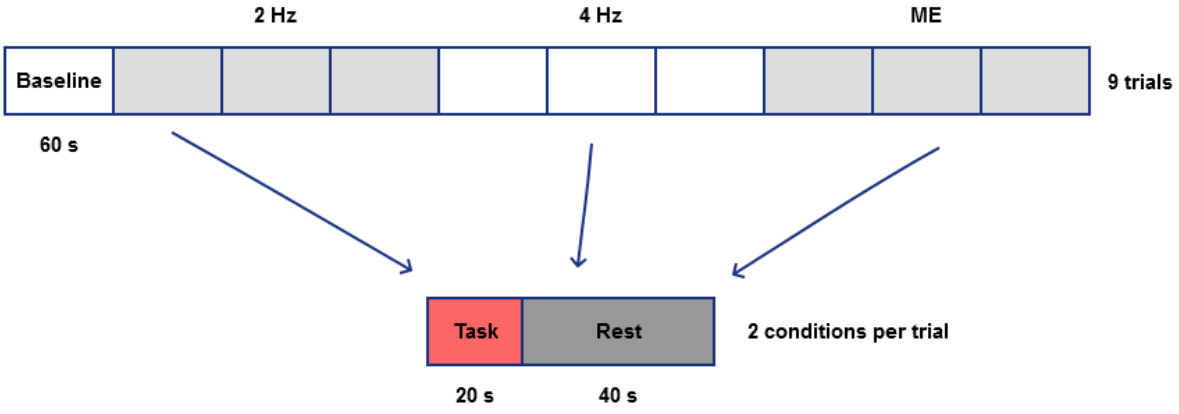


Figure 10. Finger tapping experiment performed at Philips Research.

The measurements were conducted with CW NIRSport device from NIRx Medical Technologies LLC., which measured in two wavelengths, 760 nm and 850 nm. The sampling rate was set at 7.8125 Hz with 20 channels setting. The inter-optode distance was 30 mm. The output data were light intensities received

at detector optodes which were saved in ASCII format in two files corresponding to two wavelengths. Additional files were also created which contained other information necessary for the preprocessing and analyzing, see Appendix A. The positions of probes and channels are shown in Figure 11, the presentation of channels is illustrated in Figure 12.

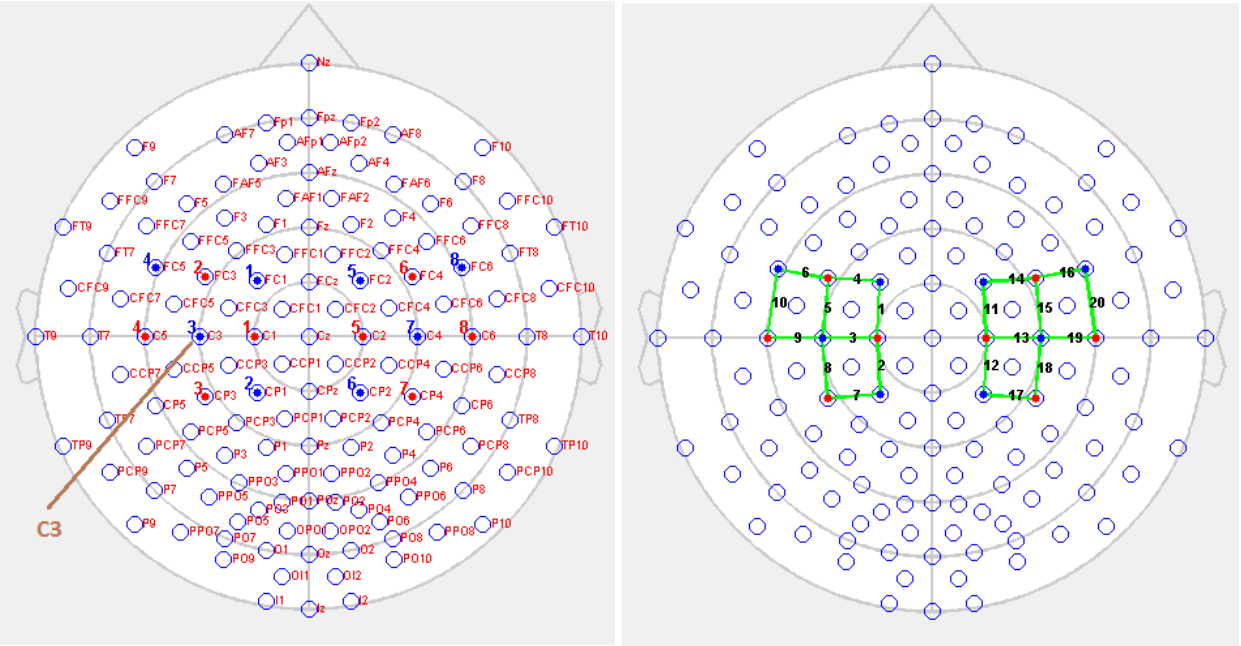


Figure 11. Topology of probes and channels in the finger tapping experiment conducted at Philips Research. Viewpoint is from the top of the head with the nose side at the top of the figures. Red and blue points with numbers present source and detector optodes. Left: Position of probes regarding to the EEG locations. Right: Topology of the active channels. Green lines with number present the channels. The software nirsLAB (NIRx Medical Technologies, LLC, version 2016.01) was used to generate these figures.

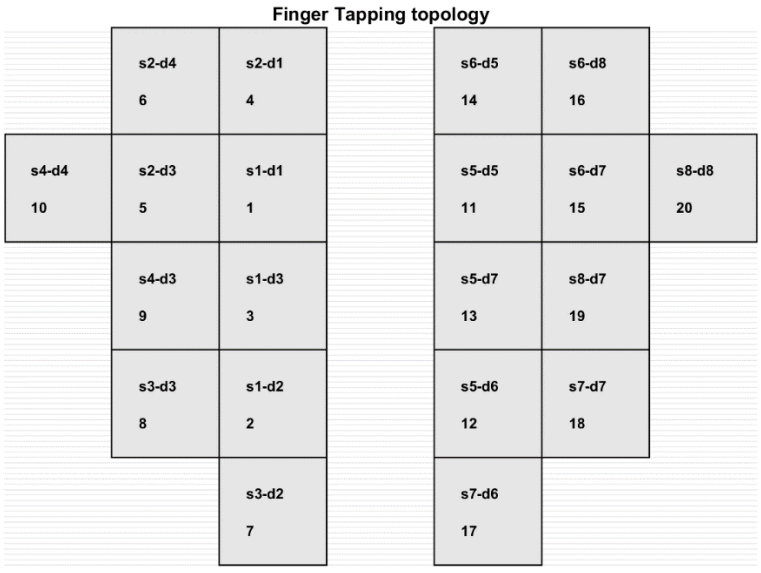


Figure 12. Presentation of channels' location in Finger Tapping experiment, each cell shows the name and index of the corresponding channel. Viewpoint is from the top of the head with the nose side at the top of the figures.

Our research question

Are there significant differences in Hb, HbO and tHb changes between the four conditions: baseline, tapping frequencies 2Hz, 4Hz and ME?

3.2.2.2. Data analysis

The data were analyzed using Matlab (The MathWorks Inc., MA, USA, version R2015b) software and the ezANOVA function of the ez package (Lawrence, 2015, version 4.3) in R (R Core Team, 2016) software.

Concentration changes were calculated from the raw intensities data for each subject (Appendix B). Following the preprocessing in the literature, the last 20 second of the rest condition preceding a tapping condition was defined as the baseline for that tapping condition. Thus each tapping condition had its own baseline.

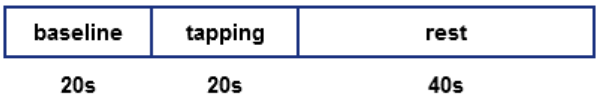


Figure 13. Finger tapping experiment: Three segments in a trial.

For visualization, each tapping trial was illustrated with a rest segment preceding the tapping condition (baseline), a tapping segment and a rest segment following the tapping condition to show the trend of concentration changes before, during and after the tapping continuously (Figure 13). Similar to the method in literature to reduce noise, samples in each trial were converted to second unit by grouping in groups of 8 samples (approx. 1 second) counting backward from the end of the segment which was applied to baseline and tapping segment, and counting forward from the beginning of the segment, which was applied to the rest segment (Figure 14). The first or last group in each segment might contain less than 8 samples as the result of the non discrete sampling rate (7.8125 Hz). The mean of each group was taken as value for the corresponding time point in second unit. As the differences in the number of samples between conditions were small, after converting to second unit, the lengths of baseline, tapping and rest segments were 20 seconds, 20 seconds and 40 seconds respectively, for all trials, all subjects.

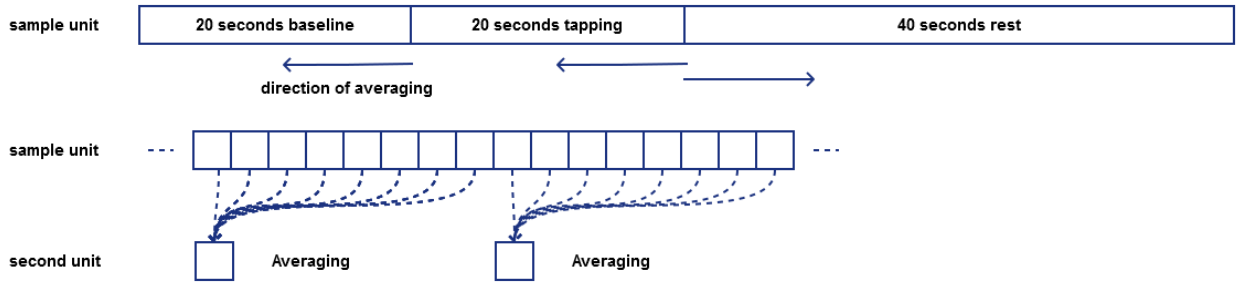


Figure 14. Conversion to second unit in Finger tapping experiment.

Baseline correction was subsequently applied for each trial by subtracting the mean of baseline segment from all time point in the trial, per type of concentration changes. Subsequently, average of 3 trials per time point for each tapping frequency of all subjects was calculated which was used to illustrate the trend of the concentration changes per tapping frequency (2Hz, 4Hz and ME).

For statistical tests, data were converted to second unit as described above however only 20 seconds of the tapping segment in each trial were taken into account after subtracting the mean of baseline from

each time point in the tapping segment. RM ANOVA was used to test the effect of frequency on each type of concentration changes. The data for RM ANOVA test were the mean of tHb, HbO or Hb changes across trials in each tapping frequency per subject, per channel, 3 frequencies x 5 subjects. Post-hoc *t*-test with Bonferroni adjustment on the same data was used for pairwise comparison when frequency was significant in the result of RM ANOVA test. For each channel, there were 3 RM ANOVA tests and up to 3 post-hoc *t*-tests, corresponding to 3 types of concentration changes and 1 effect. Finally, *t*-tests were used to test the significant differences of tHb, HbO or Hb changes at each tapping frequencies (baseline corrected already) and 0 (baseline). The data for each of these t-tests were the mean of tHb, HbO or Hb changes across trials in one tapping frequency, per subject, per channel, 1 frequency x 5 subjects. There were 9 *t*-tests per channel corresponding to 3 frequencies and 3 types of concentration changes. All statistical inferences were based on an adjusted alpha level of 5%.

Based on the probes setting of the experiment (Figure 11), for the replicated data, four channels around C3 were considered which were s4-d3, s2-d3, s1-d3 and s3-d3. In total there were 12 RM ANOVA tests, up to 12 post-hoc *t*-tests (4 channels x 3 types of concentration changes) and 36 *t*-tests (4 channels x 3 frequencies x 3 types of concentration changes).

3.2.2.3. Results

Table 5. Finger tapping replication: Mean and standard deviation across subjects of concentration changes during task condition (baseline correction applied) at channel s4-d3.

	2 Hz	4 Hz	ME
tHb	0.13 ± 0.37	0 ± 0.42	0.31 ± 0.42
HbO	0.19 ± 0.48	0.13 ± 0.46	0.43 ± 0.49
Hb	-0.06 ± 0.18	-0.13 ± 0.13	-0.12 ± 0.18

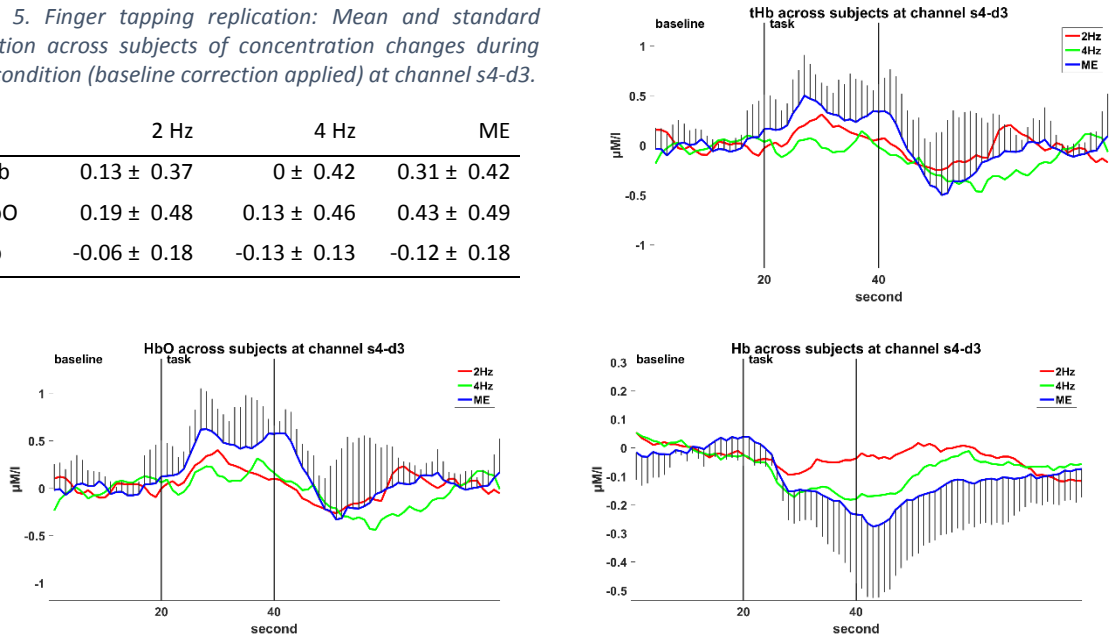


Figure 15. Finger tapping replication: mean and standard deviation across all subjects of tHb (top), HbO (bottom left) and Hb (bottom right) changes in 2 Hz (red), 4 Hz (green) and ME (blue) at channel s4-d3. Vertical lines indicate onset and end of tapping condition. Standard deviation was shown to replicate the figures in Kuboyama et al. (2004) and that applied for ME frequency only to have a clear view.

Table 6. Finger tapping replication: Mean and standard deviation across subjects of concentration changes during task condition (baseline correction applied) at channel s2-d3.

	2 Hz	4 Hz	ME
tHb	0.11 ± 0.35	-0.02 ± 0.40	0.31 ± 0.41
HbO	0.19 ± 0.39	0.15 ± 0.46	0.53 ± 0.55
Hb	-0.08 ± 0.20	-0.17 ± 0.17	-0.22 ± 0.24

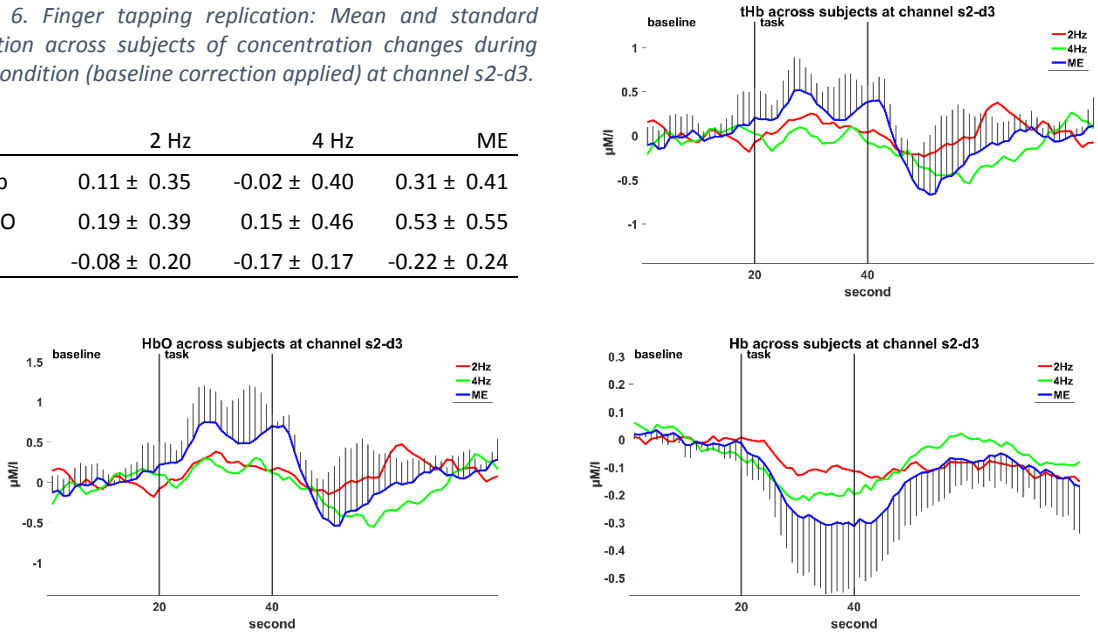


Figure 16. Finger tapping replication: mean and standard deviation across all subjects of tHb (top), HbO (bottom left) and Hb (bottom right) changes in 2 Hz (red), 4 Hz (green) and ME (blue) at channel s2-d3. Vertical lines indicate onset and end of tapping condition. Standard deviation was shown to replicate the figures in Kuboyama et al. (2004) and that applied for ME frequency only to have a clear view.

Table 7. Finger tapping replication: Mean and standard deviation across subjects of concentration changes during task condition (baseline correction applied) at channel s1-d3.

	2Hz	4Hz	ME
tHb	0.11 ± 0.47	-0.08 ± 0.70	0.57 ± 0.52
HbO	0.18 ± 0.60	0.08 ± 0.74	0.82 ± 0.63
Hb	-0.07 ± 0.29	-0.16 ± 0.25	-0.25 ± 0.28

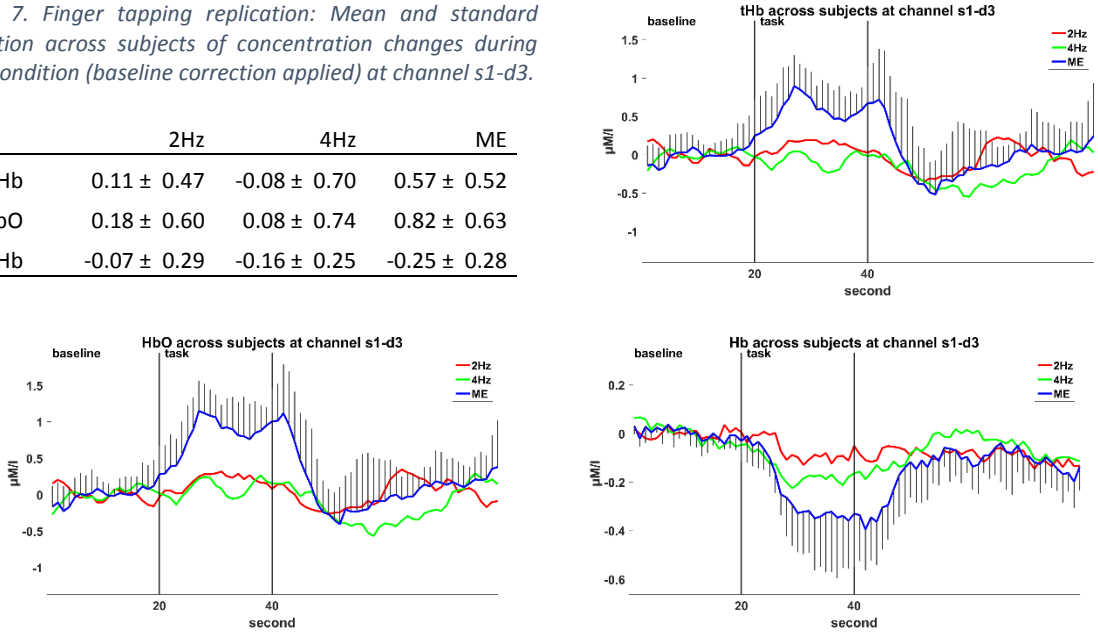


Figure 17. Finger tapping replication: mean and standard deviation across all subjects of tHb (top), HbO (bottom left) and Hb (bottom right) changes in 2 Hz (red), 4 Hz (green) and ME (blue) at channel s1-d3. Vertical lines indicate onset and end of tapping condition. Standard deviation was shown to replicate the figures in Kuboyama et al. (2004) and for ME frequency only to have a clear view.

Table 8. Finger tapping replication: Mean and standard deviation across subjects of concentration changes during task condition (baseline correction applied) at channel s3-d3.

	2 Hz	4 Hz	ME
tHb	0.03 ± 0.47	-0.06 ± 0.47	0.25 ± 0.60
HbO	0.10 ± 0.57	0.03 ± 0.59	0.51 ± 0.77
Hb	-0.06 ± 0.33	-0.09 ± 0.25	-0.26 ± 0.28

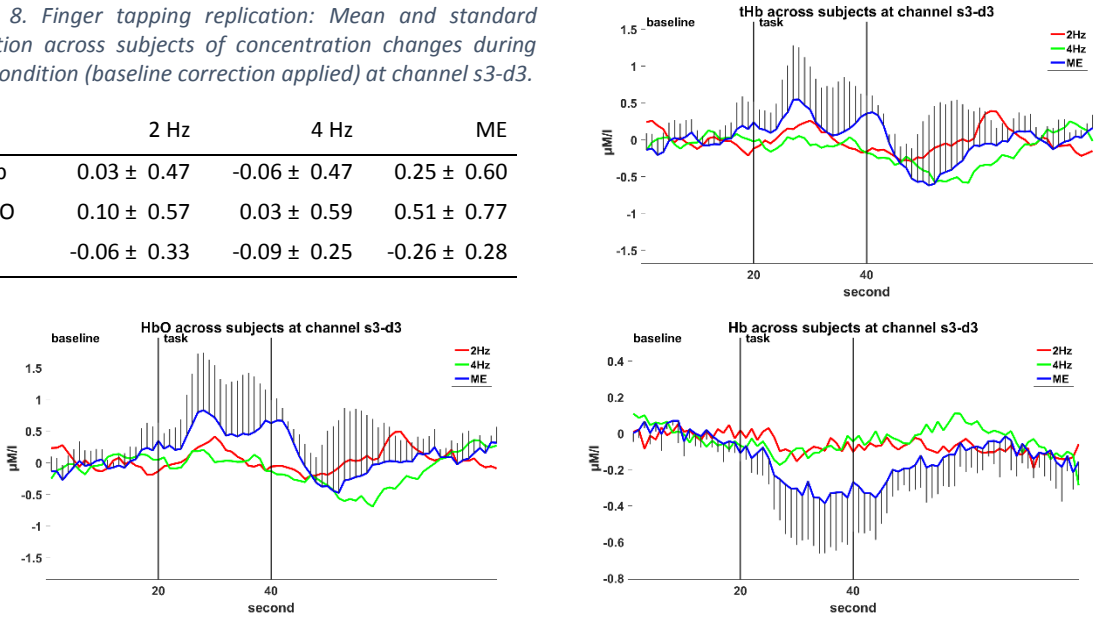


Figure 18. Finger tapping replication: mean and standard deviation across all subjects of tHb (top), HbO (bottom left) and Hb (bottom right) changes in 2 Hz (red), 4 Hz (green) and ME (blue) at channel s3-d3. Vertical lines indicate onset and end of tapping condition. Standard deviation was shown to replicate the figures in Kuboyama et al. (2004) and for ME frequency only to have a clear view.

The replication results showed that in the four considering channels:

a) The increases of HbO and tHb changes and decrease of Hb changes from baseline at ME were significant at channel s4_d3 (9), s2_d3 (5) and s1_d3 (3), same as that in the literature.

Furthermore, the decrease of Hb changes from baseline at 4 Hz was significant at channel s4_d3 (9) and s2_d3 (5), and the increase of HbO changes from baseline at 2 Hz was significant at channel s2_d3 (5), which did not occur in the result of the literature.

b) At ME the increase of tHb, HbO changes from baseline was significantly higher than that at 4 Hz and 2 Hz at channel s1_d3 (3), same as that in the literature. Besides, the increase of tHb at ME was also significantly higher than that at 4 Hz at channel s4_d3 (9).

No decrease of Hb changes from baseline at ME was significantly larger than that at 4 Hz and 2 Hz at four channels, which is different from the result in the literature.

c) The decrease of Hb, and increases of HbO and tHb changes from their baseline were not significantly different between 2Hz and 4Hz, same as that in the literature.

Among the four channels, the results at channel s1-d3 (3) were closest to that in the literature although the decrease of Hb changes from baseline at ME significantly at this channel was not significantly different from that at 4 Hz and 2 Hz.

In visualization, HbO and tHb changes at ME during the task increased, similar to that in the literature, however Hb changes at ME decreased while in the literature they increased.

Table 9. Finger tapping replication: statistical results for s4-d3, s2-d3, s1-d3 and s3-d3. Column 3, 4 and 5 contain p-values of t-tests between each tapping frequency and the corresponding baseline. Column RM ANOVA contains the result of RM ANOVA tests for Frequency effect. Column 7, 8 and 9 contain p-values of post-hoc pairwise t-test between frequencies.

		t-test			RM ANOVA	post-hoc t-test		
		***p*-value 2 Hz baseline**	***p*-value 4 Hz baseline**	***p*-value ME baseline**	**Frequency effect**	***p*-value 4Hz vs. 2Hz**	***p*-value ME vs. 2Hz**	***p*-value ME vs. 4Hz**
s4_d3	tHb	0.209	0.985	0.011*	F(2,8)=6.044, p=0.025*	0.575	0.524	0.039*
	HbO	0.257	0.184	0.009*	F(2,8)=2.812, p=0.119	-	-	-
	Hb	0.419	0.045*	0.037*	F(2,8)=0.623, p=0.56	-	-	-
s2_d3	tHb	0.114	0.746	0.019*	F(2,8)=5.21, p=0.036*	0.431	0.602	0.089
	HbO	0.039*	0.125	0.020*	F(2,8)=4.983, p=0.039*	1	0.336	0.19
	Hb	0.222	0.015*	0.051*	F(2,8)=1.592, p=0.262	-	-	-
s1_d3	tHb	0.192	0.42	0.010*	F(2,8)=16.561, p=0.001*	0.353	0.011*	0.046*
	HbO	0.207	0.589	0.003*	F(2,8)=20.992, p=0.001*	1	0.017*	0.024*
	Hb	0.497	0.083	0.022*	F(2,8)=1.331, p=0.317	-	-	-
s3_d3	tHb	0.612	0.504	0.297	F(2,8)=1.323, p=0.319	-	-	-
	HbO	0.515	0.796	0.144	F(2,8)=1.723, p=0.239	-	-	-
	Hb	0.646	0.202	0.058	F(2,8)=1.06, p=0.391	-	-	-

* significant, - not applicable

3.2.2.4. Discussion and conclusion

The results of tHb and HbO changes in the replicated data matched with that in Kuboyama et al. (2004) however that of Hb changes were not. In the replication, Hb changes decreased during the tapping condition, which matched with the typical neuronal hemodynamic response shown in figure 5 in Scholkmann et al. (2014) and the results in Gagnon et al. (2012) and Huppert, Hoge, Diamond, Franceschini, and Boas (2006) while Hb changes increased in the results of Kuboyama et al. (2004).

It should be noted that in the replicated data, the concentration changes of subjects during the last 20 seconds of the rest condition in each trial were not stable. Furthermore, the traces of HbO changes of some subjects at 2 Hz were at higher amplitude than that at 4 Hz, that explained the higher mean of HbO changes at 2Hz comparing to that that at 4Hz in the four considered channels (Table 5 to Table 8), which suggested some subjects might have tapped their finger in a wrong frequency. The traces of Hb changes of these same subjects were also at lower amplitude at 2Hz than at 4Hz, however the means of Hb changes across subjects at 2Hz and 4 Hz did not reflect that since Hb changes at 2Hz of one subject increased during the tapping condition, which canceled out the differences.

3.3. Continuous performance test (CPT) experiment

3.3.1. Literature

3.3.1.1. Method

Fallgatter and Strik investigated in 1997 the sensitivity of NIRS in concentration changes during cognitive activity in prefrontal cortex. Ten subjects, mean age 30 (27 – 33 years), 5 females, performed a modified version of the Continuous Performance Test (press a key when letter X appears after letter A and not in others). Subjects were asked to close eyes to measure the baseline (pre-task), after that, without any

instruction, subjects performed the task (CPT task). A post-stimulus segment with eyes closed ended the experiment (post-task). The length of each condition, pre-task, CPT task and post task, was normalized to 28, 138 and 14 seconds for all subjects, respectively.

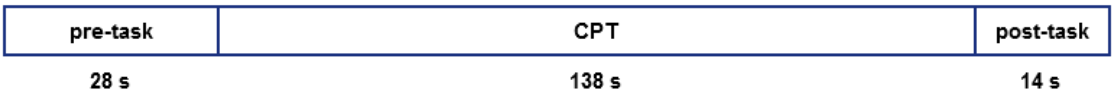


Figure 19. CPT experiment in Fallgatter and Strik (1997), lengths of segments have been normalized for all subjects.

The measurements were conducted with two 1-channel Critikon 2020 Cerebral Redox Monitor systems which provided the concentration of HbO and Hb in $\mu M/l$ as output. Each system had one sensor attached on one side of the subject's head with positions between Fp1 and F3 on the left hemisphere and between Fp2 and F4 on the right hemisphere (based on the International 10-20 system for EEG electrode placement). The sampling rate was 1Hz.

For visualization, baseline correction was applied per type of concentration, per channel and per subject, wherein the mean of pre-task condition was subtracted from all time points in the three conditions. Subsequently, the concentration changes were averaged across subjects per time point, and then smoothed using 31 points.

For statistical tests, RM ANOVA was used to test the effect of hemisphere (left and right), condition (pre-task, CPT and post-task) and the interaction between hemisphere and condition on HbO and Hb changes. The data for RM ANOVA were the average of concentration changes in each condition per subject, 2 hemispheres (left and right) x 3 conditions (pre-task, CPT task and post-task) x 10 subjects. Post-hoc *t*-tests were used to perform pairwise comparisons.

3.3.1.2. Results

The results in Fallgatter and Strik (1997) showed that:

a) During CPT condition, the Hb changes in the left hemisphere were significantly higher than that in the right hemisphere.
b) During CPT condition, the HbO changes were not significantly different between hemispheres.
c) Hb changes in the right hemisphere differed significantly between conditions.

They interpreted their results as the right frontal lobe showing significantly more brain activation during the CPT task, compared to the left frontal lobe.

3.3.2. Experiment at Philips Research

3.3.2.1. Procedure and materials

The CPT experiment was conducted at Philips Research in a combined experiment wherein the CPT and Verbal-fluency task experiment (which is described in the next section) both took place. Eleven subjects participated in the experiment after signing a written consent. Due to experiment error, data of one subject were excluded from the analysis. In total there were data of 10 subjects, mean age 26 (SD=5.16), 3 females. Subjects were asked to do the CPT and VFT tasks in counterbalance; five of them performed the CPT first and then VFT, the others performed VFT and then CPT.

In the CPT experiment, subjects were asked to keep their eyes open for 120 seconds and then close eyes for 28 seconds, after that perform the CPT for 150 seconds and finally close eyes for 120 seconds. There was always a 15-second instruction period between the conditions mentioned above wherein a sound reminded the subject to open eyes after an Eyes Closed condition or the subject read instruction on the screen. The first and last Eyes Closed condition were used as the corresponding pre-task and post-task condition in the CPT experiment in Fallgatter and Strik (1997).

Figure 20. CPT experiment performed at Philips Research. I stands for Instruction.

The measurements were conducted with CW NIRSport device from NIRx Medical Technologies LLC., which measured in two wavelengths, 760 nm and 850 nm. The sampling rate was set at 7.8125 Hz with 20 channels setting. The inter-optode distance was 30 mm. The output data were light intensities received at detector optodes which were saved in ASCII format in two files corresponding to two wavelengths. Additional files were also created which contained other information necessary for the preprocessing and analyzing, see Appendix A. The positions of probes and channels are shown in Figure 21, the presentation of channels is illustrated in Figure 22.

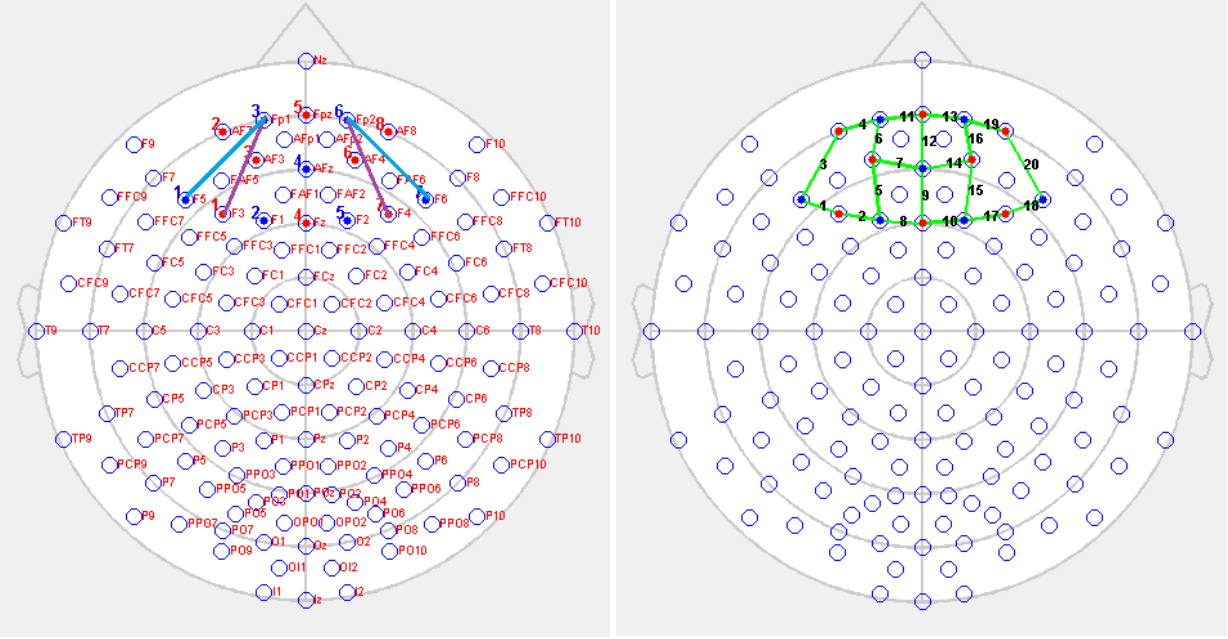


Figure 21. Topology of probes and channels in the CPT and VFT experiment conducted at Philips Research. Viewpoint is from the top of the head with the nose at the top of the figures. Red and blue points with numbers present source and detector optodes. Left: Position of probes regarding to the EEG locations. Purple lines present the positions of the channels in the CPT experiment in Fallgatter and Strik (1997). Blue lines present the positions of the channels in the VFT experiment in Herrmann et al. (2003). Right: Topology of the active channels. Green lines with number present the channels. The software nirsLAB (NIRx Medical Technologies, LLC, version 2016.01) was used to generate these figures.

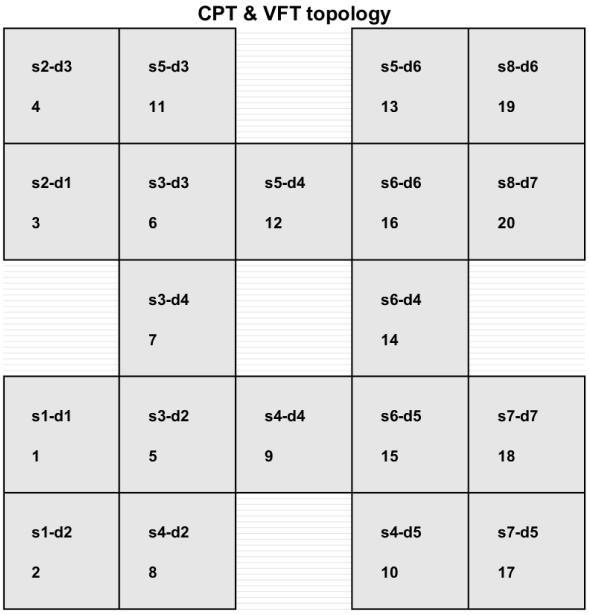


Figure 22. Presentation of channels' location in CPT and VFT experiment, each cell shows the name and index of the corresponding channel. Viewpoint is from the top of the head with the nose side at the top of the figures.

Our research question

Are there significant differences in Hb and HbO changes between three conditions pre-task, CPT and post-task, including the comparison between left and right hemisphere locations?

3.3.2.2. Data analysis

The data were analyzed using Matlab (The MathWorks Inc., MA, USA, version R2015b) software and the ezANOVA function of the ez package (Lawrence, 2015, version 4.3) in R (R Core Team, 2016) software.

Concentration changes were calculated from the raw intensities data of each subject, which were then split to two parts corresponding to two experiments CPT and VFT.

For the replication of CPT experiment, only data from the first Eyes Closed condition to the end of the last Eyes Closed condition were taken into account, including the two Instruction periods in between. These instruction periods were kept in the visualization to show the continuation of concentration changes. The first and last Eyes Closed, and CPT segments were chosen as pre-task, post-task and CPT condition, respectively.

In Fallgatter and Strik (1997), the length of the post-task condition after normalization was 14 seconds (there was no information about the length of this condition before normalization), while in the replicated data, the length of this condition was 120 seconds and was preceded by a 15-second Instruction condition. We chose the first 14 seconds of the 120-second post-task condition to represent the post-task condition for several reasons. Firstly, there was no instruction period between the CPT and the post-task condition reported in Fallgatter and Strik (1997), therefore due to the latency of hemodynamics response, it was expected that the average concentration changes of the first 14 seconds of the post-task condition which had been used in Fallgatter and Strik (1997) would be higher than that in the later period of the post-task

condition. With this choice for the replicated data, the differences between the CPT and the post-task condition would be clearer in the replication's result. Secondly, the 15-second instruction condition between the CPT and post-task condition was not chosen to represent the post-task condition because the reading during this condition may cause activation in the brain which would lead to unexpected results. Finally, 14 seconds were chosen to be the same as that in the literature.

For visualization, the lengths of conditions across subjects were normalized, wherein samples at the onset of pre-task, Instruction and CPT conditions whose condition's length was longer than the minimum condition's length were excluded. The length normalization for post-task condition was not necessary since the chosen length, 14 seconds, was much smaller than the measured length of this condition (approx. 120 seconds). After normalization, the lengths of pre-task, the first instruction, CPT, and the second instruction condition were 28.0, 14.8, 149.9 and 14.8 seconds, respectively. Baseline correction was performed where the mean of pre-task condition was subtracted from each time point from all conditions per subject. After that, average changes per time point across all subjects were calculated and then smoothed by a symmetric average moving filter with filter length 31 (the length is same as that in the paper).

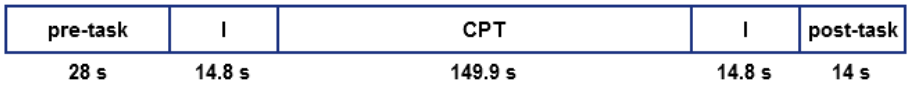


Figure 23. CPT experiment at Philips Research. I stands for Instruction. Lengths of segments have been normalized across subjects.

For statistical tests, only data of pre-task, CPT and post-task condition after normalization were taken into account. The normalization data did not alter the result of the statistical tests since up to 2 samples at the onset of each condition were excluded. RM ANOVA was used to test the significance of condition (pre-task, CPT and post-task), hemisphere (left and right) and the interaction between condition and hemisphere on HbO or Hb with data being the mean of HbO or Hb changes of each condition per hemisphere channel and per subject (3 conditions x 2 hemispheres x 10 subjects). Post-hoc *t*-test with Bonferroni adjustment on the same data was used for pairwise comparisons when any of the effects was significant in the result of RM ANOVA test. For each pair of channels representative for left and right hemispheres, there was 2 RM ANOVA tests and up to 6 post-hoc t-tests, corresponding to 2 types of concentration changes, HbO and Hb, and 3 effects. All statistical inferences were based on an adjusted alpha level of 5%.

In Fallgatter and Strik (1997), data were measured with 2-channels system with positions corresponding to those between positions Fp1 and F3 in the left hemisphere and between Fp2 and F4 in the right hemisphere. For the replication we considered either s2-d3 (channel 4) and s8-d6 (channel 19) or s3-d3 (channel 6) and s6-d6 (channel 16) to represent the channel on the left and right hemispheres (Figure 21). In total there were 4 RM ANOVA tests and up to 12 post-hoc t-tests (2 pairs of channels, 2 types of concentration changes and 3 effects).

3.3.2.3. Results

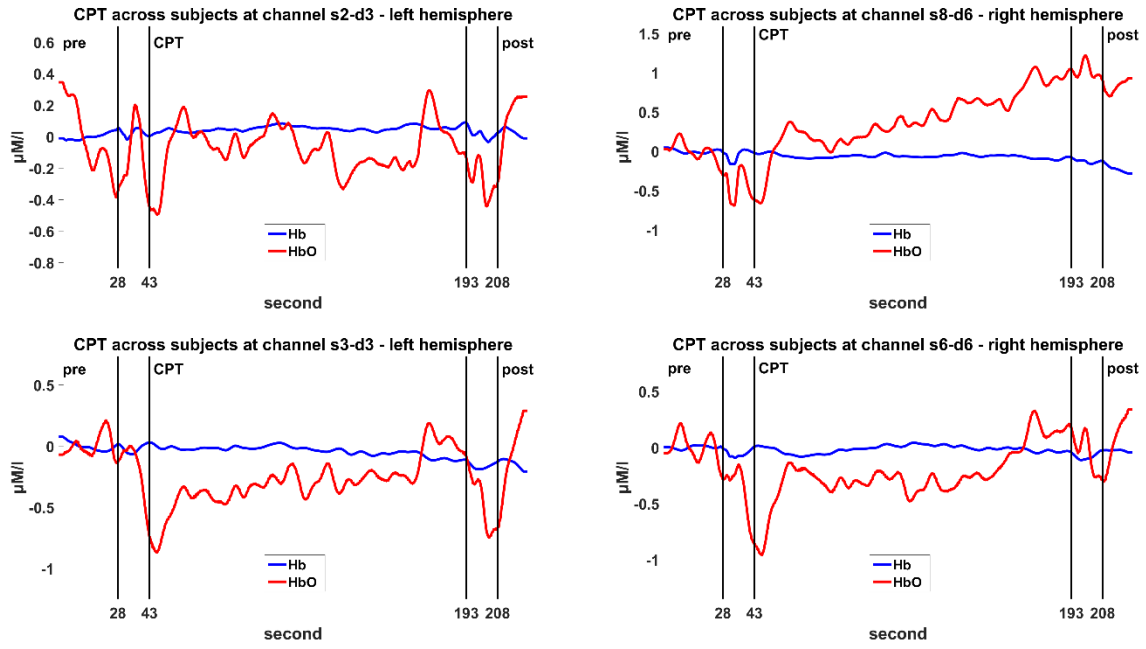


Figure 24. CPT replication: Hb (blue) and HbO (red) changes in left and right hemispheres in three conditions pre-task, CPT and post-task by order in CPT experiment. Vertical line which follows or precedes each condition indicates the end or onset of that condition, the small segments between these lines are instruction periods. Channel s2_d3 (top left) and channel s3_d3 (bottom left) represent left hemisphere. Channel s8_d6 (top right) and channel s6_d6 (bottom right) represent right hemisphere. Error bars were not shown to replicate the figures in Fallgatter and Strik (1997).

Table 10. CPT replication: Mean and standard deviation across subjects of Hb and HbO (µM) of CPT and post condition in four channels (baseline correction applied).

channel		CPT	Post	channel		CPT	Post
s2_d3	Hb	0.05 ± 0.39	0.02 ± 0.39	s8_d6	Hb	-0.07 ± 0.18	-0.23 ± 0.36
	HbO	-0.07 ± 1.11	0.14 ± 0.53		HbO	0.40 ± 0.65	0.83 ± 1.30
s3_d3	Hb	-0.04 ± 0.35	-0.15 ± 0.39	s6_d6	Hb	-0.02 ± 0.19	-0.03 ± 0.25
	HbO	-0.29 ± 0.55	-0.07 ± 0.71		HbO	-0.24 ± 0.75	0.09 ± 0.88

The replication results showed that with the pair of channel s3_d3 and s6-d6 or s2_d3 and s8-d6 representing the left and right hemispheres, respectively:

a) During CPT task, Hb changes were **not** significantly different between hemispheres, while in the result of Fallgatter and Strik (1997), Hb changes in the left hemisphere were significant higher than that in the right hemisphere.

b) During CPT task, HbO changes were not significantly different between hemispheres, same as that in Fallgatter and Strik (1997) however the average HbO changes in the replicated data did not increase during CPT task as they did in Fallgatter and Strik (1997).

c) Hb changes in the right hemisphere did not differ significantly between conditions, while they did in the result of Fallgatter and Strik (1997).

Table 11. CPT replication: statistical results for four pairs of channels representing left and right hemispheres. Column 4 and 5 contain the results of RM ANOVA for the corresponding effect. Column post-hoc t-test contains the results of post-hoc t-tests for the corresponding effect.

Hemisphere			**RM ANOVA**		**post-hoc t-test**
Left	**Right**		**Effect**		***p*-value**
s3_d3	s6_d6	Hb	Condition	F(2,18)=0.65, *p*=0.534	-
			Hemisphere	F(1,9)=0.607, *p*=0.456	-
			Interaction	F(2,18)=1.279, *p*=0.302	-
		HbO	Condition	F(2,18)=0.917, *p*=0.418	-
			Hemisphere	F(1,9)=0.703, *p*=0.423	-
			Interaction	F(2,18)=0.373, *p*=0.694	-
s2_d3	s8_d6	Hb	Condition	F(2,18)=0.804, *p*=0.463	-
			Hemisphere	F(1,9)=0.045, *p*=0.837	-
			Interaction	F(2,18)=3.039, *p*=0.073	-
		HbO	Condition	F(2,18)=1.917, *p*=0.176	-
			Hemisphere	F(1,9)=0.135, *p*=0.722	-
			Interaction	F(2,18)=1.301, *p*=0.297	-

- not applicable

3.3.2.4. Discussion and conclusion

The replicated results were not the same as that in Fallgatter and Strik (1997) due to the differences in hemodynamic response between subjects. In the replicated data, the trend of Hb changes of all subjects in the CPT condition were not consistent. In some subjects' data, Hb changes increased while that of the other decreased or stayed flat. Similarly for HbO changes. Therefore the average of concentration changes of all subjects became flat. Besides, HbO changes in the pre-task condition (baseline) were higher than that in the first half of the CPT condition by 7 over 10 subjects, which explained the decrease of HbO changes in CPT condition comparing to that in pre-task. Furthermore, motion artifacts were observed in the two 15-second instruction conditions in the data of some subjects which were the cause of the peaks in these conditions. There were no details information reported over the measurements therefore we could not verify if experimental errors existed.

3.4. Verbal-fluency task (VFT) experiment

3.4.1. Literature

3.4.1.1. Method

Herrmann et al. in 2003 investigated further the sensitivity of NIRS with cognitive activities on prefrontal cortex. Fourteen subjects participated the Verbal-Fluency Test (VFT). Subjects were asked to close eyes to measure the baseline for 60 seconds. After that they were asked to open their eyes and were instructed to pronounce as many nouns as possible which began with letter A, F and S, by order, respectively. The period for each letter was 60 seconds. The number of correct responses of subjected were recorded and served as a measure of behavioral performance. The length of four conditions (baseline, letter A, letter F and letter S) were normalized to 52, 56, 54 and 53 seconds respectively for all subjects. It was not clear how the authors normalized the data in these segments (e.g. removing samples at the beginning or at the end of each condition). Baseline correction for visualization was not reported.

Figure 25. VFT experiment in Herrmann et al. (2003), lengths of segments have been normalized across subjects.

The measurements were conducted with NIRO-300 system, which had two channels and provided Hb and HbO changes from baseline in $\mu M/l$ as output, calculated based on the modified Beer-Lambert law. The channels' positions were between Fp1 and the middle of F7 and F3 on the left hemisphere and between positions Fp2 and the middle of F4 and F8 on the right hemisphere. These positions were based on the International 10-20 system for EEG electrode placement. The sampling rate was 2Hz.

Per type of concentration changes, the average values of changes in each condition of each hemisphere per subject were calculated and these values were analyzed using RM ANOVA, 2 hemispheres x 4 conditions. Post-hoc tests were used for pairwise comparison, however no specific adjust method was reported.

3.4.1.2. Results

The results in Herrmann et al. (2003) showed that

a) The interaction between condition and hemisphere was not significant for both Hb and HbO changes.

b) There were no significant differences between hemispheres in both types of concentration.

c) HbO changes in each of the three letter conditions were significantly higher than that in baseline condition, in both hemispheres.

d) Hb changes in the letter S condition was significantly lower than that in the baseline, in both hemispheres.

e) The number of correct responses in each letter condition was similar, and no significant correlation was found between the subjects' performance and HbO changes.

Their results were interpreted as no differences were found between frontal hemispheres during the task. It was expected that only activation on the left hemisphere should have been found. The authors argued if more sensor locations could help to detect the strong left frontal activation during the VFT and if the not properly setting of the experiment caused the poor performance of the subjects.

3.4.2. Experiment at Philips Research

3.4.2.1. Procedure and materials

As described in the CPT experiment, the VFT experiment was conducted at Philips Research in a combined experiment wherein the CPT and VFT experiment both took place. In total there were data of 10 subjects, (mean age 26 (5.16), females 3, males 7), being used in this analysis. Subjects were asked to do the CPT and VFT tasks in counterbalance, five of them performed the CPT first and then VFT, the others performed VFT and then CPT.

In the VFT experiment, the letter-fluency version of the VFT was used, wherein subjects were asked to pronounce as many nouns as possible in a language that was most convenient for them. Due to

international employees it was impossible to find 'one language'- participants only. Three subjects chose Dutch, the others English for performing the task. Subjects were asked to sit still with eyes-open for 120 seconds and subsequently with closed eyes for 60 seconds, subsequently pronounce as many nouns as possible with the begin letter A, then letter F and then letter S (by order), and subsequently sit still with closed eyes for 120 seconds. There was always a 15-second instruction period in between the conditions mentioned above wherein a sound reminded subject to open eyes after an Eyes Closed condition or subject read the instruction on the screen (Figure 26). Besides some impressions about task performance, there were no recorded results of the VFT performance (i.e. the number of correct nouns in each letter condition of subjects) such as in the literature.

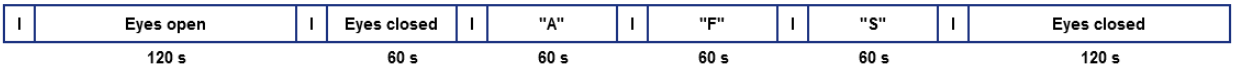


Figure 26. VFT experiment performed at Philips Research. I stands for Instruction.

The measurements were conducted with CW NIRSport device from NIRx Medical Technologies LLC., which measured in two wavelengths, 760 nm and 850 nm. The sampling rate was set at 7.8125 Hz with 20 channels setting. The inter-optode distance was 30 mm. The output data were light intensities received at detector optodes which were saved in ASCII format in two files corresponding to two wavelengths. Additional files were also created which contained other information necessary for the preprocessing and analyzing, see Appendix A. The positions of probes and channels are shown in Figure 21, the presentation of channels is illustrated in Figure 22.

Our research question

Are there significant differences in Hb and HbO changes between four conditions: baseline, letter A, letter F, and letter S, including the comparison between left and right hemisphere locations?

3.4.2.2. Data analysis

The data were analyzed using Matlab (The MathWorks Inc., MA, USA, version R2015b) software and the ezANOVA function of the ez package (Lawrence, 2015, version 4.3) in R (R Core Team, 2016) software.

Concentration changes were calculated from the raw intensities data of each subject, which were then split to two parts corresponding to two experiments CPT and VFT.

For the replication of VFT experiment, only data from the first Eyes Closed condition to the end of the last Eyes Closed condition were considered, including four Instruction periods in between. These instruction and the last Eyes Closed periods were kept in the visualization to show the trend of concentration changes continuously, however they were not used in the statistical tests. The first Eyes Closed condition was defined as baseline condition. The Instruction condition before the letter A condition was not chosen as baseline condition to avoid any activation in the local cortex which may be caused by the reading instruction in this period.

In this experiment, HbO changes of many subjects had values in large range, approx. [-5, 15] µM, while in CPT experiment the range was approx. [-4, 4] µM. The large increases often occurred slowly after the onset of letter A condition, approx. 10 seconds, and the increases were approx. 10 µM with no spikes accompanied which are completely different with motion artifacts caused by sudden head movement where the increases were fast and usually accompanied by spikes (Figure 27). Hb changes of some subjects increased in the conditions following the baseline condition at channels s2-d3 and s8-d6 but decreased at

channels s3-d3 and s6-d6 (Figure 28) while Hb changes of other subjects decreased or stayed flat during the time courses in all channels. These matched with the description of artifacts in a similar experiment in Brigadoi et al. (2014) where subjects were asked to pronounce the color name of words. The authors suggested that the movement of the jaw during speech caused these effects. It is however difficult to differentiate these artifacts with the changes caused by brain activation. A solution approach was to correct the shift in the data however as no ground true was known, we did not know if the changes was completely caused by the jaw movement or by brain activation or both, to determine the start and end of the correction. The concentration changes during post condition could be used as reference baseline to correct the shift, however that needs time to investigate as the concentration changes in this condition were not always stable in the data of all subjects (Figure 27, Figure 28). Therefore, in the scope of this thesis, we did not correct the shift. Besides, as discussed in chapter 2.3, spike removal was not applied at this phase since we were not able to fully automate this process.

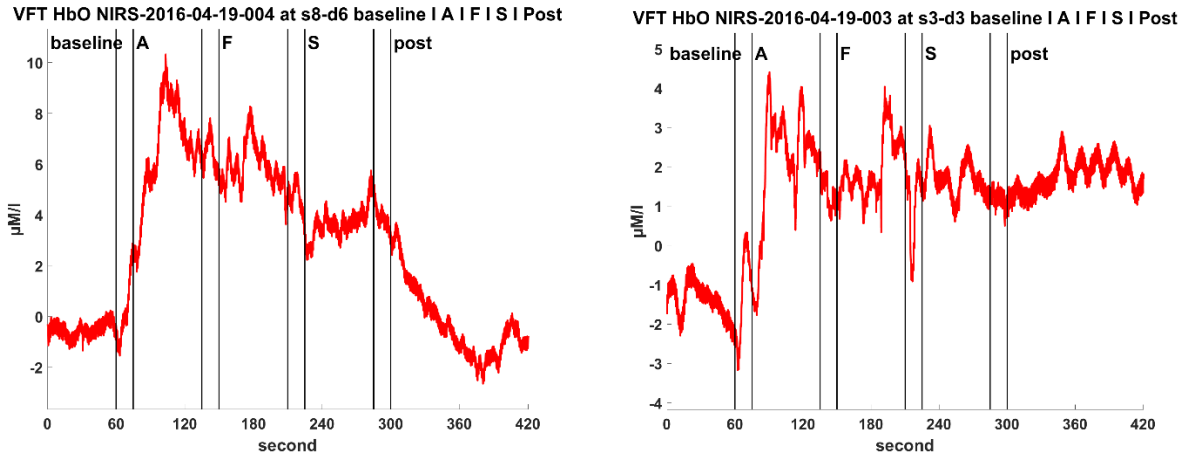


Figure 27. VFT experiment: HbO changes of two subjects at anterior channels s8-d6 (left) and s3-d3 (right) which increased sharply after the onset of the letter A condition. Vertical lines indicate the end or onset of conditions. Conditions by order of appearing are baseline, letter A, letter F, letter S and post; small segments between vertical lines are instruction conditions.

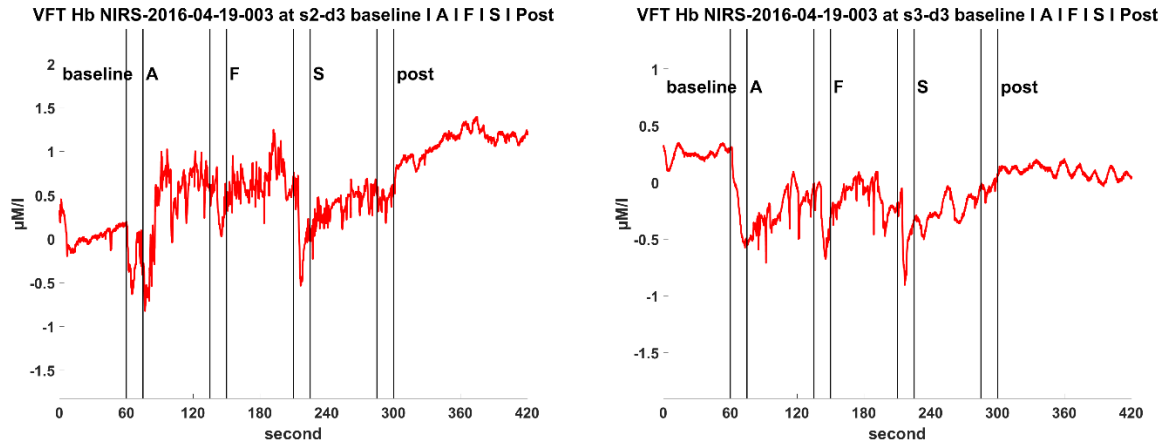


Figure 28. VFT experiment: Hb changes of a subject at anterior channels s2-d3 (left) and s3-d3 (right) which had opposite trends during the time course. Vertical lines indicate the end or onset of conditions. Conditions by order of appearing are baseline, letter A, letter F, letter S and post; small segments between vertical lines are instruction conditions.

The preprocessing and analyzing steps were the same as in the CPT experiment. The lengths of conditions were normalized across subjects wherein samples at the onset of a condition of subjects whose condition's length was longer than the minimum condition's length were excluded. The lengths of each condition: baseline, the first instruction, letter A, the 2nd instruction, letter F, the 3rd instruction, letter S, the 4th instruction and post-task after normalization were 59.9, 14.1, 59.9, 14.8, 59.9, 14.8, 59.4, 14.8 and 119.8 seconds, respectively. Data at channel s8_d6 (right hemisphere) of one subject were excluded from visualization as there were many spikes in that channel; for statistical tests, data at this channel of the subject were retained to keep the balance design for RM ANOVA tests however samples whose values were outside the range [-5, 15] µM for HbO and [-2.5,1] µM for Hb changes, of that subject's data at this channel were removed.

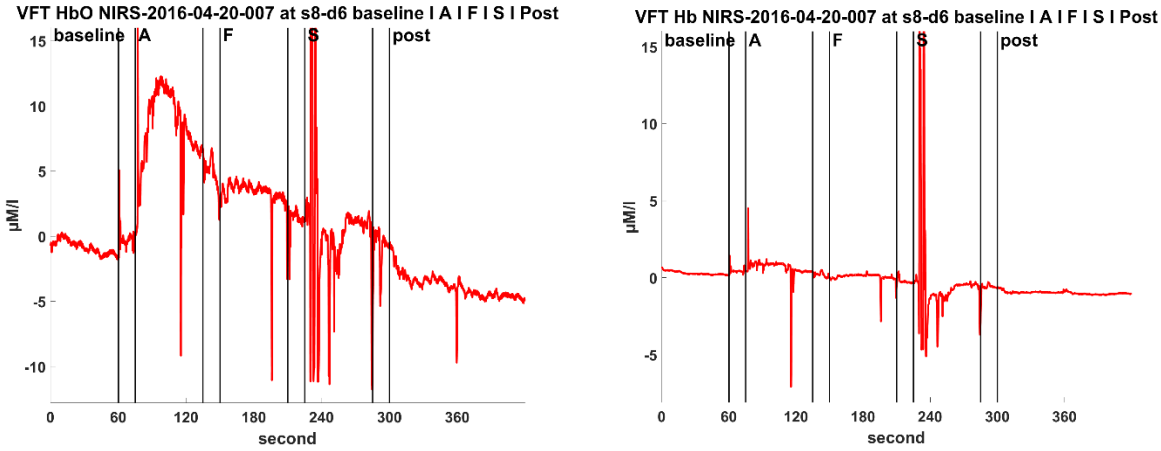


Figure 29. VFT experiment: HbO (left) and Hb (right) changes of one subjects at channel s8-d6 which contained many spikes. Vertical lines indicate the end or onset of conditions. Conditions by order of appearing are baseline, letter A, letter F, letter S and post; small segments between vertical lines are instruction conditions.

For visualization, baseline correction was applied to remove individual's differences before averaging per time point across subjects. For statistical tests, only data of baseline, and the three letter conditions were taken into account. The data in the normalized length conditions were used, which did not alter the result of the statistical tests as up to 2 samples at the onset of each condition were excluded. RM ANOVA was used to test the significance of condition (baseline, letter A, letter F and letter S), hemisphere (left and right) and the interaction between condition and hemisphere on HbO and Hb changes with data being the average HbO and Hb changes of each condition per subject (4 conditions x 2 hemispheres x 10 subjects). Post-hoc *t*-test with Bonferroni adjustment was used for pairwise comparisons when any of the effects in the results of RM ANOVA test was significant. For each pair of channels representative for left and right hemispheres there were 2 RM ANOVA tests and up to 6 post-hoc t-tests corresponding to 2 concentration changes and 3 effects.

In Herrmann et al. (2003), data were measured with 2-channel system with positions between Fp1 and the middle of F7 and F3 in the left hemisphere and between positions Fp2 and the middle of F4 and F8 in the right hemisphere. For the replication we considered either s2-d3 (channel 4) and s8-d6 (channel 19) or s3-d3 (channel 6) and s6-d6 (channel 16) to represent the left and right hemispheres, respectively (Figure 21). In total there were 4 RM ANOVA tests and up to 12 post-hoc t-tests (2 pairs of channels x 2 types of concentration changes x 3 effects).

3.4.2.3. Replication results

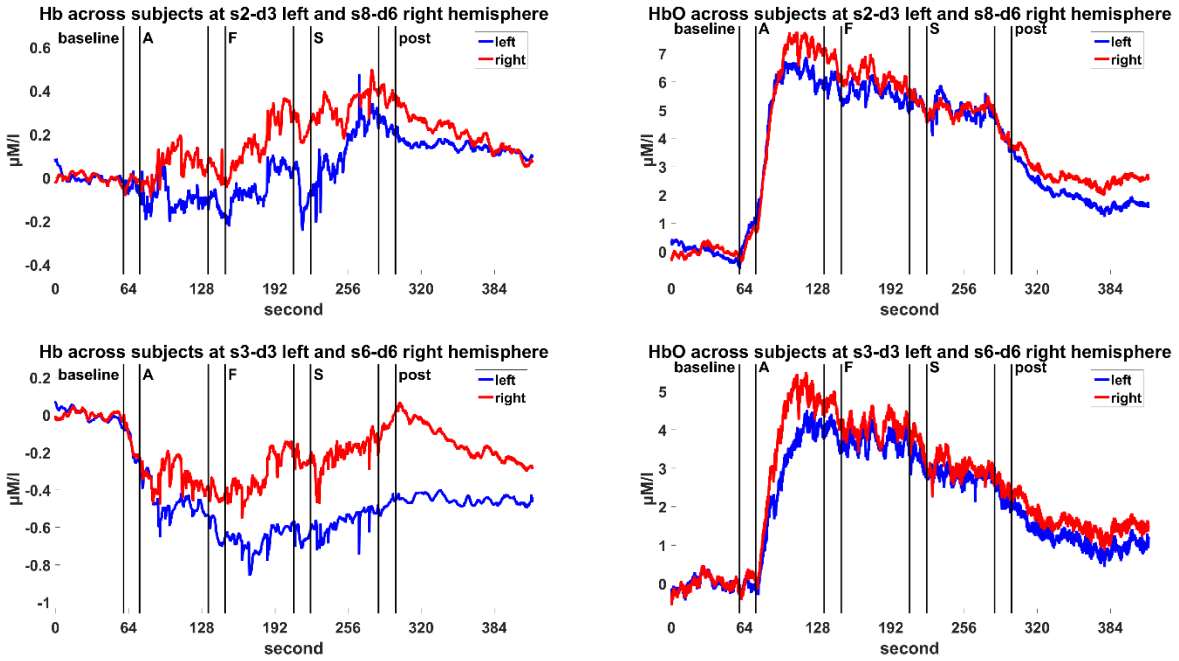


Figure 30. VFT experiment - Replication: mean across all subjects of Hb (left), HbO (right) changes between pairs of channels representative for left (blue) and right (red) hemispheres. Vertical lines indicate the end or onset of conditions. Conditions by order of appearing are baseline, letter A, letter F, letter S and post; small segments between vertical lines are instruction conditions. Pairs of channels in the first row are s2-d3 and s8-d6, and in the last row s3-d3 and s6-d6. Error bars were not shown to replicate the figures in Herrmann et al. (2003).

Table 12. VFT experiment - Replication: Mean and standard deviation across subjects of Hb and HbO changes during letter conditions (baseline correction applied) at channels s2-d3 and s3-d3, representative for the left hemisphere, and at channels s6-d6 and s8-d6, representative for the right hemisphere.

Left hemisphere					**Right hemisphere**				
channel		**lette A**	**letter F**	**letter S**	**channel**		**lette A**	**letter F**	**letter S**
s2_d3	Hb	-0.09 ± 0.67	-0.04 ± 0.86	0.13 ± 0.96	s8_d6	Hb	0.09 ± 0.47	0.13 ± 0.65	0.26 ± 0.94
	HbO	5.47 ± 4.34	5.59 ± 5.03	4.97 ± 3.89		HbO	6.31 ± 5.35	5.94 ± 5.88	4.83 ± 4.39
s3_d3	Hb	-0.46 ± 0.37	-0.67 ± 0.53	-0.57 ± 0.55	s6_d6	Hb	-0.33 ± 0.44	-0.32 ± 0.59	-0.21 ± 0.89
	HbO	2.90 ± 4.44	3.69 ± 5.46	2.81 ± 4.79		HbO	3.82 ± 4.10	4.08 ± 4.65	3.02 ± 3.21

The replication results showed that:

a) The interaction between condition and hemisphere was not significant for both Hb and HbO changes, in both pairs of channels s2-d3 and s8-d6, and s3-d3 and s6-d6 which were representative for left and right hemispheres. Same as that in the literature.

b) There were no significant differences between hemispheres in both types of concentration, in both pairs of channels. Same as that in the literature.

c) HbO changes in each letter condition A, F and S were significantly higher than that in the baseline condition, in both hemispheres and in both pairs of channels representative for left and right hemispheres. The same as that in Herrmann et al. (2003).

Furthermore, in the pair s3-d3 and s6-d6, the HbO changes in the letter F condition were significantly higher than that in the letter S condition in both hemispheres. This was not the case in the results of Herrmann et al. (2003).

d) Hb changes in the letter A and F condition were significantly lower than that in the baseline condition, in both hemispheres and in the pair s3-d3 and s6-d6. In Herrmann et al. (2003), the significant differences were observed in Hb changes between the letter S and the baseline condition.

e) The number of correct responses in each letter condition was not archived therefore there were no results regarding to the correlation between the subjects' performance and concentration changes.

In visualization, the trends of HbO changes in both hemispheres, representing in both pairs of channels s2-d3 and s8-d6, and s3-d3 and s6-d6, were consistent, which increased sharply in the letter A condition and decreased gradually in the subsequent conditions, while in the results of Herrmann et al. (2003), HbO changes increased gradually in the letter A condition, stayed in the same level during the letter F and S conditions and only decreased back to baseline in the post condition. The trends of Hb changes in both hemispheres were less consistent comparing to that in Herrmann et al. (2003). In the pair s2-d3 and s8-d6, Hb changes increased in the three letter conditions and gradually decreased in the post condition, while in Herrmann et al. (2003) the trend was decreasing. In the pair s3-d3 and s6-d6, Hb changes decreased in the letter A condition and increased in the letter F and S conditions, while in Herrmann et al. (2003), the trend was decreasing in all conditions following the baseline condition.

Table 13. VFT experiment - Replication: Statistical results for pairs of channels representative for left and right hemispheres: s2-d3 and s8-d6, and s3-d3 and s6-d6. Bold p values are the corrected p values based on Greenhouse-Geisser correction, degree of freedom is not corrected in that case, ns. stands for not significant.

Hemisphere			**RM ANOVA**			**Pairwise post-hoc t-tests**
Left	**Right**		**Effect**		**Greenhouse-Geisser epsilon**	
s2_d3	s8_d6	Hb	condition	F(3,27)=0.29, p=0.68	0.475	ns.
			hemisphere	F(1,9)=0.146, p=0.711		ns.
			interaction	F(3,27)=0.704, **p=0.486**	0.559	ns.
		HbO	condition	F(3,27)=11.17, **p=0.002***	0.506	A vs. baseline p<0.001* F vs. baseline p<0.001* S vs. baseline p<0.001*
			hemisphere	F(1,9)=0.44, p=0.524		ns.
			interaction	F(3,27)=0.923, p=0.443		ns.
s3_d3	s6_d6	Hb	condition	F(3,27)=5.073, **p=0.027***	0.535	A vs. baseline p=0.002* F vs. baseline p=0.007*
			hemisphere	F(1,9)=2.473, p=0.15		ns.
			interaction	F(3,27)=2.777, **p=0.123**	0.383	ns.
		HbO	condition	F(3,27)=5.266, **p=0.032***	0.452	A vs. baseline p=0.01* F vs. baseline p=0.01* S vs. baseline p=0.02* F vs. S p=0.04*
			hemisphere	F(1,9)=0.515, p=0.491		ns.

	interaction	F(3,27)=1.381, **p=0.275**	0.414	ns.

3.4.2.4. Discussion and conclusion

The replicated results were not completely the same as that in Herrmann et al. (2003). Although the statistical results were similar in both experiments, the trend of HbO changes across all subjects in the replicated data was not similar to that in Herrmann et al. (2003). In the replicated data, large increases of HbO changes after the onset of condition A were observed, particularly at channels s2-d3 and s8-d6. The increases of HbO changes at channels s3-d3 and s6-d6 were also large but less than those at the aforementioned channels. At channels s2-d3 and s8-d6 in the data of many subjects, Hb changes increased after the onset of condition A and only decreased after the onset of condition post-task which was different from that in Herrmann et al. (2003), while they decreased at channels s3-d3 and s6-d6 and thus matched with the results in Herrmann et al. (2003). It was suggested that the jaw movement might have caused these effects (Brigadoi et al., 2014). Besides, motion artifacts observed in the replicated data of some subjects were not processed properly. Motion artifacts should be expected due to the nature of the experiment (jaw movement), however it was not reported in Herrmann et al. (2003). The 15-second instruction periods in this experiment, in contrast to those in the CPT experiment, seemed not affecting obviously to the trend of HbO and Hb changes across subjects. Probably its effect was small comparing to the large influence of motion artifacts. Furthermore, in Herrmann et al. (2003), no baseline correction was reported for visualization, and the concentration changes during baseline in figure 1 (across subjects) and in figure 2 (one subject) were not around 0. The output of their device was the concentration changes from a baseline, which was not from the baseline condition mentioned in the experiment. It suggested that the authors did not correct the baseline for each subject based on the mean of the baseline condition as we did with the replicated data. These together might be the reasons which explain the differences between the results of the replicated data and that in Herrmann et al. (2003).

4. A suggested approach for data analysis

This chapter describes a proposed approach for processing and analyzing fNIRS data, and its application for the data of each experiment conducted at Philips Research.

Several noise removed methods have been discussed in chapter 2.3. In the scope of this thesis, for the four datasets, we proposed an approach wherein low-pass filter with cutoff 0.5 Hz and subsequently high-pass filter with cutoff 0.01 Hz were applied on the optical density to remove high frequency noise and the very slow drift in the signal due to their simplicity and effectiveness reported in literature (see 2.3). The low-pass filter with cutoff 0.5 Hz which was used in various applications (Keles el al., 2016; Brigadoi et al., 2014; Cui el al., 2011) does not impact signal components which have phases longer than 2 seconds including the typical hemodynamic response. Similarly, the high-pass filter with cutoff 0.01 Hz removes only very slow drift in the signal corresponding to features longer than 1 minutes.

Subsequently, concentration changes were calculated and block averaging was applied to further reduce the variance caused by physiological noise.

MA correction was not considered in this approach yet since they occurred mostly in the data of CPT and particularly VFT experiment and coupled with baseline shift which needed more time to investigate. Besides, the correction of spikes, results of MAs, required manual work which we skipped in this work due to time constraint. As expected, the physiological cardiac artifact (~1 Hz) was removed by the low-pass filter with cutoff 0.5 Hz. As discussed earlier in the review in Chapter 2.4., physiological noise could be further reduced using PCA method although it is not a straight forward task due to the hypothesis on temporal correlation between baseline and task condition. We find this method promising with some modification of data selection for baseline and task condition. However due to time constraint, we did not consider it in this approach for the four dataset in this study.

To evaluate the proposed approach, we compared the statistical test results of this approach with the results of the approach in literature (chapter 3) for each experiment.

The next section will describe the method applied and the corresponding results for each experiment. The data were analyzed using Matlab software (The MathWorks Inc., MA, USA, version R2015b), the function hmrBandpassFilt from the HOMER2 software (Huppert et al., 2009) and the ezANOVA function of the ez package (Lawrence, 2015, version 4.3) in R software (R Core Team, 2016). The low-pass filtering with Butterworth filter and filter order 3, and high-pass filtering with Butterworth filter and filter order 5 setting were implemented in the function hmrBandpassFilt.

4.1. Visual stimulation experiment

4.1.1. Data analysis

Concentration changes were calculated from the filtered raw intensities data of each subject using low-pass and high-pass filter as described above. After that the same steps as in the replication (section 3.1.2.2) were performed.

4.1.2. Results

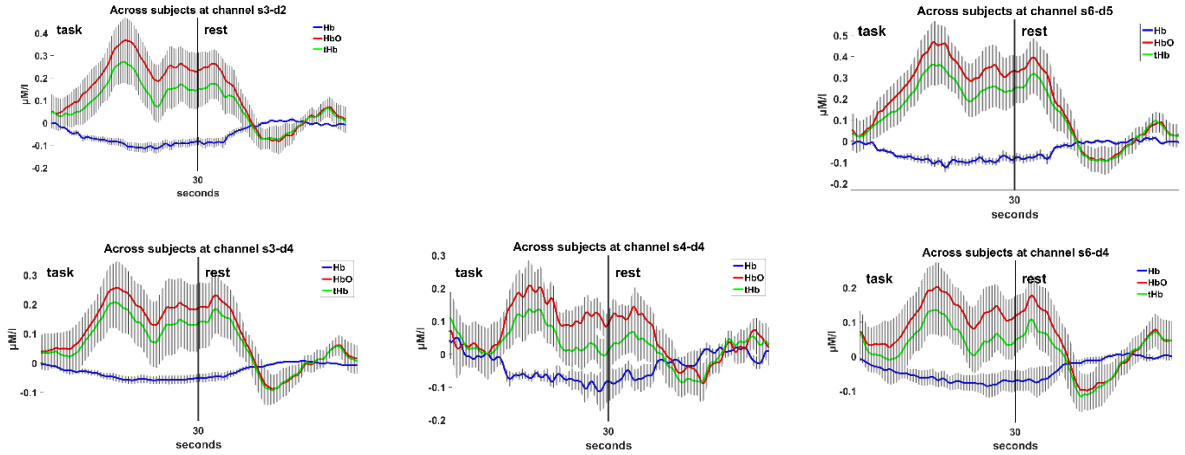


Figure 31. Visual stimulation - Proposed approach: HbO (red), tHb (green) and Hb (blue) changes at s3_d2, s6_d5 (top, left to right), s3_d4, s4_d4 and s6_d4 (bottom, left to right) corresponding to 5% above O_1, 5% above O_2, O_1, Oz and O_2. Data are presented as mean and standard error of the mean (SEM). Vertical line separates two conditions: stimulus presentation (task) and baseline (rest).

Table 14. Visual stimulation - Proposed approach: Mean across subjects of concentration changes (μM) of the last 15 seconds of the Task condition for set of locations s3_d4, s4_d4, s6_d4, s3_d2 and s6_d5 corresponding to O_1, Oz, O_2, 5% above O_1 and 5% above O_2 (baseline correction applied). SEM stands for standard error of the mean.

		HbO (Mean ± SEM)	Hb (Mean ± SEM)	tHb (Mean ± SEM)
O_1	s3_d4	0.19 ± 0.04	-0.06 ± 0.01	0.13 ± 0.04
Oz	s4_d4	0.13 ± 0.02	-0.08 ± 0.02	0.05 ± 0.02
O_2	s6_d4	0.13 ± 0.03	-0.08 ± 0.01	0.06 ± 0.03
5% above O_1	s3_d2	0.26 ± 0.04	-0.10 ± 0.01	0.16 ± 0.04
5% above O_2	s6_d5	0.35 ± 0.05	-0.09 ± 0.01	0.26 ± 0.04

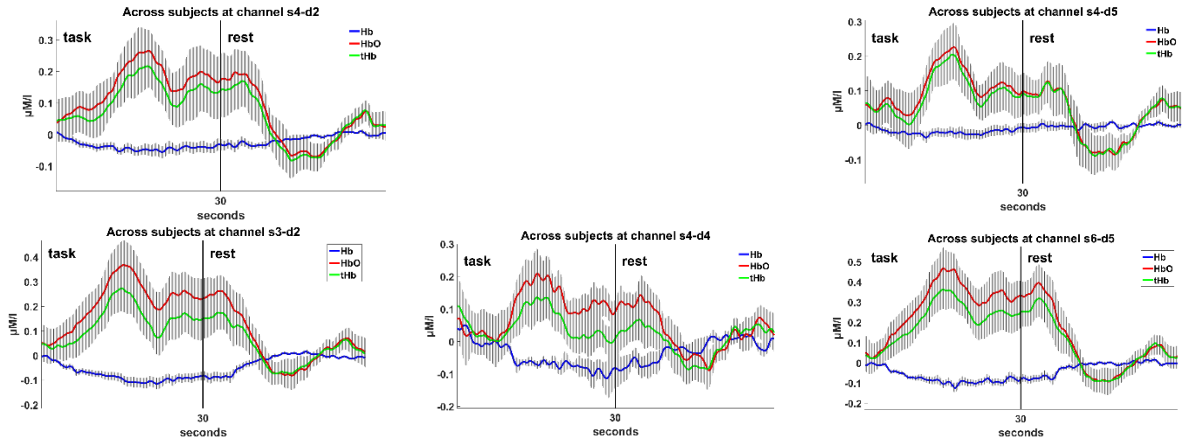


Figure 32. Visual stimulation - Proposed approach: HbO (red), tHb (green) and Hb (blue) changes at s4_d2, s4_d5 (top, left to right), s3_d2, s4_d4, and s6_d5 (bottom, left to right) corresponding to 5% above O_1, 5% above O_2, O_1, Oz and O_2. Data are presented as mean and standard error of the mean (SEM). Vertical line separates two conditions: stimulus presentation (task) and baseline (rest).

Table 15. Visual stimulation - Proposed approach: Mean across subjects of concentration changes (μM) of the last 15 seconds of the Task condition for set of locations s3_d2, s4_d4, s6_d5, s4_d2 and s4_d5 corresponding to O_1, Oz, O_2, 5% above O_1 and 5% above O_2 (baseline correction applied). SEM stands for standard error of the mean.

		HbO (Mean ± SEM)	Hb (Mean ± SEM)	tHb (Mean ± SEM)
O_1	s3_d2	0.26 ± 0.04	-0.10 ± 0.01	0.16 ± 0.04
Oz	s4_d4	0.13 ± 0.02	-0.08 ± 0.02	0.05 ± 0.02
O_2	s6_d5	0.35 ± 0.05	-0.09 ± 0.01	0.26 ± 0.04
5% above O_1	s4_d2	0.19 ± 0.04	-0.04 ± 0.01	0.15 ± 0.04
5% above O_2	s4_d5	0.13 ± 0.02	-0.02 ± 0.01	0.11 ± 0.03

The results of the proposed approach showed that:

a) The effect of condition was significant for HbO, Hb and tHb changes, same as that in the replication results.

   The effect of location was **not** significant for HbO, Hb and tHb changes, same as that in the replication results.

   The effect of the interaction between location and condition was significant for HbO changes in both sets, and was significant for Hb in one set, same as that in the replication results.

   The effect of the interaction between location and condition was **not** significant for tHb changes in both sets. It was significant in the replication results however there were no significant pairwise comparison in the replication results, therefore I consider they are the same in both approaches.

b) HbO, Hb and tHb changes at Oz were not significantly different than that at O_1 and O_2, same as that in the replication results.

c) HbO, Hb and tHb changes at Oz were not significantly different from that at 5% above O_1 and 5% above O_2, same as that in the replication results.

d) The increase observed in HbO changes was greater than the decrease observed in Hb changes, same as that in the replication results.

e) The trend of tHb was similar to that of HbO, same as that in the replication results.

f) During Task condition, Hb changes at O_1 (s3_d2) were significantly different from that at 5% above O_2 (s4_d5), which was not the case in the replication results.

In visualization, the trends of Hb, HbO and tHb changes across trials and across subjects were similar to that in the replication results but less noisy.

Table 16. Visual stimulation - Proposed approach: Results of RM ANOVA for set of locations s3_d4, s4_d4, s6_d4, s3_d2 and s6_d5 corresponding to O_1, Oz, O_2, 5% above O_1 and 5% above O_2. Bold p values are the corrected p values based on Greenhouse-Geisser correction, degree of freedom is not corrected in that case, ns. stands for not significant, ni. stands for not important where interaction between task condition at one location and rest condition at another location is significant.

	Effect		Greenhouse-Geisser epsilon	Pairwise comparison
Hb	Condition	F(1,10)=71.454, p<.001*		*p*<.001*
	Location	F(4,40)=0.575, ***p*=0.483**	0.2809753	

	Interaction	F(4,40)=0.653, ***p*=0.486**	0.3588556	
HbO	Condition	F(1,10)=13.773, *p*=0.004*		*p*<.001*
	Location	F(4,40)=0.086, ***p*=0.917**	0.4983800	
	Interaction	F(4,40)=5.721, *p*=0.001*		ni.
tHb	Condition	F(1,10)=6.675, *p*=0.027*		*p*<.001*
	Location	F(4,40)=0.152, *p*=0.961		
	Interaction	F(4,40)=3.123, ***p*=0.083**	0.3818262	

* significant

Table 17. Proposed approach: Results of RM ANOVA for set of locations s3_d2, s4_d4, s6_d5, s4_d2 and s4_d5 corresponding to O_1, Oz, O_2, 5% above O_1 and 5% above O_2. Bold p values are the corrected p values based on Greenhouse-Geisser correction, degree of freedom is not corrected in that case, ns. stands for not significant, ni. stands for not important where interaction between task condition at one location and rest condition at another location is significant.

	Effect		Greenhouse-Geisser epsilon	Pairwise comparison
Hb	Condition	F(1,10)=44.683, *p*<.001*		*p*<.001*
	Location	F(4,40)=0.789, ***p*=0.424**	0.3281714	
	Interaction	F(4,40)=3.588, *p*=0.014*		two conditions at s3_d2, *p*=0.015* task condition at s4_d5 and s3_d2, *p*=0.024* two conditions at s6_d5, *p*=0.013*
HbO	Condition	F(1,10)=18.071, *p*=0.002*		*p*<.001*
	Location	F(4,40)=0.086, ***p*=0.883**	0.4070861	
	Interaction	F(4,40)=4.843, *p*=0.003*		ni.
tHb	Condition	F(1,10)=8.945, *p*=0.014*		*p*<.001*
	Location	F(4,40)=0.452, *p*=0.770		
	Interaction	F(4,40)=2.581, ***p*=0.101**	0.4962804	

* significant

4.1.3. Discussion and conclusion

The traces of HbO, tHb and Hb were similar between two approaches however they were smoother in the proposed approach by using low-pass and high-pass filters as expected (see Figure 7, Figure 8, Figure 31, and Figure 32). By design, the proposed approach does minimize instrumental noise and physiological noise with very low frequency as expected. The statistical results in the proposed approach were similar to that of the replication approach although an additional processing step had been applied. Here, the large number of repeated measures in this experiment (10 trials) plays important role in the results of both approaches. When considering data of the first 2 trials only, the results of the proposed approach were better with more significant effects than that of the replication approach. Less number of repetitions has a great impact in practice, particularly for a quick test in which a small number of trials is required.

4.2. Finger tapping experiment

4.2.1. Data analysis

Concentration changes were calculated from the filtered raw intensities data of each subject as described above. Instead of grouping samples into group of 8 (approx. 1 second), block averaging was applied for each tapping frequency after baseline correction where baseline was defined as the last 20 second of the rest condition preceding each tapping condition. Thus each tapping condition had its own baseline.

For visualization, similar to the replication, each tapping trial was illustrated with a 20-seconds rest segment preceding the tapping period (baseline), a 20-seconds tapping segment and a 40-seconds rest segment following the tapping period to show the trend of concentration changes before, during and after the tapping continuously. The lengths of tapping and rest segments across trials and across subjects were normalized, wherein samples at the onset the tapping segment whose condition's length was longer than the minimum condition's length were excluded, while for the rest segment, samples from the onset up to the minimum length of all rest segment or to the 40^{th} seconds were included. Length normalization was not necessary for baseline as its length (20 seconds) was always smaller than the length of the corresponding rest condition (approx. 40 seconds). After normalization, the lengths of tapping condition and rest condition were 19.97 and 39.94 seconds, respectively. Baseline correction was applied per trial where the mean of the baseline was subtracted from each time point in 3 segments. Finally, average changes per time point across all trials and across all subjects for each tapping frequency were used to illustrate the trend of concentration changes per tapping frequency (2Hz, 4Hz and ME).

For statistical tests, all the time points in the normalized length of the tapping condition in each trial were taken into account after subtracting the mean of the corresponding baseline. The number of time points which were excluded at the onset of the tapping condition in each trial was up to 1 and thus did not alter the statistical results. Similar to the tests in the replication, RM ANOVA was used to test the significant differences of tHb, HbO or Hb changes between tapping frequencies with frequency being the effect of the test. The data for each RM ANOVA test were the mean of tHb, HbO or Hb changes across trials in each tapping frequency per subject, per channel, 3 frequencies x 5 subjects. Post-hoc t-test with Bonferroni adjustment on the same data was used for pairwise comparison when frequency was significant in the result of RM ANOVA test. For each channel, there were 3 RM ANOVA tests and up to 3 post-hoc t-tests, corresponding to 3 types of concentration changes. Finally, t-tests were used to test the significant differences of tHb, HbO or Hb changes at each tapping frequencies (baseline corrected already) and 0 (baseline). The data for each of these t-tests were the mean of tHb, HbO or Hb changes across trials in one tapping frequency, per subject, per channel, 1 frequency x 5 subjects. There were 9 t-tests per channel corresponding to 3 frequencies and 3 types of concentration changes. All statistical inferences were based on an adjusted alpha level of 5%.

Similar to the replication, four channels around C3 were considered which were s4-d3, s2-d3, s1-d3 and s3-d3. In total there were 12 RM ANOVA tests, up to 12 post-hoc t-tests (4 channels x 3 types of concentration changes) and 36 t-test (4 channels x 3 frequencies x 3 types of concentration changes).

4.2.2. Results

Table 18. Mean and standard deviation across subjects of concentration changes during task condition (baseline correction applied) at channel s4-d3.

	2 Hz	4 Hz	ME
tHb	0.14 ± 0.29	0.02 ± 0.27	0.35 ± 0.31
HbO	0.15 ± 0.40	0.13 ± 0.31	0.44 ± 0.37
Hb	-0.02 ± 0.13	-0.11 ± 0.12	-0.09 ± 0.14

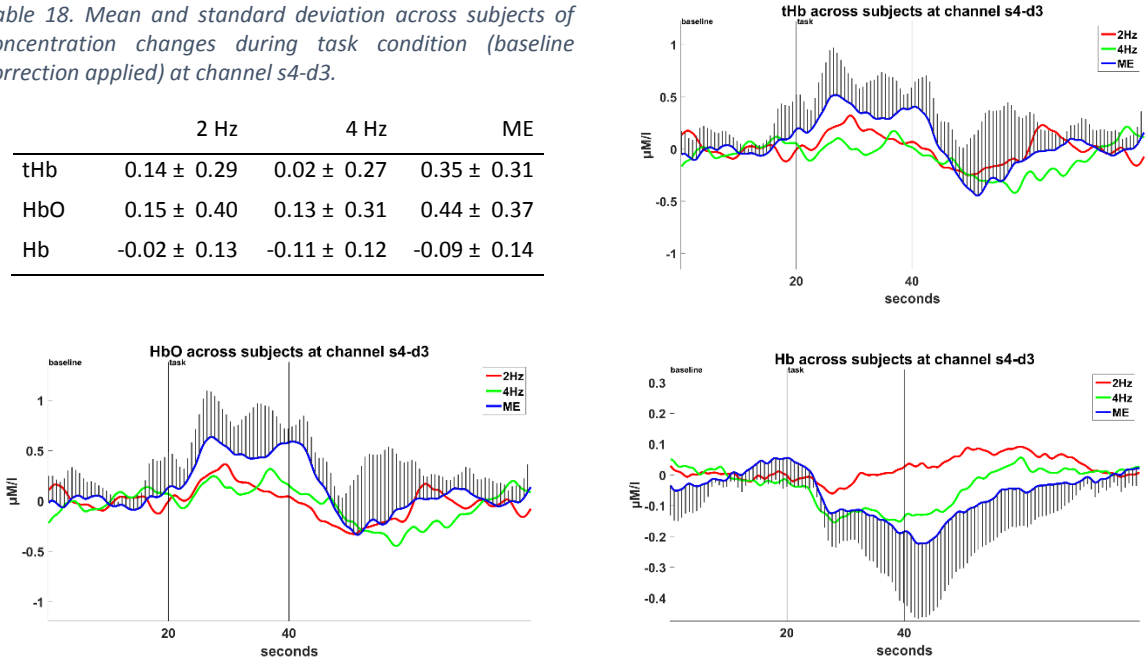


Figure 33. Finger tapping experiment – Proposed approach: mean and standard deviation across all subjects of tHb (top), HbO (bottom left) and Hb (bottom right) changes in 2 Hz (red), 4 Hz (green) and ME (blue) at channel s4-d3. Vertical lines indicate onset and end of tapping condition. Standard deviation was shown to replicate the figures in Kuboyama et al. (2004) and that applied for ME frequency only to have a clear view.

Table 19. Mean and standard deviation across subjects of concentration changes during task condition (baseline correction applied) at channel s2-d3.

	2 Hz	4 Hz	ME
tHb	0.09 ± 0.20	-0.01 ± 0.25	0.35 ± 0.32
HbO	0.13 ± 0.23	0.14 ± 0.29	0.55 ± 0.43
Hb	-0.04 ± 0.10	-0.15 ± 0.10	-0.20 ± 0.19

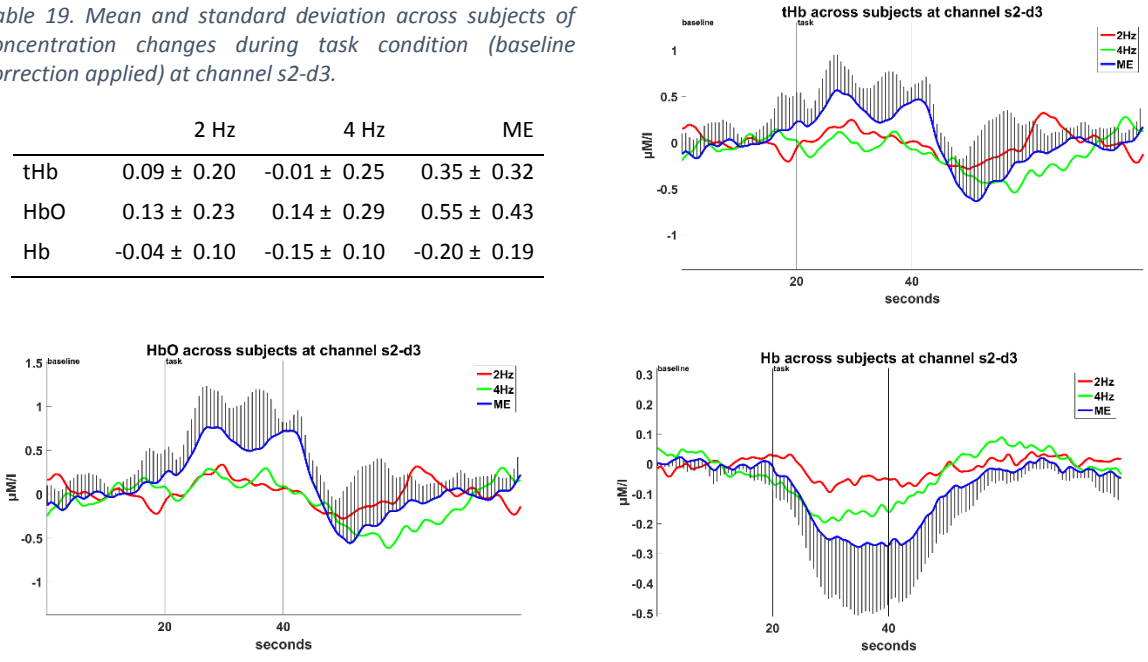


Figure 34. Finger tapping experiment – Proposed approach: mean and standard deviation across all subjects of tHb (top), HbO (bottom left) and Hb (bottom right) changes in 2 Hz (red), 4 Hz (green) and ME (blue) at channel s2-d3. Vertical lines indicate onset and end of tapping condition. Standard deviation was shown to replicate the figures in Kuboyama et al. (2004) and that applied for ME frequency only to have a clear view.

Table 20. Mean and standard deviation across subjects of concentration changes during task condition (baseline correction applied) at channel s1-d3.

	2 Hz	4 Hz	ME
tHb	0.12 ± 0.31	-0.05 ± 0.32	0.60 ± 0.37
HbO	0.15 ± 0.40	0.07 ± 0.41	0.82 ± 0.45
Hb	-0.03 ± 0.18	-0.12 ± 0.15	-0.22 ± 0.21

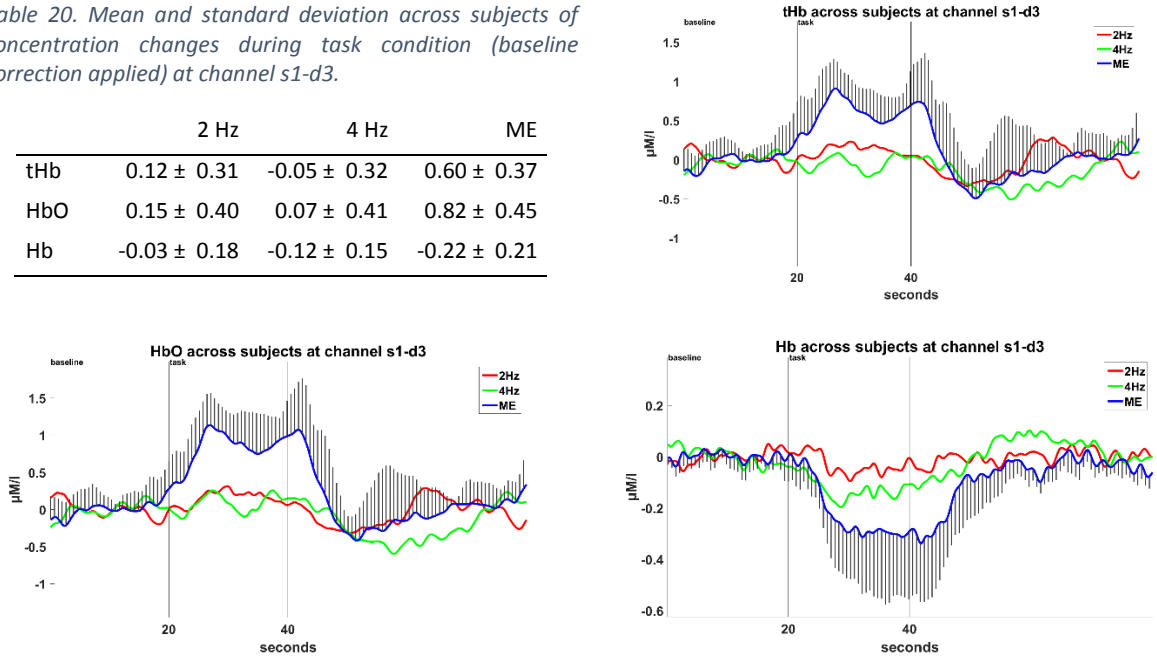


Figure 35. Finger tapping experiment – Proposed approach: mean and standard deviation across all subjects of tHb (top), HbO (bottom left) and Hb (bottom right) changes in 2 Hz (red), 4 Hz (green) and ME (blue) at channel s1-d3. Vertical lines indicate onset and end of tapping condition. Standard deviation was shown to replicate the figures in Kuboyama et al. (2004) and that applied for ME frequency only to have a clear view.

Table 21. Mean and standard deviation across subjects of concentration changes during task condition (baseline correction applied) at channel s3-d3.

	2 Hz	4 Hz	ME
tHb	0.05 ± 0.26	-0.05 ± 0.3	0.26 ± 0.51
HbO	0.07 ± 0.4	0 ± 0.4	0.51 ± 0.67
Hb	-0.02 ± 0.25	-0.06 ± 0.15	-0.25 ± 0.24

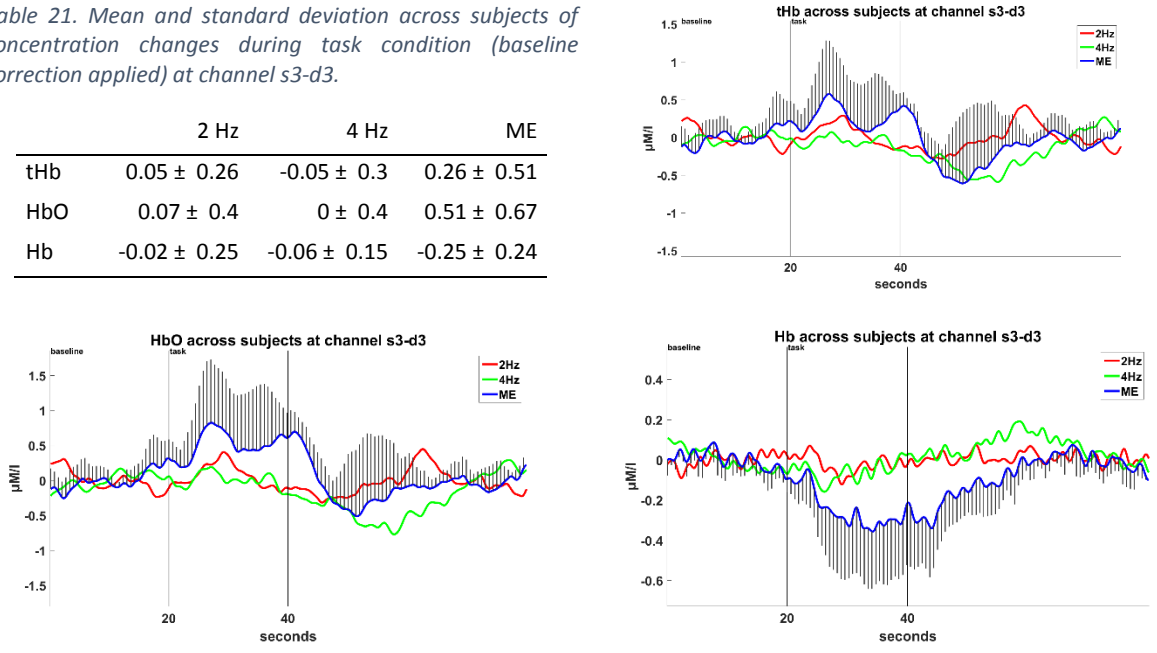


Figure 36. Finger tapping experiment – Proposed approach: mean and standard deviation across all subjects of tHb (top), HbO (bottom left) and Hb (bottom right) changes in 2 Hz (red), 4 Hz (green) and ME (blue) at channel s3-d3. Vertical lines indicate onset and end of tapping condition. Standard deviation was shown to replicate the figures in Kuboyama et al. (2004) and that applied for ME frequency only to have a clear view.

The results of the proposed approach show that:

a) The increase of tHb, HbO, and Hb changes from baseline at ME was significant at channel s1_d3 (3), same as that in the replication.

At channel s4_d3 (9) and s2_d3 (5), only the increase of tHb and HbO changes from baseline was significant, that of Hb was not significant.

The decrease of Hb changes from baseline at 4 Hz was significant at channel s2_d3 (5) only while it was also significant at channel s4_d3 (9) in the replication. The increase of HbO changes from baseline at 2 Hz was not significant at channel s2_d3 (5), while it was significant in the replication.

b) At ME the increase of tHb, HbO changes from baseline was significantly higher than that at 4 Hz and 2 Hz at channel s1_d3 (3), same as that in the replication.

The increase of tHb at ME was also significantly higher than that at 4 Hz at channel s4_d3 (9), same as that in the replication.

c) The decrease of Hb, and increase of HbO and tHb changes from their baseline was not significantly different between 2Hz and 4Hz, same as that in the replication.

The results at channel s3_d3 was not significant, similar to that in the replication.

Table 22. Proposed approach: statistical results for s4-d3, s2-d3, s1-d3 and s3-d3. Column 3, 4 and 5 contain p-values of t-test between tapping frequency and baseline. Column RM ANOVA contains the result of RM ANOVA for Frequency effect. Column 7, 8 and 9 contain p-value of post-hoc pairwise t-test between frequencies.

		t-test			**RM ANOVA**	**post-hoc t-test**		
		p-value 2 Hz baseline	*p*-value 4 Hz baseline	*p*-value ME baseline	Frequency effect	*p*-value 4Hz vs. 2Hz	*p*-value ME vs. 2Hz	*p*-value ME vs. 4Hz
s4_d3	tHb	0.254	0.782	0.011*	F(2,8)=5.132, *p*=0.037*	0.836	0.54	0.049*
	HbO	0.394	0.257	0.016*	F(2,8)=2.28, *p*=0.165	1	0.68	0.077
	Hb	0.816	0.086	0.074	F(2,8)=0.905, *p*=0.442	1	1	1
s2_d3	tHb	0.056	0.863	0.018*	F(2,8)=8.577, *p*=0.01*	0.462	0.238	0.053
	HbO	0.081	0.217	0.020*	F(2,8)=6.789, *p*=0.019*	1	0.209	0.085
	Hb	0.412	0.026*	0.061	F(2,8)=2.489, *p*=0.144	0.527	0.536	0.975
s1_d3	tHb	0.151	0.599	0.005*	F(2,8)=19.229, *p*=0.001*	0.48	0.002*	0.037*
	HbO	0.299	0.619	0.004*	F(2,8)=17.268, *p*=0.001*	1	0.028*	0.023*
	Hb	0.727	0.118	0.050*	F(2,8)=1.676, *p*=0.247	0.882	0.738	0.918
s3_d3	tHb	0.472	0.597	0.264	F(2,8)=1.679, *p*=0.246	1	1	0.435
	HbO	0.661	0.981	0.137	F(2,8)=2.13, *p*=0.181	1	0.848	0.436
	Hb	0.862	0.367	0.073	F(2,8)=1.42, *p*=0.297	1	0.944	0.662

* significant

4.2.3. Conclusion

The study for this dataset came to similar conclusion as in the Visual stimulation dataset.

4.3. Continuous performance test (CPT) experiment

4.3.1. Data analysis

Concentration changes were calculated from the filtered raw intensities data of each subject using low-pass and high-pass filter as described above. After that the same steps as in the replication (section 3.3.2.2) were performed except the smoothing step which was applied for visualization in the replication approach.

4.3.2. Results

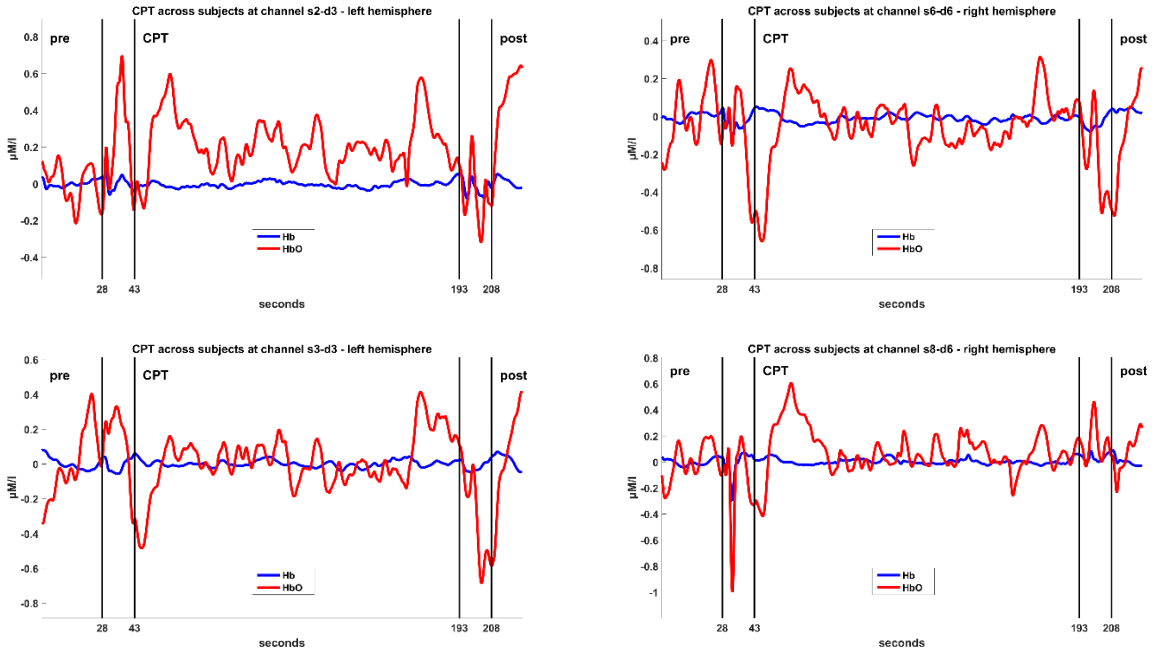


Figure 37. CPT experiment - Proposed approach: Hb (blue) and HbO (red) changes in left and right hemispheres in three condition pre-task, CPT and post-task in CPT experiment. Vertical line which follows or precedes each condition indicates the end or onset of that condition, the small segments between these lines are instruction periods. Channel s2_d3 (top left) and channel s3_d3 (bottom left) represent left hemisphere. Channel s6_d6 (top right) and channel s8_d6 (bottom right) represent right hemisphere.

Table 23. CPT experiment - Proposed approach: Mean and standard deviation across subjects of Hb and HbO (μM) of CPT and post condition in four channels (baseline correction applied).

Channel		CPT	Post	channel		CPT	Post
s2_d3	Hb	-0.01 ± 0.11	0.01 ± 0.10	s6_d6	Hb	-0.01 ± 0.07	0.03 ± 0.10
	HbO	0.21 ± 0.32	0.45 ± 0.47		HbO	-0.04 ± 0.18	-0.09 ± 0.27
s3_d3	Hb	0.00 ± 0.09	0.03 ± 0.08	s8_d6	Hb	0.00 ± 0.03	0.00 ± 0.10
	HbO	0.02 ± 0.13	0.02 ± 0.28		HbO	0.08 ± 0.18	0.06 ± 0.36

In this approach, the statistical results of both pair of channels s3_d3 and s6-d6, and s2_d3 and s8-d6, representing the left and right hemispheres respectively, were the same with that in the replication. Although RM ANOVA test return significant result for the effect condition in HbO changes for the pair s2-d3 and s8-d6, no significant differences between conditions were found in the post-hoc pairwise t-test. The traces of HbO and Hb in this approach were different from that in the replication.

Table 24. CPT experiment - Proposed approach: statistical results for four pairs of channels representing left and right hemispheres. Column 4 and 5 contain the results of RM ANOVA for the corresponding effect. Column post-hoc t-test contains results of post-hoc t-test for the corresponding effects.

Hemisphere				RM ANOVA	post-hoc t-test
Left	**Right**		**Effect**		**p-values**
s3_d3	s6_d6	Hb	Condition	F(2,18)=1.07, p=0.364	-
			Location	F(1,9)=1.313, p=0.281	-
			Interaction	F(2,18)=0.572, p=0.574	-
		HbO	Condition	F(2,18)=0.116, p=0.891	-
			Location	F(1,9)=0.012, p=0.916	-
			Interaction	F(2,18)=0.898, p=0.425	-
s2_d3	s8_d6	Hb	Condition	F(2,18)=0.019, p=0.981	-
			Location	F(1,9)=0.282, p=0.608	-
			Interaction	F(2,18)=0.15, p=0.862	-
		HbO	Condition	F(2,18)=4.807, p=0.021*	ns.
			Location	F(1,9)=0.386, p=0.55	-
			Interaction	F(2,18)=3.102, p=0.07	-

- not applicable

4.3.3. Discussion and conclusion

It was difficult to evaluate the visualization of the proposed approach and the replication approach since the very slow drift in the original signal was removed by the high-pass filter in the first approach, while in the second approach, the traces of HbO and Hb changes across subjects were smoothed.

The proposed approach clearly showed the presence of low frequency components in the traces of HbO changes while it was masked in the replication approach due to smoothing.

By design, the proposed approach did minimize instrumental noise, however physiological noise was still present in the filtered signal. Furthermore, due to the presence of MAs which have not been corrected as mentioned earlier, the statistical results of the proposed approach on this dataset did not strongly indicate better performance to that of the replication approach.

4.4. Verbal-fluency test (VFT) experiment

4.4.1. Data analysis

Concentration changes were calculated from the filtered raw intensities data of each subject using low-pass and high-pass filter as described above. After that the same steps as in the replication (section 3.4.2.2) were performed.

4.4.2. Results

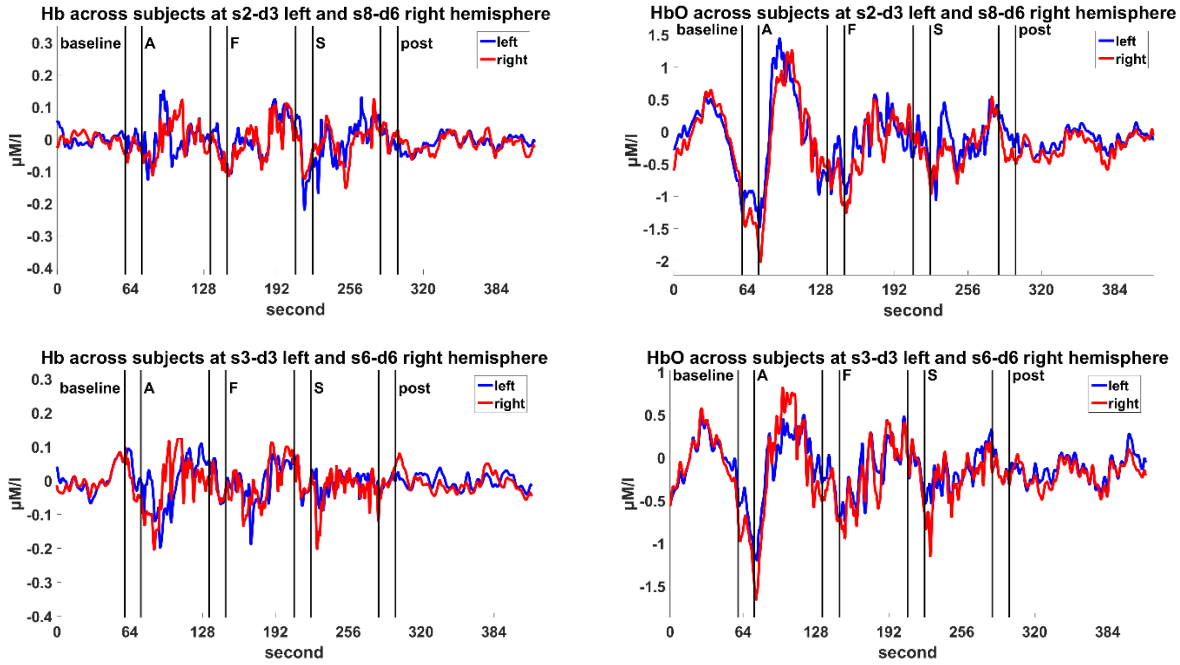


Figure 38. VFT experiment: mean across all subjects of Hb (left), HbO (right) changes between pairs of channels representative for left (blue) and right (red) hemispheres. Vertical lines indicate the end or onset of conditions. Conditions by order of appearing are baseline, letter A, letter F, letter S and post; small segments between vertical lines are instruction conditions. Pairs of channels in the first row are s2-d3 and s8-d6, and in the last row s3-d3 and s6-d6.

Table 25. VFT experiment: Mean and standard deviation across subjects of Hb and HbO changes during letter conditions (baseline correction applied) at channels s2-d3 and s3-d3, representative for the left hemisphere, and at channels s6-d6 and s8-d6, representative for the right hemisphere.

Left hemisphere					**Right hemisphere**				
channel		**lette A**	**letter F**	**letter S**	**channel**		**lette A**	**letter F**	**letter S**
s2_d3	Hb	-0.00 ± 0.04	0.01 ± 0.05	0.00 ± 0.05	s8_d6	Hb	0.01 ± 0.03	-0.00 ± 0.04	0.01 ± 0.08
	HbO	0.12 ± 0.18	-0.05 ± 0.18	-0.14 ± 0.17		HbO	0.08 ± 0.18	-0.15 ± 0.14	-0.17 ± 0.27
s3_d3	Hb	-0.01 ± 0.03	-0.00 ± 0.06	-0.01 ± 0.05	s6_d6	Hb	-0.01 ± 0.03	-0.00 ± 0.03	-0.03 ± 0.03
	HbO	-0.03 ± 0.19	-0.13 ± 0.17	-0.16 ± 0.12		HbO	-0.04 ± 0.13	-0.18 ± 0.17	-0.26 ± 0.20

The results showed that:

a) The interaction between condition and hemisphere was not significant for both Hb and HbO changes, in both pairs of channels s2-d3 and s8-d6, and s3-d3 and s6-d6. Same as that in the replication.

b) There were no significant differences between hemispheres in both types of concentration, in both pairs of channels. Same as that in the replication.

c) HbO changes in the letter A condition were **not** significantly different from that in the baseline condition in both hemispheres and in both pairs of channels, while it was the case in the replication.

HbO changes in the letter F condition were significantly **lower** than that in the baseline condition in both hemispheres and in the pair of channels s3-d3 and s6-d6, while that was significantly **higher** in the replication.

HbO changes in the letter S condition were significantly **lower** than that in the baseline condition in both hemispheres and in both pairs of channels, while that was significantly **higher** in the replication.

HbO changes in the letter A condition were significantly **higher** than that in the letter F and S conditions while there were no significant differences in HbO changes between these conditions in the replication.

There were no significant differences in HbO changes between the letter F and S conditions in the pair of channels s3-d3 and s6-d6 while that was the case in the replication.

Hb changes in the letter A and F condition were not significantly different from that in the baseline condition, in both hemispheres in the pair s3-d3 and s6-d6, while that was the case in the replication.

In visualization, the traces of HbO changes were similar in both hemispheres, in both pairs of channels s2-d3 and s8-d6, and s3-d3 and s6-d6 representative for left and right hemispheres. Same to that in the replication. The traces of Hb changes were similar in both hemispheres, in both pairs of channels, however that in the replication were less similar between hemispheres. The trends of HbO changes in this approach

Table 26. VFT experiment – proposed approach: Statistical results for pairs of channels representative for left and right hemispheres: s2-d3 and s8-d6, and s3-d3 and s6-d6. Bold p values are the corrected p values based on Greenhouse-Geisser correction, degrees of freedom are not corrected in that case, ns. stands for not significant.

Hemisphere			**RM ANOVA**			**Pairwise post-hoc t-tests**
Left	**Right**		**Effect**		**Greenhouse-Geisser epsilon**	
s2_d3	s8_d6	Hb	condition	F(3,27)=1.156, **p=0.324**	0.448	ns.
			hemisphere	F(1,9)=1.691, p=0.226		ns.
			interaction	F(3,27)=1.367, **p=0.276**	0.387	ns.
		HbO	condition	F(3,27)=9.781, p<.001*		S vs. baseline p=0.005*, A vs. F p<0.001*, A vs. S p=0.002*
			hemisphere	F(1,9)=0.856, p=0.379		ns.
			interaction	F(3,27)=2.009, **p=0.171**	0.579	ns.
s3_d3	s6_d6	Hb	condition	F(3,27)=1.171, **p=0.332**	0.659	ns.
			hemisphere	F(1,9)=0.003, p=0.959		ns.
			interaction	F(3,27)=0.847, **p=0.444**	0.659	ns.
		HbO	condition	F(3,27)=9.466, p<0.001*		F vs. baseline p=0.002*, S vs. baseline p<0.001*, A vs. F p=0.009*, A vs. S p=0.002*
			hemisphere	F(1,9)=0.587, p=0.463		ns.
			interaction	F(3,27)=2.29, **p=0.131**	0.658	ns.

4.4.3. Discussion and conclusion

The visualization and statistical test results in the proposed approach differed from that in the replication. In the proposed approach, the very slow drift was removed which made the traces of HbO and Hb changes across subjects differed from that in the replication.

By design, the proposed approach did minimize instrumental noise. However, similar to the case of CPT experiment, physiological noise was still present in the filtered signal and MAs have not been corrected as mentioned in earlier, the statistical results of the proposed approach on this dataset did not indicate better performance to that of the replication approach.

4.5. Conclusion

In summary, the proposed approach with high-pass and low-pass filters which by design minimized high frequency noise, the heart rate artifact and slow drift, gave smoother traces of HbO, Hb and tHb changes and comparable statistical results for the data of Visual stimulation, Finger tapping, and CPT experiments with all repetitions. In practice when a quick test with a smaller number of repetitions is required, the proposed approach should give a better results as showed in the case of Visual stimulation experiment when considering data of the first 2 repetitions only. The obtained results is with the exception of VFT due to the presence of physiological noise and MAs which needs special treatment as indicated in the methodology section 2. For this experiment, the application of the proposed approach should couple with another method for physiological noise and MAs.

In general, given a new dataset without much prior information, the proposed approach is recommended.

5. Spatial analysis

The first section in this chapter describes the solution approach for the research question of finding the locations of the strongest signal during stimulation. In the subsequent sections, the method and results for each experiment conducted at Philips Research are discussed.

In literature, an approach to create topographic views of measured locations was to use the mean of concentration changes during stimulation at each location as the indicator to quantify the most activated area via contour maps (Takahashi et al., 2000). A similar approach was used in this thesis. The positive parts of area under curve (AUC) of HbO changes during stimulation after baseline corrected were used and considered to represent the strength of activity. Instead of contour maps of the means of concentration changes, we use image segmentation. Details are provided in the section below.

5.1. Solution approach

Single subject level

In this work, the area under curve (AUC) of HbO changes in the task condition per channel after baseline corrected was considered to represent for the strength of the signal in the channel where AUC was made by areas of the curve HbO changes in the positive parts during task condition that was corresponding to the peaks during the task condition (Figure 39). This choice was based on the typical hemodynamic response that the increase of HbO changes was often larger than the decrease of Hb changes during brain activation (Wolf el al., 2001; Scholkmann et al., 2014). Besides, the positive parts of the HbO changes curve were corresponding to the increase of HbO changes comparing to baseline, hence we expected to filter out channels whose HbO changes were below baseline level, which we considered as not active.

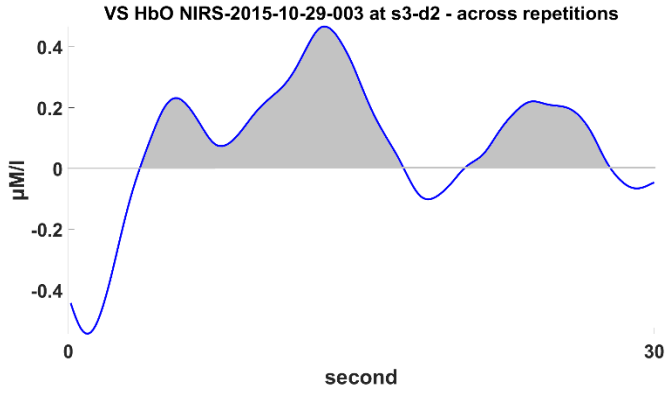


Figure 39. Blue curve presents the mean of HbO changes during Task condition at channel s3-d2 of one subject across repetitions in Visual stimulation experiment. AUC is defined as the sum of gray areas which are made by the blue curve and the positive part of x axis.

We used image segmentation to separate locations by considering the topology of the channels as a 2D gray image in which each channel corresponded to a pixel in the image and the value of each pixel was the representative value of the corresponding channel after scaling into a gray level. Thus, identifying the locations of strongest signal was equivalent to the locating of a group of pixels with highest gray intensity. In this work, it was performed by using the multi-thresholding version of the Otsu image clustering method (Otsu, 1979) with 2 thresholds to classify data into 3 groups in order to separate the strongest signal group better. Otsu's is a fast clustering method for gray image which minimizes the within-class

variance and maximizing the between-class variance, simultaneously. The Otsu method for two classes assumed the data following bimodal distribution and the sample sizes of the two groups were not too unbalanced. Besides, no spatial information was taken into account. However, real-world data do not necessary follow such conditions, in that case the threshold given by the two-group Otsu method is not satisfactory (Figure 40).

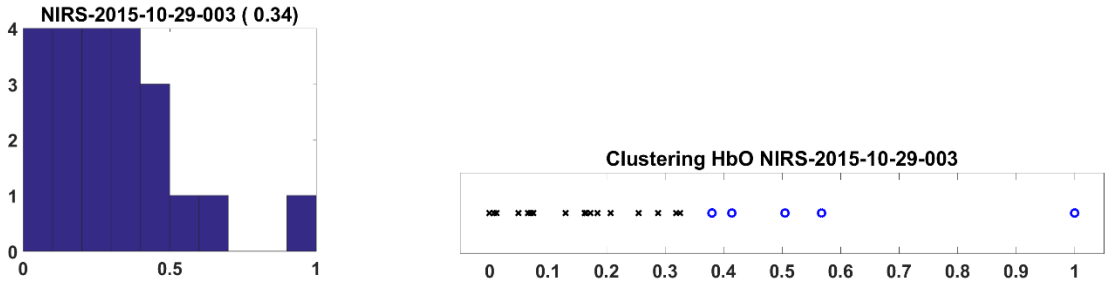


Figure 40. Otsu method does not work well with these data. Left: histogram of 22 grayscale values representative for 22 channels of a subject in Visual stimulation experiment. Right: two clusters based on the threshold 0.34 given by the Matlab's implementation of the Otsu method.

To apply the two-threshold Otsu method, the AUC values of all channels per subject were mapped to grayscale. The grayscale values in range [0, 1] were calculated using the formula

$$v_{gray} = \frac{v_{raw} - min_{raw}}{max_{raw} - min_{raw}} \quad (0)$$

where v_{gray} denotes the grayscale value, v_{raw} is the value before mapping, min_{raw} and max_{raw} are the minimum and maximum value of all channels' values before mapping, respectively.

As the result, channels in the group with highest values were considered as having strongest signal.

Group level

The biggest challenge in combining multiple subjects is to deal with differences in scale between subjects and differences in distribution within subject. In literature, values at subject level were normalized to reduce the differences between subjects (Huppert et al., 2006). In this work, a similar method was applied. The combination was performed after Otsu clustering at subject level. Per subject, high value channels were mapped to pixels with intensity 1 and the other as intensity 0. At group level, the combined grayscale image of multiple subjects was mapped from the sum of intensities of all subjects' binary images per pixel position. Similarly, the two-threshold version of Otsu method was applied again on the combined grayscale image to identify the highest value group for strong signal channels representative at group level.

5.2. Visual stimulation experiment

5.2.1. Method

The data were analyzed using Matlab (The MathWorks Inc., MA, USA, version R2015b) software.

In this analysis, the calculation of HbO changes follows the proposed approach described in chapter 4. Per trial, HbO changes in the task condition were considered after correcting baseline with the baseline was defined as the last 15 seconds of the rest condition preceding the task condition and the length of the task condition was normalized across subjects (see 3.1.2.2). Subsequently, the means of HbO changes

across trials were obtained for each subject, at each channel. AUC for each channel, per subject was calculated based on these means of HbO changes in the Task condition. For each subject, there were 22 AUC values corresponding to 22 channels. The analysis at single subject level and at group level follows the approach as described previously. Data of one subject were excluded due to motion artifacts which occurred in many repetitions and in many channels. In total there were data of 11 subjects.

5.2.2. Results

Single subject level

The results of one subject are shown as follows. Result of all 11 subjects are described in Appendix D. Per subject there are one graphs and two images. The first graph shows three groups of grayscale values based on the two thresholds as results of the two-threshold Otsu method. Red color indicates values which belong to the strongest group. Blue and black indicate the other groups by decreasing order of the values. The left image illustrates the topology of the channels in grayscale image and the right image illustrates the topology of the channels in 3 groups where red cell(s) represents the strongest location(s), gray and black cells represent the other locations by order of decreasing strength of signal; the number in each cell is the channel index. Three cells with horizontal stripes in the middle of each image indicate non-measured areas.

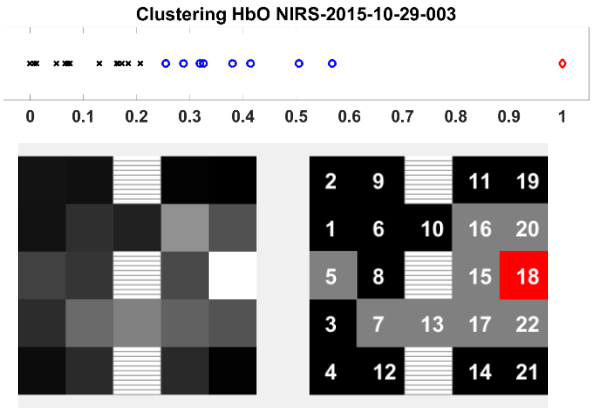
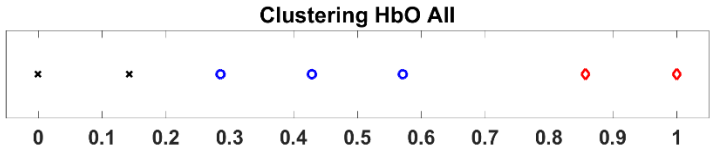


Figure 41. Visual stimulation experiment. Results of a subject. Top graph shows three groups of grayscale values as the results of the two-threshold Otsu method, red values indicate group with highest values; bottom-left image shows the topology of the channels in grayscale; bottom-right image shows the topology of the channels grouped by strength of signal, red cells illustrate locations of strongest signal, gray and black cells illustrate locations of weaker signal by decreasing order, each number represents the corresponding channel index. Cells with stripes indicate non-measured areas.

Group level

The results across subjects are shown as follows with the same presentation as that at single subject level.



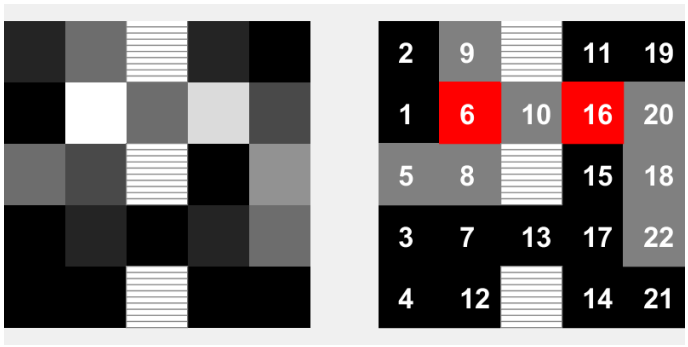


Figure 42. Visual stimulation experiment – Group level. Top: three groups of grayscale values as the results of two-threshold Otsu method, red color indicate group with highest values. Bottom left: topology of the channels in grayscale image. Bottom right: topology of the channels grouped by strength of signal, red cells illustrate locations of strongest signal, gray and black cells illustrate locations of weaker signal by decreasing order, each number represents the corresponding channel index. Cells with stripes indicate non-measured areas.

5.2.3. Discussion and conclusion

The results at single subject level showed that locations of strongest signal varied between subjects.

At group level, location corresponding to channels 6 (s3-d2) and 16 (s6-d5) have the strongest signal based on the data of 11 subjects. 8 subjects had s6-d5 as one of most activated locations and 6 subjects had s3-d2. These locations are near or overlay the EEG position O_1 and O_2 which match with the reported locations from Takahashi et al. (2000) and Wijeakumar et al. (2012).

In some subjects' results, for example NIRS-2015-10-29-003, NIRS-2015-11-09-002 and NIRS-2015-11-10-003, the locations of strongest signal were not in the neighbor of O_1 and O_2, particularly the last one. It was suggested that the positions of the optodes might not be at the desired locations.

5.3. Finger tapping experiment

5.3.1. Method

The data were analyzed using Matlab (The MathWorks Inc., MA, USA, version R2015b) software.

In this analysis, the calculation of HbO changes follows the proposed approach described in chapter 4. The tapping condition at maximum effort was considered since only at this frequency the increase of HbO changes from baseline was significant. Per trial, HbO changes in the tapping condition were considered after correcting baseline with the baseline was defined as the last 15 seconds of the rest condition preceding the task condition and the lengths of the tapping condition were normalized across 3 trials and across subjects (see 3.2.2.2). Subsequently, the means of HbO changes across 3 trials were obtained for each subject, at each channel. AUC for each channel, per subject was calculated based on these means of HbO changes in the tapping condition. For each subject, there were 20 AUC values corresponding to 20 channels. The analysis at single subject level and at group level follows the steps described previously. Data of all 5 subjects were taken in this analysis.

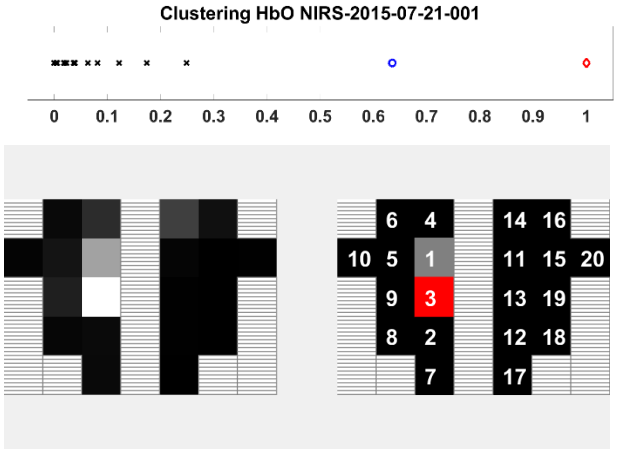
To access the significant difference between left and right hemispheres during the finger tapping condition at single subject level, a paired t-test was used to test the mean of HbO changes of the tapping condition per subject at 7 channels around EEG position C3 in the left hemispheres and their symmetric channels on the right hemisphere. Maximum effort was considered since only at this frequency the increase of HbO changes from baseline was significant. The adjusted alpha level was set at 5%. The

selected channels in the left hemisphere were 10, 5, 1, 9, 3, 8 and 2 (s4-d4, s2-d3, s1-d1, s4-d3, s1-d3, s3-d3 and s1-d2) and the corresponding channels in the right hemisphere 20, 15, 11, 19, 13, 18 and 12 (s8-d8, s6-d7, s5-d5, s8-d7, s5-d7, s7-d7 and s5-d6) (see Figure 11 for the topology of the channels in this experiment).

5.3.2. Results

Single subject level

The results of a subject are shown as follows. Results of all 5 subjects are described in Appendix E. Per subject there are one graphs and two images. The first graph shows three groups of grayscale values based on the two thresholds as results of the two-threshold Otsu method. Red color indicates values which belong to the strongest group. Blue and black indicate the other groups by decreasing order of the values. The left image illustrates the topology of the channels in grayscale image and the right image illustrates the topology of the channels in 3 groups where red cell(s) represents the strongest location(s), gray and black cells represent the other locations by order of decreasing strength of signal; the number in each cell is the channel index. Cells with horizontal stripes each image indicate non-measured areas. The result of paired t-test was also shown for each subject for easily referencing.

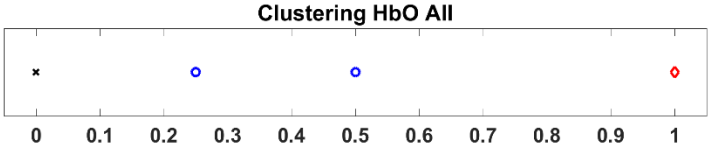


Paired *t*-test result: the mean of HbO changes of 7 channels in the left hemisphere was significantly higher than that in the right hemisphere (t=3.2221, p=0.0181).

Figure 43. Finger tapping experiment. Results of a subject. Top graph shows three groups of grayscale values as the results of the two-threshold Otsu method, red values indicate group with highest values; bottom-left image shows the topology of the channels in grayscale; bottom-right image shows the topology of the channels grouped by strength of signal, red cells illustrate locations of strongest signal, gray and black cells illustrate locations of weaker signal by decreasing order, each number represents the corresponding channel index. Cells with stripes indicate non-measured areas.

Group level

The results across subjects are shown as follows with the same presentation as that at single subject level.



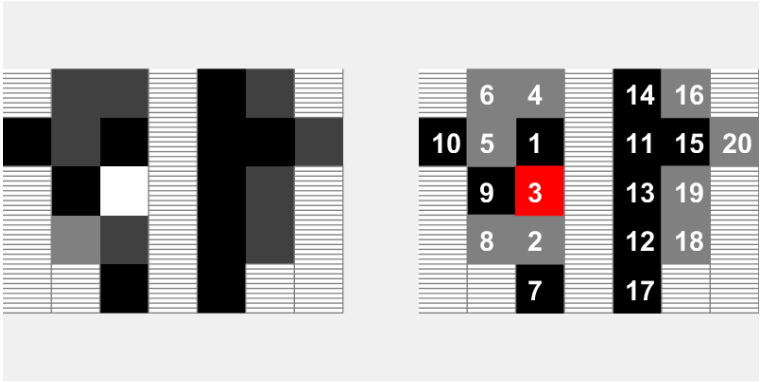


Figure 44. Finger tapping experiment – Group level. Top: three groups of grayscale values as the results of the two-threshold Otsu method, red values indicate group with highest values; Bottom-left: topology of the channels in grayscale; Bottom-right: topology of the channels grouped by strength of signal, red cells illustrate locations of strongest signal, gray and black cells illustrate locations of weaker signal by decreasing order, each number represents the corresponding channel index. Cells with stripes indicate non-measured areas.

5.3.3. Discussion and conclusion

The results at single subject level showed that locations of strongest signal varied between subjects. The strongest location of subject NIRS-2015-07-21-002 was at channel 20, on the right hemisphere while it was expected on the left hemisphere as all subjects were reported as right-handed however there was no archived information to verify if that was an experimental error.

Despite the small number of subjects in the data, 4 over 5 subjects had the strongest signal at channel 3 (s1-d3), which was near the EEG position C3, and that matched with the location reported in Kuboyama et al. (2004).

5.4. Continuous performance test (CPT) experiment

5.4.1. Method

The data were analyzed using Matlab (The MathWorks Inc., MA, USA, version R2015b) software.

In this analysis, the calculation of HbO changes follows the proposed approach described in chapter 4. Per subject, HbO changes in the whole CPT condition were considered after correcting baseline wherein the baseline was defined as the last 15 seconds of the pre-task condition and the length of the CPT condition was normalized across subjects (see 3.3.2.2). AUC for each channel, per subject was calculated based on the HbO changes in the CPT condition. For each subject, there were 20 AUC values corresponding to 20 channels. The analysis at single subject level and at group level follows the approach as described at the beginning of this chapter. Data of all 10 subjects were taken in this analysis wherein the data at channels 7 (s3-d4), 9 (s4-d4), 12 (s5-d4) and 14 (s6-d4) of all subjects were excluded since the signal at these channels was noisy.

5.4.2. Results

Single subject level

The results of a subject are shown as follows. Results of all 10 subject are described in Appendix F. Per subject there are one graphs and two images. The first graph shows three groups of grayscale values based on the two thresholds as results of the two-threshold Otsu method. Red color indicates values which belong to the strongest group. Blue and black indicate the other groups by decreasing order of the values.

The left image illustrates the topology of the channels in grayscale image and the right image illustrates the topology of the channels in 3 groups where red cell(s) represents the strongest location(s), gray and black cells represent the other locations by order of decreasing strength of signal; the number in each cell is the channel index. Three cells with horizontal stripes in the middle of each image indicate non-measured areas.

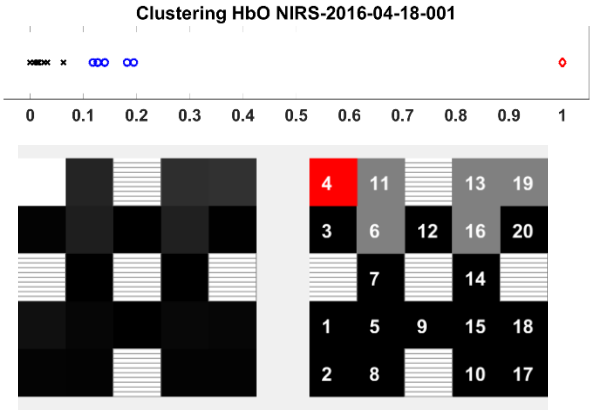


Figure 45. CPT experiment. Results of a subject. Top graph shows three groups of grayscale values as the results of the two-threshold Otsu method, red values indicate group with highest values; bottom-left image shows the topology of the channels in grayscale; bottom-right image shows the topology of the channels grouped by strength of signal, red cells illustrate locations of strongest signal, gray and black cells illustrate locations of weaker signal by decreasing order, each number represents the corresponding channel index. Cells with stripes indicate non-measured areas.

Group level

The results across subjects are shown as follows with the same presentation as that at single subject level.

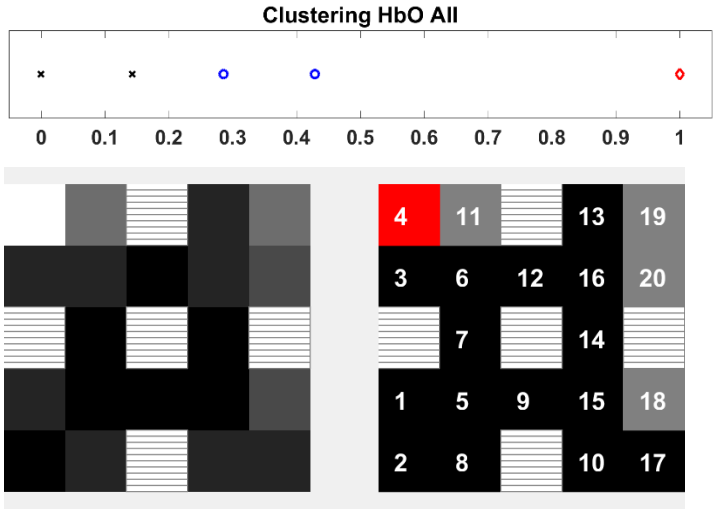


Figure 46. CPT experiment – Group level. Top: three groups of grayscale values as the results of the two-threshold Otsu method, red values indicate group with highest values; Bottom-left: topology of the channels in grayscale; Bottom-right: topology of the channels grouped by strength of signal, red cells illustrate locations of strongest signal, gray and black cells illustrate locations of weaker signal by decreasing order, each number represents the corresponding channel index. Cells with stripes indicate non-measured areas.

5.4.3. Discussion and conclusion

The results at single subject level showed that locations of strongest signal varied between subjects. At group level, location corresponding to channel 4 (s2-d3) had the strongest signal based on the data of 10 subjects. 7 subjects had this channel as one of the most activated locations. The location is in the left hemisphere while the results in Fallgatter and Strik (1997) showed brain activation in the right hemisphere. As discussed in the replicated result of CPT experiment (see 3.3.2.4), motion artifacts and probably experimental errors might have influenced to the measured data of some subjects in this experiment.

5.5. Verbal-fluency test (VFT) experiment

5.5.1. Method

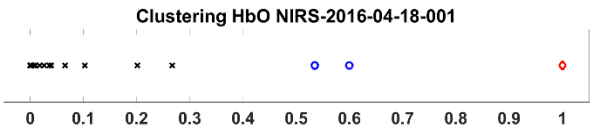
The data were analyzed using Matlab (The MathWorks Inc., MA, USA, version R2015b) software.

In this analysis, the calculation of HbO changes follows the proposed approach described in chapter 4. Per subject, HbO changes in the three letter conditions A, F and S were considered after correcting baseline wherein the baseline was defined as the last 15 seconds of the baseline condition and the lengths of the letter conditions were normalized across three letter conditions and across subjects (see 3.4.2.2). AUC for each channel, per subject was calculated based on the mean of HbO changes across three letter conditions. For each subject, there were 20 AUC values corresponding to 20 channels. The analysis at single subject level and at group level follows the approach as described at the beginning of this chapter. Data of all 10 subjects were taken in this analysis wherein the data at channels 7 (s3-d4), 9 (s4-d4), 12 (s5-d4) and 14 (s6-d4) of all subjects were excluded since the signal at these channels was noisy. Data at channel s8-d6 of one subject which contains many spikes were also excluded.

5.5.2. Results

Single subject level

The results of a subject are shown as follows. Results of all 10 subjects are described in Appendix G. Per subject there are one graphs and two images. The first graph shows three groups of grayscale values based on the two thresholds as results of the two-threshold Otsu method. Red color indicates values which belong to the strongest group. Blue and black indicate the other groups by decreasing order of the values. The left image illustrates the topology of the channels in grayscale image and the right image illustrates the topology of the channels in 3 groups where red cell(s) represents the strongest location(s), gray and black cells represent the other locations by order of decreasing strength of signal; the number in each cell is the channel index. Three cells with horizontal stripes in the middle of each image indicate non-measured areas.



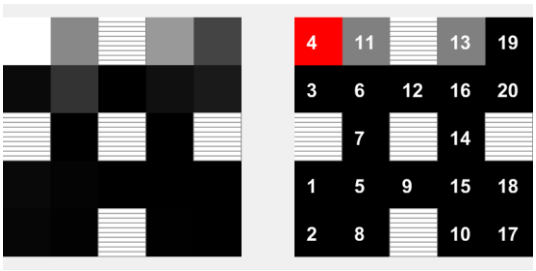


Figure 47. VFT experiment. Results of a subject. Top graph shows three groups of grayscale values as the results of the two-threshold Otsu method, red values indicate group with highest values; bottom-left image shows the topology of the channels in grayscale; bottom-right image shows the topology of the channels grouped by strength of signal, red cells illustrate locations of strongest signal, gray and black cells illustrate locations of weaker signal by decreasing order, each number represents the corresponding channel index. Cells with stripes indicate non-measured areas.

Group level

The results across subjects are shown as follows with the same presentation as that at single subject level.

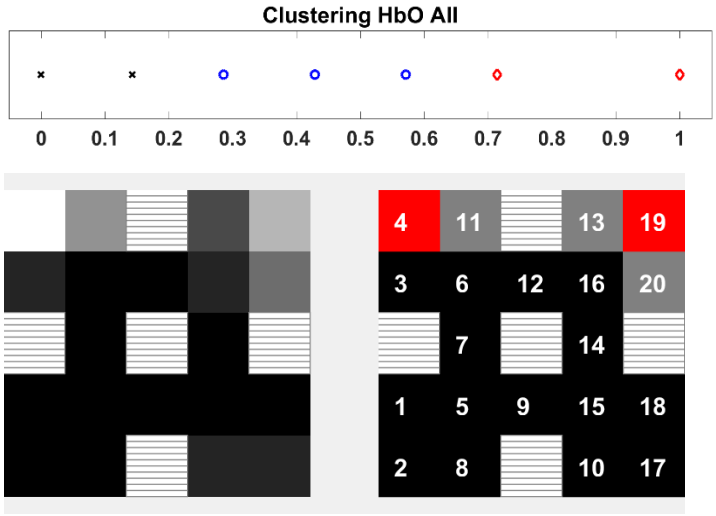


Figure 48. VFT experiment – Group level. Top: three groups of grayscale values as the results of the two-threshold Otsu method, red values indicate group with highest values; Bottom-left: topology of the channels in grayscale; Bottom-right: topology of the channels grouped by strength of signal, red cells illustrate locations of strongest signal, gray and black cells illustrate locations of weaker signal by decreasing order, each number represents the corresponding channel index. Cells with stripes indicate non-measured areas.

5.5.3. Discussion and conclusion

The results at single subject level showed that locations of strongest signal were around the anterior brain area in both hemispheres. At group level, locations corresponding to channels 4 (s2-d3) and channel 19 (s8-d6), in the left and right hemispheres respectively, had the strongest signal based on the data of 10 subjects. 7 subjects had channel s2-d3 and 6 subjects had channel s8-d6 as one of the most activated locations. Although these locations were in line with the results in Herrmann et al. (2003) that activation occurred in both hemispheres near the EEG positions between Fp1 and F7/F3 and between Fp2 and F4/F8 in the left and right hemispheres respectively, motion artifacts might have greatly influenced to the measured data of some subjects in this experiment as discussed in chapter 3.4.2.4, thus the results in this analysis should be considered with caution.

5.6. Conclusion

Locations of strongest signal for each experiment have been identified at single subject level and at group level based on HbO changes. The results at group level of the Visual stimulation were around O_1 and O_2 and that of Finger tapping experiment was around C3, which agreed with the reported locations in literature. The results at group level of the VFT experiment were in line with the results in Herrmann et al. (2003) that activation occurred in both hemispheres near the EEG positions between Fp1 and F7/F3 and between Fp2 and F4/F8 in the left and right hemispheres respectively. However these should be taken with caution as motion artifacts might have greatly influenced to the measured data of some subjects in this experiment. The result at group level of the CPT experiment was around location Fp1 in the left hemisphere while the results in Fallgatter and Strik (1997) showed brain activation in the right hemisphere. As discussed in the replicated result of CPT experiment (see 3.3.2.4), motion artifacts and probably experimental errors might have influenced to the measured data of some subjects in this experiment.

The results at single subjects were not consistent. This again confirmed the similar results in other studies about the reproducibility of fNIRS (Plichta et al., 2006; Plichta et al., 2007; Kono et al., 2007; Schecklmann et al., 2008; Biallas et al., 2012).

6. Conclusions

The topic of this thesis is to preprocess and analyze fNIRS data of four experiments conducted at Philips Research which were available prior this work. We first evaluated the reproducibility of the reported methods in literature on these datasets. The results of the literature corresponding to the Visual stimulation and the Finger tapping experiments were generally reproducible. For the VFT experiment, although the traces of HbO and Hb changes across subjects in the replicated data showed strong effect of motion artifacts, the results in literature of the VFT experiment were also considered as reproducible. However for the CPT experiment, the observed data suggested that experimental errors might have occurred such as the unstable of concentration changes in baseline condition, the existence of two 15-second instruction conditions in between the desired conditions and motion artifacts in the data of some subjects led to unexpected concentration changes in the CPT task. This may hinder the reproducibility of the results in literature of this experiment.

Subsequently, an approach to preprocess and analyze fNIRS data was proposed using low-pass filter with cutoff 0.5 Hz, high-pass filter with cutoff 0.01 Hz and averaging trials which helped removing high frequency noise, the cardiac artifact, the very low frequency components and reducing further the variance of other physiological components. In general, given a new dataset without much prior information, the proposed approach is recommended as it gave smoother traces of HbO, Hb and tHb and comparable statistical results for the data of Visual stimulation, Finger tapping, and CPT experiments. The proposed method is also recommended for a quick test with a smaller number of trials as in the case of Visual stimulation experiment when considering data of the first 2 trials only.

Finally, an imaging segmentation method was used to identify the locations of strongest signal. A new method was proposed to identify locations of strongest signal for each experiment at single subject level and at group level based on HbO changes. The analysis for locations of strongest signal for each experiment was highly agreement with the reported locations in literature at group level as showed in the results of the Visual stimulation experiment which were around position O_1 and O_2 and of the Finger tapping experiment around C3. Furthermore, the results at single subjects were not consistent and this again confirmed the similar results in other studies about the reproducibility of fNIRS (Plichta et al., 2006; Plichta et al., 2007; Kono et al., 2007; Schecklmann et al., 2008; Biallas et al., 2012). The difference between left and right hemispheres in the Finger tapping experiment at single subject level agreed with the locations of strongest signal. As a general issue in data with MA presence, the results of the CPT and VFT should be taken with caution since MAs have not been corrected.

For future work, the correction for baseline shift of MAs should be investigated. A prominent approach in reducing physiological noise and also MAs is using reference channels (also called short separation channels) wherein besides the standard channels whose source-detector optode distance is around 3 cm, additional channels are set whose source-detector optode distance is less than 1 cm such that the channels are sensitive only to signal in the superficial layers. The signals in these reference channels could be used to subtract noise from the signal in the standard channels (Gagnon et al., 2011; Zhang et al., 2010). Another approach that is also worth to be investigated is using additional measure of physiological noise such as respiration, heart rate and blood pressure (Kirilina et al., 2013; Patel, Katura, Maki, & Tachtsidis, 2011).

As recommendation for future experiments, MAs should be prevented or reduced with a proper way to secure the optodes on the head which takes into account the comfortability for each group of subjects

(e.g. cap is not too tight, children are difficult to sit still) and the motions necessary in the experiment (e.g. head movement, jaw movement) (Huppert et al., 2009). Furthermore, a prototype should be implemented and tested thoroughly to optimize the steps and to determine sufficient preprocessing and analyzing methods for the purpose of the experiment.

7. Acknowledgements

I would like to thank my supervisors at Philips Research, Marian Dekker and Linda Bremer - van der Heiden, and at the Radboud University Nijmegen, Jason Farquhar and Tom Heskes, for their suggestions, feedback, help and support during my research. I would also like to thank Fons de Lange, Lennaert Bronts, Maurice Stassen, Sabina Manzari, Paul Lemmens and Ad Denissen for their help and feedback on my work. I thank my friends Georgiana, Hanieh, Laura, Henk, Andra, Ada and Ivy for sharing their experience, knowledge and feedback which has helped me to improve my work. I am also very grateful to my husband and our children for their endless help and support.

Bibliography

Bauernfeind, G., Wriessnegger, S. C., Daly, I., & Müller-Putz, G. R. (2014). Separating heart and brain: on the reduction of physiological noise from multichannel functional near-infrared spectroscopy (fNIRS) signals. *Journal of Neural Engineering*, *11*(5), 056010. http://doi.org/10.1088/1741-2560/11/5/056010

Biallas, M., Trajkovic, I., Haensse, D., Marcar, V., & Wolf, M. (2012). Reproducibility and sensitivity of detecting brain activity by simultaneous electroencephalography and near-infrared spectroscopy. *Experimental Brain Research*, *222*(3), 255–264. http://doi.org/10.1007/s00221-012-3213-6

Boas, D. a, Gaudette, T., Strangman, G., Cheng, X., Marota, J. J., & Mandeville, J. B. (2001). The accuracy of near infrared spectroscopy and imaging during focal changes in cerebral hemodynamics. *NeuroImage*, *13*(1), 76–90. http://doi.org/10.1006/nimg.2000.0674

Brigadoi, S., Ceccherini, L., Cutini, S., Scarpa, F., Scatturin, P., Selb, J., … Cooper, R. J. (2014). Motion artifacts in functional near-infrared spectroscopy: A comparison of motion correction techniques applied to real cognitive data. *NeuroImage*, *85*, 181–191. http://doi.org/10.1016/j.neuroimage.2013.04.082

Bunge, S. A., & Kahn, I. (2010). Cognition: An Overview of Neuroimaging Techniques. *Encyclopedia of Neuroscience*, *2*, 1063–1067. http://doi.org/10.1016/B978-008045046-9.00298-9

Cope, M., & Delpy, D. T. (1988). System for long-term measurement of cerebral blood and tissue oxygenation on newborn infants by near infra-red transillumination. *Medical & Biological Engineering & Computing*, *26*(3), 289–294. http://doi.org/10.1007/BF02447083

Cope, M. (1991). The application of near infrared spectroscopy to non invasive monitoring of cerebral oxygenation in the newborn infant. *PhD Thesis*, Dept. of Medical Physics and Bioengineering, University College London.

Cui, X., Bray, S., & Reiss, A. L. (2010). Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. *NeuroImage*, *49*(4), 3039–3046. http://doi.org/10.1016/j.neuroimage.2009.11.050

Cui, X., Bray, S., Bryant, D. M., Glover, G. H., & Reiss, A. L. (2011). A quantitative comparison of NIRS and fMRI across multiple cognitive tasks. *NeuroImage* *54*(4), 2808–2821. http://doi.org/10.1016/j.neuroimage.2010.10.069

Cutini, S., Moro, S. B., & Bisconti, S. (2012). Review: Functional near infrared optical imaging in cognitive neuroscience: An introductory review. *Journal of Near Infrared Spectroscopy*, *20*(1), 75–92. http://doi.org/10.1255/jnirs.969

Delpy, D. T., Cope, M., van der Zee, P., Arridge, S., Wray, S., & Wyatt, J. (1988). Estimation of optical pathlength through tissue from direct time of flight measurement. *Physics in Medicine and Biology*, *33*(12), 1433–1442. http://doi.org/10.1088/0031-9155/33/12/008

Duncan, a, Meek, J. H., Clemence, M., Elwell, C. E., Tyszczuk, L., Cope, M., & Delpy, D. T. (1995). Optical pathlength measurements on adult head, calf and forearm and the head of the newborn infant using phase resolved optical spectroscopy. *Physics in Medicine and Biology*, *40*(2), 295–304. http://doi.org/10.1088/0031-9155/40/2/007

Duncan, A., Meek, J.H., Clemence, M., Elwell, C.E., Fallon, P., Tyszczuk, L., Cope, M., & Delpy, D.T. (1996). Measurement of cranial optical path length as a function of age using phase resolved near infrared spectroscopy. *Pediatric Research*, 39, 889–894. doi:10.1203/00006450-199605000-00025

Essenpreis, M., Elwell, C. E., Cope, M., van der Zee, P., Arridge, S. R., & Delpy, D. T. (1993). Spectral dependence of temporal point spread functions in human tissues. *Applied Optics*, *32*(4), 418–425. http://doi.org/10.1364/AO.32.000418

Fallgatter, A. J., & Strik, W. K. (1997). Right frontal activation during the continuous performance test assessed with near-infrared spectroscopy in healthy subjects. *Neuroscience Letters*, *223*(2), 89–92. http://doi.org/10.1016/S0304-3940(97)13416-4

Fekete, T., Rubin, D., Carlson, J. M., & Mujica-Parodi, L. R. (2011). The nirs analysis package: Noise reduction and statistical inference. *PLoS ONE*, *6*(9). http://doi.org/10.1371/journal.pone.0024322

Ferrari, M., & Quaresima, V. (2012). A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *NeuroImage*, *63*(2), 921–935. http://doi.org/10.1016/j.neuroimage.2012.03.049

Gagnon, L, Perdue K, Greve DN, Goldenholz D, Kaskhedikar G, Boas DA. Improved recovery of the hemodynamic response in diffuse optical imaging using short optode separations and state-space modeling. *NeuroImage*. 2011; 56 (3):1362–1371. [PubMed: 21385616]

Gagnon, L., Yucel, M.A., Dehaes, M., Cooper, R.J., Perdue, K.L., Selb, J., Huppert, T.J., Hoge, R.D., Boas, D.A. (2012). Quantification of the cortical contribution to the NIRS signal over the motor cortex using concurrent NIRS-fMRI measurements. *NeuroImage* 59, 3933–3940.

Herrmann, M. J., Ehlis, A. C., & Fallgatter, A. J. (2003). Frontal activation during a verbal-fluency task as measured by near-infrared spectroscopy. *Brain Research Bulletin*, *61*(1), 51–56. http://doi.org/10.1016/S0361-9230(03)00066-2

Huppert T.J., Hoge R.D., Diamond S.G., Franceschini M.A., Boas, D. A. (2006). A temporal comparison of BOLD, ASL, and NIRS hemodynamic responses to motor stimuli in adult humans. *NeuroImage*, *29*(2), 368–382. http://doi.org/10.1016/j.neuroimage.2005.08.065

Huppert, T. J., Diamond, S. G., Franceschini, M. A., & Boas, D. A. (2009). HomER: a review of time-series analysis methods for near-infrared spectroscopy of the brain. *Applied Optics*, *48*(10), 280–298.

Irani, F., Platek, S. M., Bunce, S., Ruocco, A. C., & Chute, D. (2007). Functional near infrared spectroscopy (fNIRS): an emerging neuroimaging technology with important applications for the study of brain disorders. *The Clinical Neuropsychologist*, *21*(1), 9–37. http://doi.org/10.1080/13854040600910018

Keles, H. O., Barbour, R. L., & Omurtag, A. (2016). Hemodynamic correlates of spontaneous neural activity measured by human whole-head resting state EEG + fNIRS. *NeuroImage*, *138*, 76–87. http://doi.org/10.1016/j.neuroimage.2016.05.058

Kittler, J., & Illingworth, J. (1985). On threshold selection using clustering criteria. *Systems, Man and Cybernetics, IEEE Transactions on*, (5), 652–655. http://doi.org/10.1109/TSMC.1985.6313443

Kono, T., Matsuo, K., Tsunashima, K., Kasai, K., Takizawa, R., Rogers, M. A., … Kato, N. (2007). Multiple-time replicability of near-infrared spectroscopy recording during prefrontal activation task in healthy men. *Neuroscience Research*, *57*(4), 504–512. http://doi.org/10.1016/j.neures.2006.12.007

Kuboyama, N., Nabetani, T., Shibuya, K.-I., Machida, K., & Ogaki, T. (2004). The effect of maximal finger tapping on cerebral activation. *Journal of Physiological Anthropology and Applied Human Science*, *23*(4), 105–10. http://doi.org/JST.JSTAGE/jpa/23.105 [pii]

Lawrence, M. A. (2015). ez: Easy Analysis and Visualization of Factorial Experiments. R package version 4.3. https://CRAN.R-project.org/package=ez

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, a. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843), 150–7. http://doi.org/10.1038/35084005

Moaveni, M.K. (1970). A Multiple Scattering Field Theory Applied to Whole Blood. *PhD Thesis*, Dept. of Electrical Engineering, University of Washington.

Obrig, H. (2014). NIRS in clinical neurology - a "promising" tool? *NeuroImage*, *85*, 535–546. http://doi.org/10.1016/j.neuroimage.2013.03.045

O'Haver, T. (2016). A Pragmatic Introduction to Signal Processing. Retrieved from https://terpconnect.umd.edu/~toh/spectrum/IntroToSignalProcessing.pdf

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, *9*(1), 62–66. http://doi.org/10.1109/TSMC.1979.4310076

Patel, S., Katura, T., Maki, A., & Tachtsidis, I. (2011). Quantification of systemic interference in optical topography data during frontal lobe and motor cortex activation: An independent component analysis. *Advances in Experimental Medicine and Biology*, *701*, 45–51. http://doi.org/10.1007/978-1-4419-7756-4-7

Plichta, M. M., Herrmann, M. J., Baehne, C. G., Ehlis, A. C., Richter, M. M., Pauli, P., & Fallgatter, A. J. (2006). Event-related functional near-infrared spectroscopy (fNIRS): Are the measurements reliable? *NeuroImage*, *31*(1), 116–124. http://doi.org/10.1016/j.neuroimage.2005.12.008

Plichta, M. M., Herrmann, M. J., Baehne, C. G., Ehlis, A. C., Richter, M. M., Pauli, P., & Fallgatter, A. J. (2007). Event-related functional near-infrared spectroscopy (fNIRS) based on craniocerebral correlations: Reproducibility of activation? *Human Brain Mapping*, *28*(8), 733–741. http://doi.org/10.1002/hbm.20303

Scarpa, F., Cutini, S., Scatturin, P., Dell'Acqua, R., & Sparacino, G. (2010). Bayesian filtering of human brain hemodynamic activity elicited by visual short-term maintenance recorded through functional near-infrared spectroscopy (fNIRS). *Optics Express*, *18*(25), 26550–26568. http://doi.org/10.1364/OE.18.026550

Schecklmann, M., Ehlis, A. C., Plichta, M. M., & Fallgatter, A. J. (2008). Functional near-infrared spectroscopy: A long-term reliable tool for measuring brain activity during verbal fluency. *NeuroImage*, *43*(1), 147–155. http://doi.org/10.1016/j.neuroimage.2008.06.032

Schmitt, J. M. (1986). Optical measurement of blood oxygenation by implantable telemetry. *Technical Report* G558-15, Stanford (as compiled and posted by Prahl, S. A. on http://omlc.org/spectra/hemoglobin/moaveni.html).

Scholkmann, F., Spichtig, S., Muehlemann, T., & Wolf, M. (2010). How to detect and reduce movement artifacts in near-infrared imaging using moving standard deviation and spline interpolation. *Physiological Measurement*, *31*(5), 649–662. http://doi.org/10.1088/0967-3334/31/5/004

Scholkmann, F., & Wolf, M. (2012). Measuring brain activity using functional near infrared spectroscopy : a short review. *Spectroscopy Europe*, *24*(4), 6 – 10.

Scholkmann, F., Kleiser, S., Metz, A. J., Zimmermann, R., Mata Pavia, J., Wolf, U., & Wolf, M. (2014). A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *NeuroImage*, *85*, 6–27. http://doi.org/10.1016/j.neuroimage.2013.05.004

Smith, M. (2011). Shedding light on the adult brain: a review of the clinical applications of near-infrared spectroscopy. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, *369*(1955), 4452–69. http://doi.org/10.1098/rsta.2011.0242

Spichtig, S., Scholkmann, F., Chin, L., Lehmann, H., & Wolf, M. (2012). Assessment of intermittent UMTS electromagnetic field effects on blood circulation in the human auditory region using a near-infrared system. *Bioelectromagnetics*, *33*(1), 40–54. http://doi.org/10.1002/bem.20682

Takahashi, K., Ogata, S., Atsumi, Y., Yamamoto, R., Shiotsuka, S., Maki, A., … Igawa, M. (2000). Activation of the visual cortex imaged by 24 channel near-infrared spectroscopy. *Journal of Biomedical Optics*, *5*(1), 93–96.

Takatani, S., & Graham, M. D. (1979). Theoretical analysis of diffuse reflectance from a two-layer tissue model. *IEEE Transactions on Bio-Medical Engineering*, *26*(12), 656–664. http://doi.org/10.1109/TBME.1979.326455

Villringer, A, & Chance, B. (1997). Non-invasive optical spectroscopy and imaging of human brain function. *Trends in Neurosciences*, *20*(10), 435–442. http://doi.org/10.1016/S0166-2236(97)01132-6

Wijeakumar, S., Shahani, U., Simpson, W. A., & McCulloch, D. L. (2012). Localization of hemodynamic responses to simple visual stimulation: An fNIRS study. *Investigative Ophthalmology and Visual Science*, *53*(4), 2266–2273. http://doi.org/10.1167/iovs.11-8680

Wilcox, T., Bortfeld, H., Woods, R., Wruck, E., & Boas, D. A. (2005). During Object Processing in Infants. *Changes*, *10*(1), 1–16.

Wolf, M., Wolf, U., Toronov, V., Michalos, A., Paunescu, L. A., Choi, J. H., & Gratton, E. (2002). Different time evolution of oxyhemoglobin and deoxyhemoglobin concentration changes in the visual and motor cortices during functional stimulation: a near-infrared spectroscopy study. *Neuroimage*, *16*(3 Pt 1), 704–712. http://doi.org/S1053811902911286 [pii]

Zhang, Y., Brooks, D. H., Franceschini, M. A., & Boas, D. A. (2005). Eigenvector-based spatial filtering for reduction of physiological interference in diffuse optical imaging. *Journal of Biomedical Optics*, *10*(1), 11014. http://doi.org/10.1117/1.1852552

Zhang, Q., Strangman, G. E., & Ganis, G. (2010). Adaptive filtering to reduce global interference in non-invasive NIRS measures of brain activation: How well and when does it work?, *48*(Suppl 2), 1–6. http://doi.org/10.1097/MPG.0b013e3181a15ae8.Screening

Yücel, M. A., Selb, J., Boas, D. A., Cash, S. S., & Cooper, R. J. (2014). Reducing motion artifacts for long-term clinical NIRS monitoring using collodion-fixed prism-based optical fibers. *NeuroImage*, *85*, 192–201. http://doi.org/10.1016/j.neuroimage.2013.06.054

Software

ezANOVA function of the ez package in R, version 4.3. Lawrence, M. A., 2015.

HOMER2 software, function hmrBandpassFilt. Huppert, T. J., Diamond, S. G., Franceschini, M. A., & Boas, D. A. (2009). HomER: a review of time-series analysis methods for near-infrared spectroscopy of the brain. *Applied Optics*, *48*(10), 280–298.

Matlab software, version R2015b. The MathWorks Inc., MA, USA.

nirsLAB version 2016.01. NIRx Medical Technologies, LLC.

R software, 2016. R Core Team.

Appendix A. Data files

CW NIRSport device of NIRx Medical Technologies, LLC. was used to conduct the experiments at the wavelength 760 nm and 850 nm. The measurements were performed using the software NIRStar (version 14.1, NIRx Medical Technologies, LLC.), which saved the data of each subjects in a separate directory in several files. All the files which were used in the preprocessing and analyzing in this thesis are described as follows.

Table 27. Files from the data of each subject which are used for preprocessing and analyzing

Intensities files	*.wl1, .wl2*	*Raw data, intensities received at detector optodes*
Configuration file	*_config.txt*	*Configuration of the measurement*
Event log file	*.evt*	*Stimuli*
Information file	*.inf*	*Subject demographics*
Header file	*.hdr*	*Channels*
Probe info file	*_probeInfo.mat*	*Probes setup*
Topology file	*.tpl*	*Topology of channels*

Intensities received at detector optodes are saved in ASCII format in two files $.wl1$ and $.wl2$ corresponding to two wavelengths 760 nm and 850 nm, respectively. There are 64 columns corresponding to 64 channels made up by 8 sources and 8 detector optodes however depending on the setting when the experiment was conducted, the number of active channels may be smaller. Information of active channels is stored in the $.hdr$ file.

Configuration of the measurement such as sampling rate, number of sources and detector optodes, is saved in ._config.txt file.

Each row in the event file $.evt$ contains the index of the sample when the stimulus began and the corresponding event ID. The event ID is in binary format with lower bits first. For example, the first line below shows event 2 = 0000 0010 (binary) which begins at sample 1173, the second line shows event 3 = 0000 0011 (binary) which begins at sample 1330. The event IDs are from the experiment application which are used to differentiate conditions.

```
1173    0       1       0       0       0       0       0       0
1330    1       1       0       0       0       0       0       0
```

Subject demographics such as age, gender, are saved in .inf file.

Information about the measurement is stored in $.hdr$ file, such as wavelength in [ImagingParameters]; event ID of stimuli and the corresponding sample index, time index in [Markers]; the active channels in [DataStructure]; the distance between source and detector of each valid channel [ChannelsDistance].

Information about the probes positions is saved in $\_probeInfo.mat$ file. Using the software nirsLAB (NIRx Medical Technologies, LLC.), one could see the visualization of probes and channels' position regarding to the EEG positions, see Figure 5.

The topology file $.tpl$ shows the position of channels in the form of a table with cell value like `204.000` which means channel made up by source s2 and detector d4, s2-d4. This information is also used to visualize channels' position, see Figure 6.

More information about these files could be found in the manuals of NIRStar and nirsLAB.

Appendix B. Calculating concentration changes

As aforementioned, the amount of hemoglobin in the tissue affects the attenuation of light which passes through the tissue. Using the modified Beer-Lambert law, we could estimate the attenuation of the transmitted light A as

$$A = -log_{10}\frac{I}{I_0} = \epsilon * c * d * DPF + G \quad (1)$$

where I_0 is the light intensity emitted at the source optode which is constant during the time course, I is the light intensity received at the detector optode, ϵ is the extinction coefficient of hemoglobin at a specific wavelength, c is the hemoglobin concentration, d is the distance between source and detector optodes, differential pathlength factor DPF is the factor which when multiplying with d will define the pathlength of the light travelling from source to detector optode, and G accounts for part of scattered light which is not received at detector optode (losses) (Delpy et al., 1988; Villringe & Chance, 1997; Boas et al, 2001).

If ϵ, d, DPF and G are known, we could calculate the *absolute* hemoglobin concentration, however for continuous wave NIRS devices (CW-NIRS), which is described in the next section, that is not possible.

For CW-NIRS, considering at a specific location, a change in hemoglobin concentration causes change in the light attenuation. In that case, the extinction coefficient ϵ and distance d are still the same, assuming DPF and G also unchanged, the change in the attenuation is

$$\Delta A = A_{after} - A_{before} = log_{10}\frac{I_{before}}{I_{after}} = \epsilon * \Delta c * d * DPF \quad (2)$$

where ΔA denotes the change in attenuation, I_{before} and I_{after} are the received intensities before and after the concentration changed, respectively, and Δc is the change in concentration of hemoglobin.

There are several chromophores in the tissue however the most significant contributions to the absorption are from deoxyhemoglobin (Hb) and oxyhemoglobin (HbO). To identify these amounts, equation (2) is rewritten as

$$\Delta A^{\lambda} = log_{10}\frac{I^{\lambda}_{before}}{I^{\lambda}_{after}} = (\epsilon^{\lambda}_{Hb} * \Delta\mathrm{Hb} + \epsilon^{\lambda}_{HbO} * \Delta\mathrm{HbO}) * d * DPF^{\lambda} \quad (3)$$

where λ denotes a specific wavelength, $\epsilon^{\lambda}_{Hb}$and ϵ^{λ}_{HbO} are the corresponding extinction coefficient of Hb and HbO at the specific wavelength, $\Delta\mathrm{Hb}$ and $\Delta\mathrm{Hb}O$ are the Hb and HbO changes, respectively, and DPF^{λ} is the differential path length factor at the specific wavelength.

Based on the received intensities of two wavelengths and the known extinction coefficient of Hb and HbO and DPF at each wavelength, the Hb and HbO changes could be determined by the following equations.

$$\Delta\mathrm{Hb} = \frac{\epsilon^{\lambda_2}_{HbO} * \frac{\Delta A^{\lambda_1}}{DPF^{\lambda_1}} - \epsilon^{\lambda_1}_{HbO} * \frac{\Delta A^{\lambda_2}}{DPF^{\lambda_2}}}{d * (\epsilon^{\lambda_1}_{Hb} * \epsilon^{\lambda_2}_{HbO} - \epsilon^{\lambda_1}_{HbO} * \epsilon^{\lambda_2}_{Hb})} \quad (4)$$

$$\Delta \mathrm{HbO} = \frac{\epsilon_{Hb}^{\lambda_1} * \frac{\Delta A^{\lambda_2}}{DPF^{\lambda_2}} - \epsilon_{Hb}^{\lambda_2} * \frac{\Delta A^{\lambda_1}}{DPF^{\lambda_1}}}{d * (\epsilon_{Hb}^{\lambda_1} * \epsilon_{HbO}^{\lambda_2} - \epsilon_{HbO}^{\lambda_1} * \epsilon_{Hb}^{\lambda_2})} \quad (5)$$

If natural logarithm (ln) is used in equation (1), the term extinction coefficient (ϵ) is replaced by absorption coefficient (α) where $\alpha = \ln(10) * \epsilon$. The unit of extinction coefficient is $cm^{-1}mole^{-1}l$ hence the unit for absorption coefficient. In practice, absorption coefficient is commonly used and in $cm^{-1}millimole^{-1}l$ unit (Cope, 1991, p. 58). DPF and ΔA are unitless. The unit of distance d is cm and thus the unit of hemoglobin changes is $millimole/l$ (mM/l). In order to have easy reading numbers of hemoglobin changes, the molar abortion coefficient could be converted to $cm^{-1}micromole^{-1}l$ unit by dividing over 1000, in this case the unit of concentration changes is $micromole/l$ ($\mu M/l$).

In this thesis, concentration changes for each channel, defined by a pair of source and detector optodes, were calculated using equation (4) and (5) where

$$\Delta A^{\lambda} = ln\frac{mean(\sum I^{\lambda})}{I^{\lambda}} \quad (6)$$

with I being the raw intensity received at the corresponding detector optode at a time point at a specific wavelength. The mean of raw intensities over the time course was used as the reference point to keep the calculated concentration changes along the time course around the x axis, which made the visualization of concentration changes easier to follow.

There are many sources which provide different values for DPF and extinction coefficients which lead to different quantification of concentration changes however the trend of them are relatively the same as illustrated in Appendix C. In this thesis, the molar absorption coefficients were calculated based on the molar extinction coefficients from W. B. Gratzer [1] and DPF values advised by NIRx Medical Technologies, the manufactory of the CW-NIRS device which was used to perform the measurements at Philips Research (Table 28).

Table 28. Parameters to calculate concentration changes, advised by NIRx.

	DPF^{λ}	$\alpha_{HbO}^{\lambda}\ (cm^{-1}\mu M^{-1}l)$	$\alpha_{Hb}^{\lambda}(cm^{-1}\mu M^{-1}l)$
$\lambda_1 = 760\ nm$	6.40	$1.4865865 * 10^{-3}$	$3.843707 * 10^{-3}$
$\lambda_2 = 850\ nm$	5.75	$2.526391 * 10^{-3}$	$1.798643 * 10^{-3}$

The unit of the molar absorption coefficient is set as $cm^{-1}\mu M^{-1}l$ such that the unit of the calculated concentration changes is $\mu M/l$ to make the values of concentration changes easier to read.

[1] http://omlc.org/spectra/hemoglobin/summary.html

Appendix C. Accuracy of NIRS

There are many sources which provide different values for DPF and extinction coefficients which lead to different quantification of concentration changes. To verify the similarity of the trend of concentration changes when using different values of DPFs and absorption coefficients, I used combination values from two sets as shown in Table 29 and Table 30. The DPF values in set 1 and set 2 were from the low and high curve, respectively, in Essenpreis et al. (1993, Fig 4.). The molar absorption coefficients in set 1 and 2 were calculated based on the molar extinct coefficient from W. B. Gratzer [2] and J. M. Schmitt[3], respectively. The results in Figure 49, Figure 50 and Figure 51 confirmed that the trend of concentration changes stayed the same.

Table 29. The DPF values were from Essenpreis et al. (1993). The molar absorption coefficients in set 1 and 2 were calculated based on the molar extinct coefficient from W. B. Gratzer and J. M. Schmitt, respectively.

	Set 1	*Set 2*
DPF^{λ_1}	6.40	7.25
DPF^{λ_2}	5.75	6.38
$\epsilon_{Hb}^{\lambda_1}$ $\epsilon_{HbO}^{\lambda_1}$	$3.843707 * 10^{-3}$ $1.4865865 * 10^{-3}$	$3.96116 * 10^{-3}$ $1.3818 * 10^{-3}$
$\epsilon_{Hb}^{\lambda_2}$ $\epsilon_{HbO}^{\lambda_2}$	$1.798643 * 10^{-3}$ $2.526391 * 10^{-3}$	$1.8424 * 10^{-3}$ $2.44118 * 10^{-3}$

Table 30. Combination of setting values for the comparisons.

	DPFs	Molar absorption coefficients
Option 1	Set 1	Set 1
Option 2	Set 2	Set 1
Option 3	Set 1	Set 2
Option 4	Set 1	Set 1
Option 5	Set 1	Set 2
Option 6	Set 2	Set 1

[2] http://omlc.org/spectra/hemoglobin/summary.html
[3] http://omlc.org/spectra/hemoglobin/moaveni.html

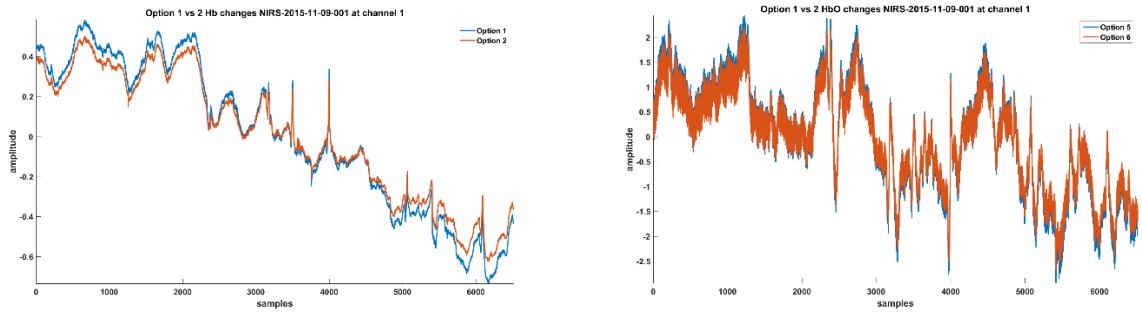


Figure 49. Comparison of concentration changes calculated based on option 1 (blue) and 2 (orange), different DPFs, same absorption coefficients. Left: Hb changes. Right: HbO changes.

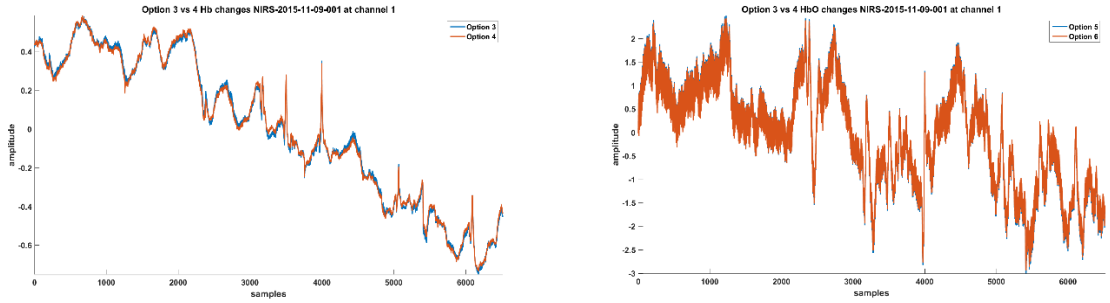


Figure 50. Comparison of concentration changes calculated based on option 3 (blue) and 4 (orange), same DPFs, different absorption coefficients. Left: Hb changes. Right: HbO changes.

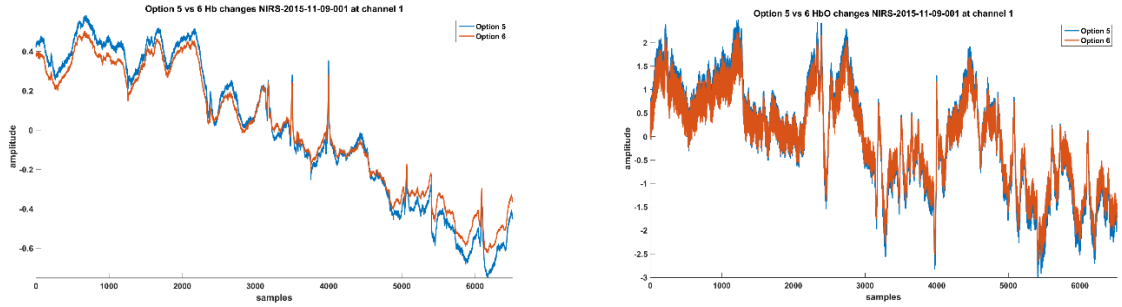
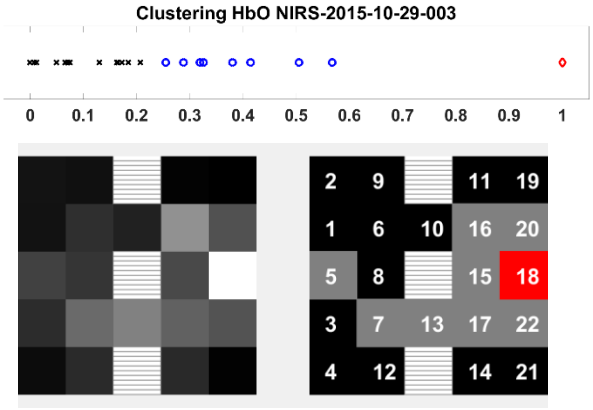
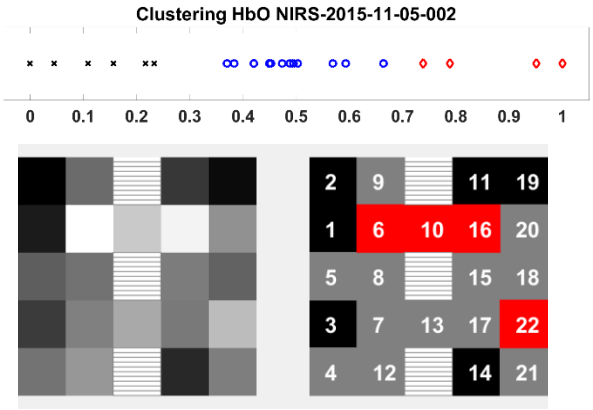


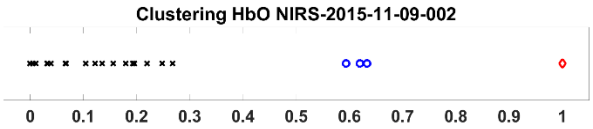
Figure 51. Comparison of concentration changes calculated based on option 5 (blue) and 6 (orange), different DPFs, different absorption coefficients. Left: Hb changes. Right: HbO changes.

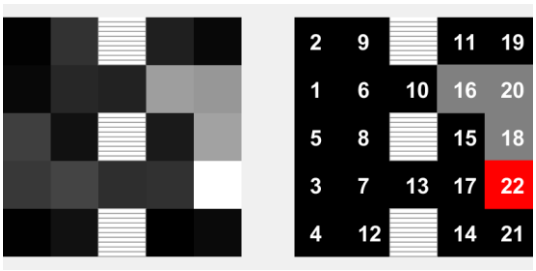
Appendix D. Spatial analysis results at single subject level for Visual stimulation experiment

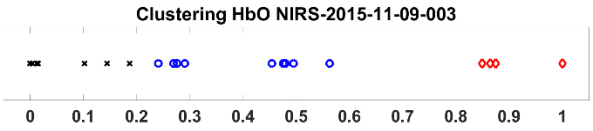
The results of 11 subject are shown as follows. Per subject there are one graphs and two images. The first graph shows three groups of grayscale values based on the two thresholds as results of the two-threshold Otsu method. Red color indicates values which belong to the strongest group. Blue and black indicate the other groups by decreasing order of the values. The left image illustrates the topology of the channels in grayscale image and the right image illustrates the topology of the channels in 3 groups where red cell(s) represents the strongest location(s), gray and black cells represent the other locations by order of decreasing strength of signal; the number in each cell is the channel index. Three cells with horizontal stripes in the middle of each image indicate non-measured areas.

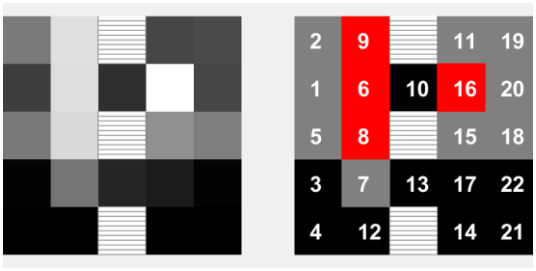


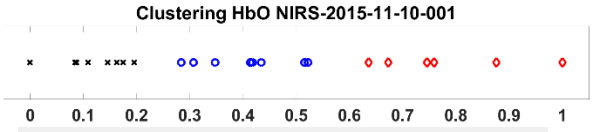


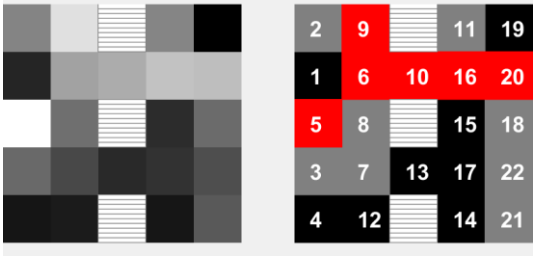


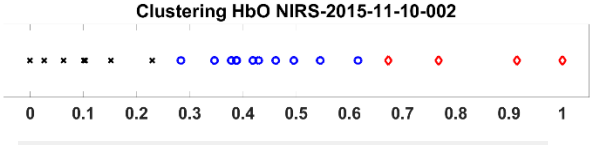

2 9 11 19
1 6 10 16 20
5 8 15 18
3 7 13 17 22
4 12 14 21

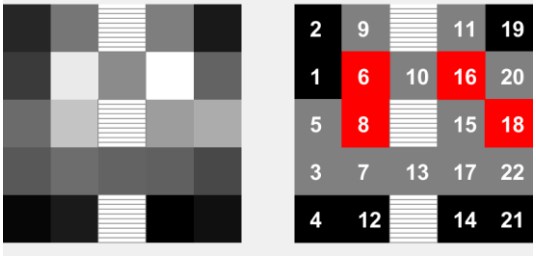

Clustering HbO NIRS-2015-11-09-003
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

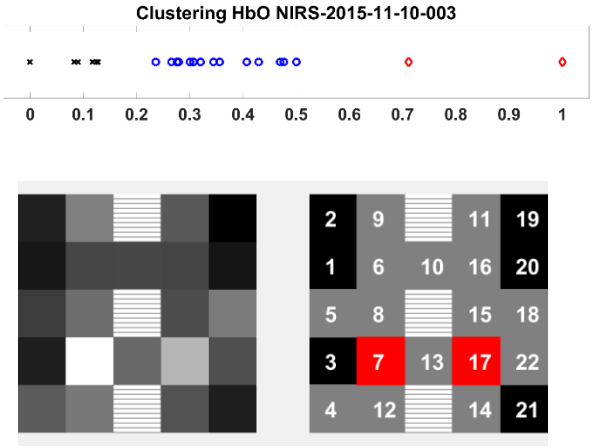

2 9 11 19
1 6 10 16 20
5 8 15 18
3 7 13 17 22
4 12 14 21

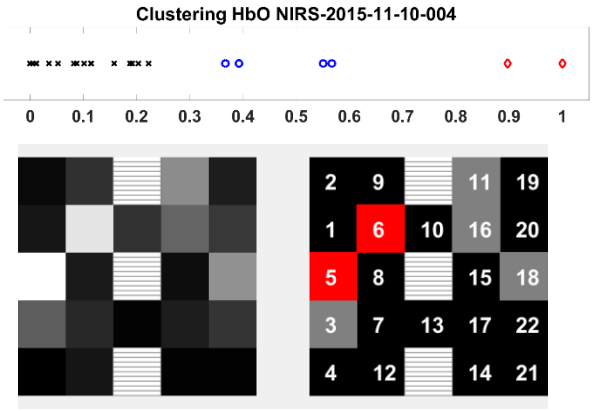

Clustering HbO NIRS-2015-11-10-001
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

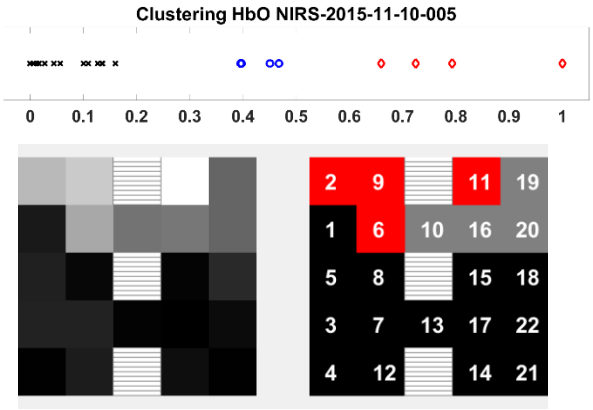

2 9 11 19
1 6 10 16 20
5 8 15 18
3 7 13 17 22
4 12 14 21

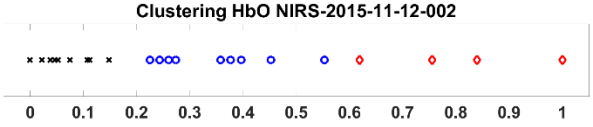

Clustering HbO NIRS-2015-11-10-002
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1


2 9 11 19
1 6 10 16 20
5 8 15 18
3 7 13 17 22
4 12 14 21


Clustering HbO NIRS-2015-11-10-003
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
2 9 11 19
1 6 10 16 20
5 8 15 18
3 7 13 17 22
4 12 14 21


Clustering HbO NIRS-2015-11-10-004
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
2 9 11 19
1 6 10 16 20
5 8 15 18
3 7 13 17 22
4 12 14 21


Clustering HbO NIRS-2015-11-10-005
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
2 9 11 19
1 6 10 16 20
5 8 15 18
3 7 13 17 22
4 12 14 21


Clustering HbO NIRS-2015-11-12-002
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

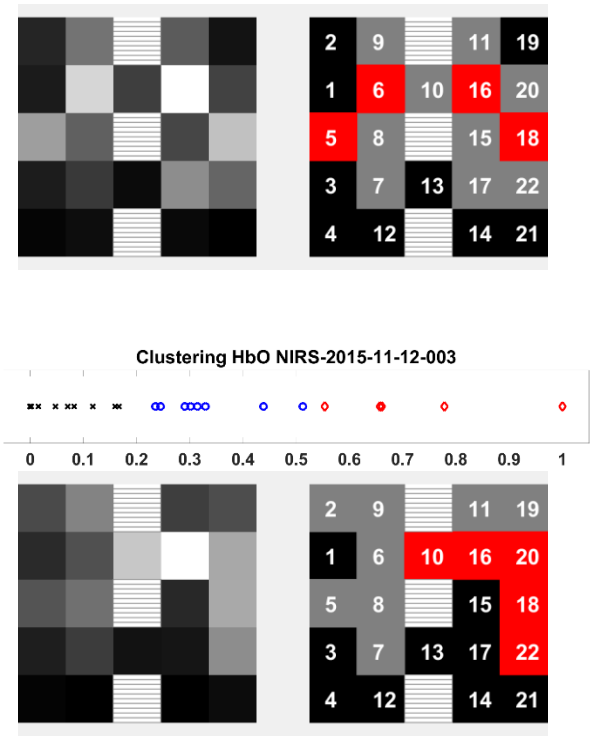
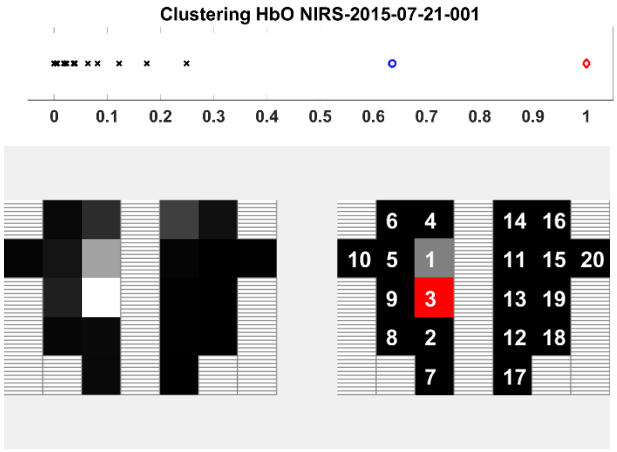


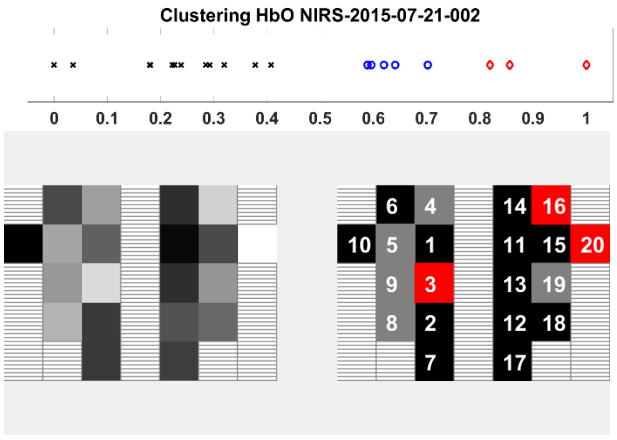
Figure 52. Visual stimulation experiment. Results of 11 subjects. Per subject: top graph shows three groups of grayscale values as the results of the two-threshold Otsu method, red values indicate group with highest values; bottom-left image shows the topology of the channels in grayscale; bottom-right image shows the topology of the channels grouped by strength of signal, red cells illustrate locations of strongest signal, gray and black cells illustrate locations of weaker signal by decreasing order, each number represents the corresponding channel index. Cells with stripes indicate non-measured areas.

Appendix E. Spatial analysis results at single subject level for Finger tapping experiment

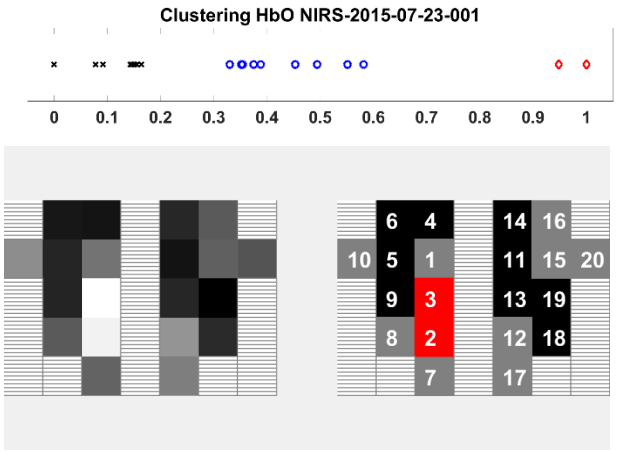
The results of 5 subject are shown as follows. Per subject there are one graphs and two images. The first graph shows three groups of grayscale values based on the two thresholds as results of the two-threshold Otsu method. Red color indicates values which belong to the strongest group. Blue and black indicate the other groups by decreasing order of the values. The left image illustrates the topology of the channels in grayscale image and the right image illustrates the topology of the channels in 3 groups where red cell(s) represents the strongest location(s), gray and black cells represent the other locations by order of decreasing strength of signal; the number in each cell is the channel index. Cells with horizontal stripes each image indicate non-measured areas. The result of paired t-test was also shown for each subject for easily referencing.



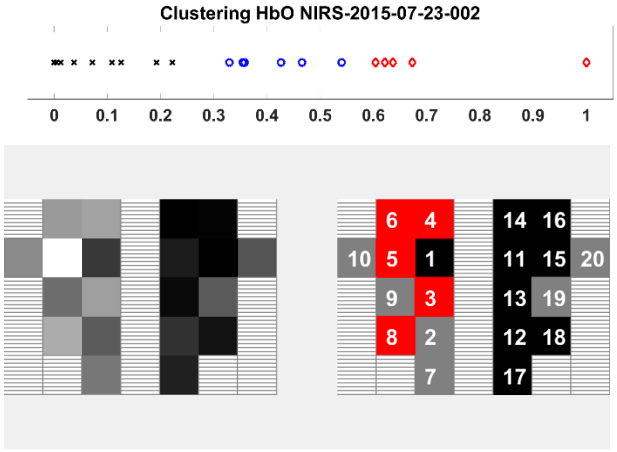
Paired *t*-test result: the mean of HbO changes of 7 channels in the left hemisphere was significantly higher than that in the right hemisphere (*t*=3.2221, *p*=0.0181).



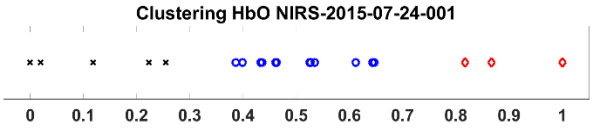
Paired *t*-test result: the mean of HbO changes of 7 channels in the left hemisphere was not significantly different from that in the right hemisphere (*t*=0.4425, *p*=0.6736).

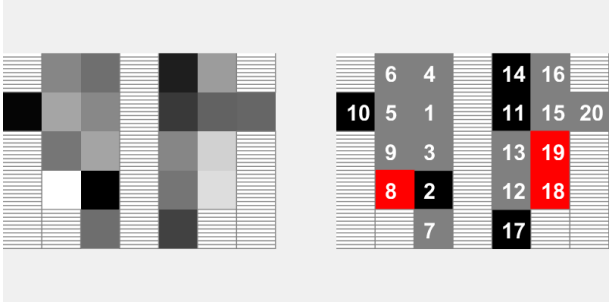


Paired *t*-test result: the mean of HbO changes of 7 channels in the left hemisphere was considered as significantly higher than that in the right hemisphere (*t*= 2.4247, *p*= 0.0515).



Paired *t*-test result: the mean of HbO changes of 7 channels in the left hemisphere was significantly higher than that in the right hemisphere (*t*= 2.9377, *p*= 0.0260).



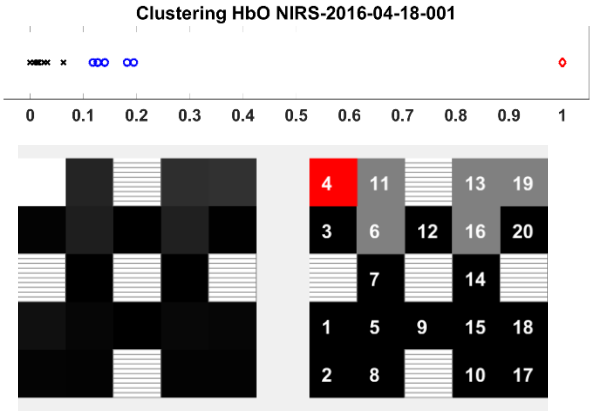


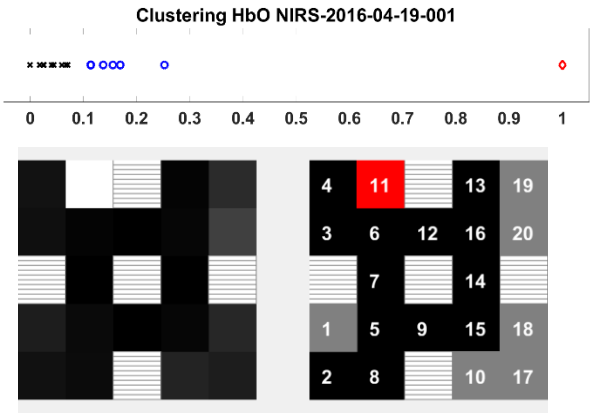
Paired *t*-test result: the mean of HbO changes of 7 channels in the left hemisphere was not significantly different from that in the right hemisphere (*t*= -0.4326, *p*= 0.6804).

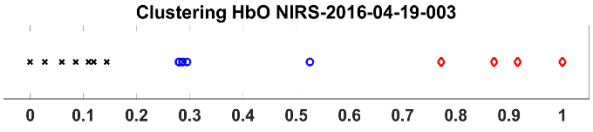
Figure 53. Finger tapping experiment. Results of 5 subjects. Per subject: top graph shows three groups of grayscale values as the results of the two-threshold Otsu method, red values indicate group with highest values; bottom-left image shows the topology of the channels in grayscale; bottom-right image shows the topology of the channels grouped by strength of signal, red cells illustrate locations of strongest signal, gray and black cells illustrate locations of weaker signal by decreasing order, each number represents the corresponding channel index. Cells with stripes indicate non-measured areas.

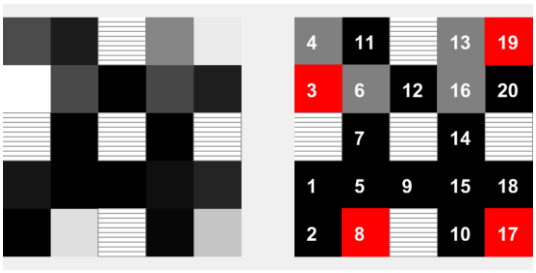
Appendix F. Spatial analysis results at single subject level for CPT experiment

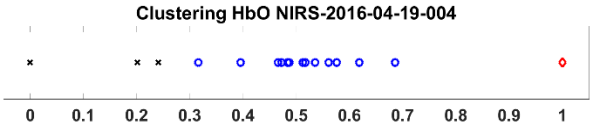
The results of 10 subject are shown as follows. Per subject there are one graphs and two images. The first graph shows three groups of grayscale values based on the two thresholds as results of the two-threshold Otsu method. Red color indicates values which belong to the strongest group. Blue and black indicate the other groups by decreasing order of the values. The left image illustrates the topology of the channels in grayscale image and the right image illustrates the topology of the channels in 3 groups where red cell(s) represents the strongest location(s), gray and black cells represent the other locations by order of decreasing strength of signal; the number in each cell is the channel index. Three cells with horizontal stripes in the middle of each image indicate non-measured areas.

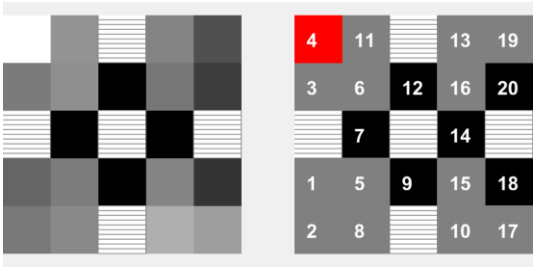


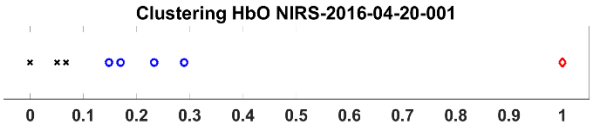


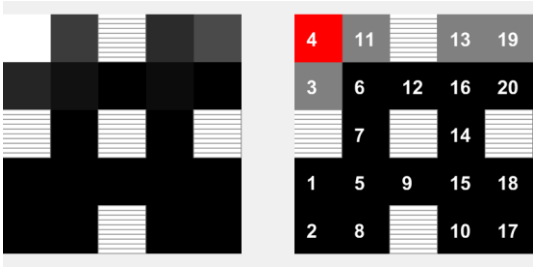


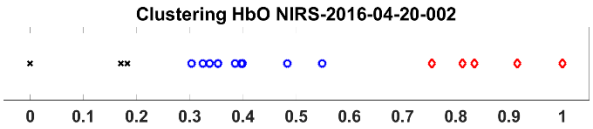

4 11 13 19
3 6 12 16 20
7 14
1 5 9 15 18
2 8 10 17

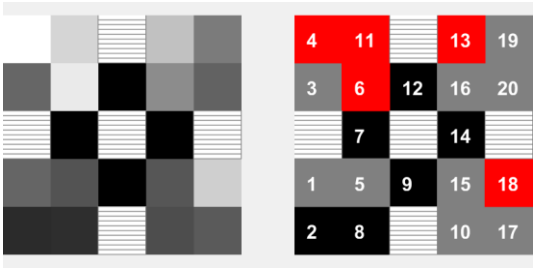

Clustering HbO NIRS-2016-04-19-004
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

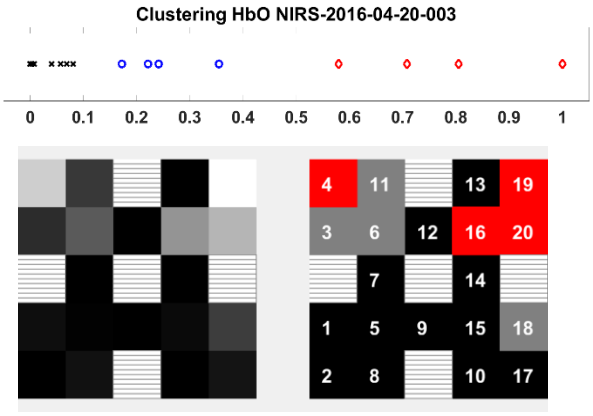

4 11 13 19
3 6 12 16 20
7 14
1 5 9 15 18
2 8 10 17

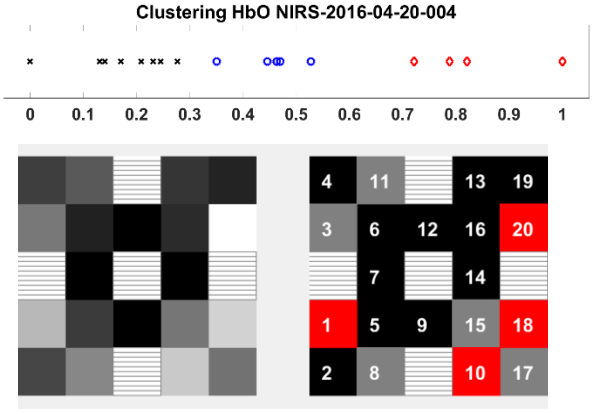

Clustering HbO NIRS-2016-04-20-001
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

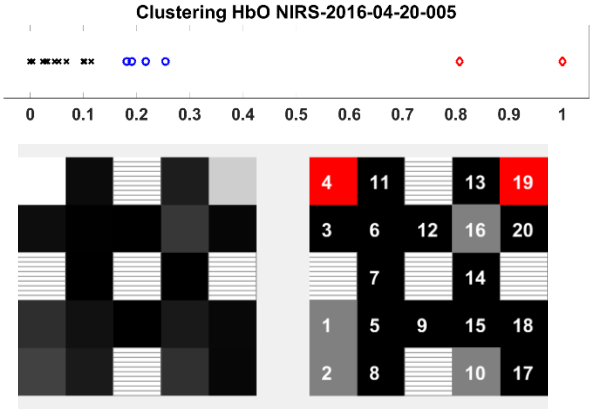

4 11 13 19
3 6 12 16 20
7 14
1 5 9 15 18
2 8 10 17

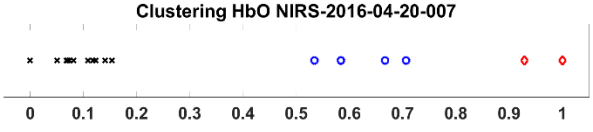

Clustering HbO NIRS-2016-04-20-002
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1


4 11 13 19
3 6 12 16 20
7 14
1 5 9 15 18
2 8 10 17


Clustering HbO NIRS-2016-04-20-003
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
4 11 13 19
3 6 12 16 20
7 14
1 5 9 15 18
2 8 10 17


Clustering HbO NIRS-2016-04-20-004
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
4 11 13 19
3 6 12 16 20
7 14
1 5 9 15 18
2 8 10 17


Clustering HbO NIRS-2016-04-20-005
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
4 11 13 19
3 6 12 16 20
7 14
1 5 9 15 18
2 8 10 17


Clustering HbO NIRS-2016-04-20-007
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

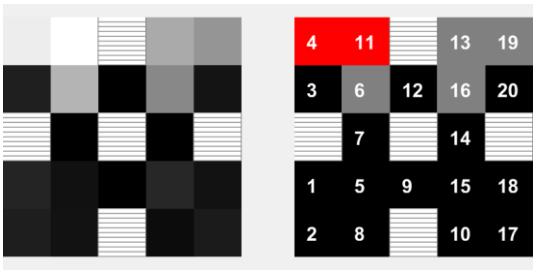
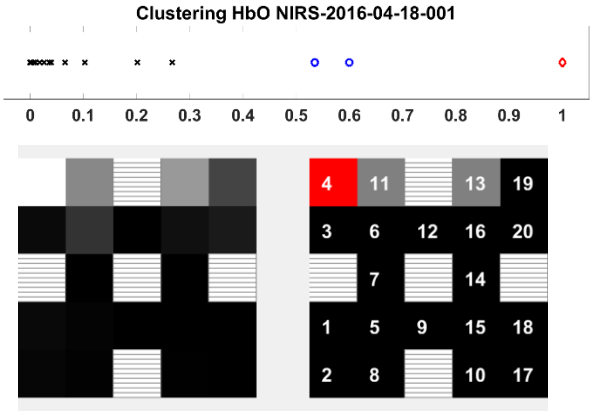
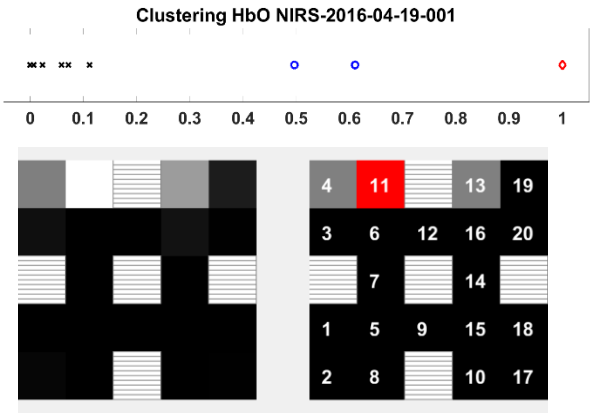


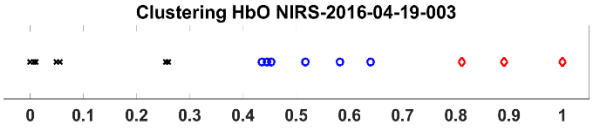
Figure 54. CPT experiment. Results of 10 subjects. Per subject: top graph shows three groups of grayscale values as the results of the two-threshold Otsu method, red values indicate group with highest values; bottom-left image shows the topology of the channels in grayscale; bottom-right image shows the topology of the channels grouped by strength of signal, red cells illustrate locations of strongest signal, gray and black cells illustrate locations of weaker signal by decreasing order, each number represents the corresponding channel index. Cells with stripes indicate non-measured areas.

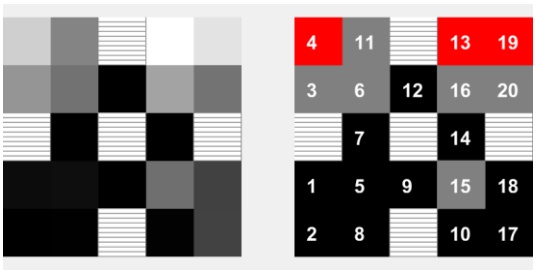
Appendix G. Spatial analysis results at single subject level for VFT experiment

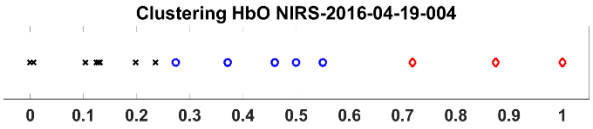
The results of 10 subject are shown as follows. Per subject there are one graphs and two images. The first graph shows three groups of grayscale values based on the two thresholds as results of the two-threshold Otsu method. Red color indicates values which belong to the strongest group. Blue and black indicate the other groups by decreasing order of the values. The left image illustrates the topology of the channels in grayscale image and the right image illustrates the topology of the channels in 3 groups where red cell(s) represents the strongest location(s), gray and black cells represent the other locations by order of decreasing strength of signal; the number in each cell is the channel index. Three cells with horizontal stripes in the middle of each image indicate non-measured areas.

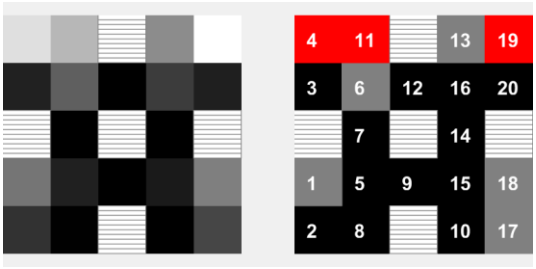


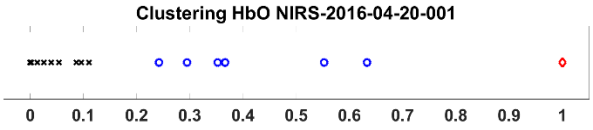


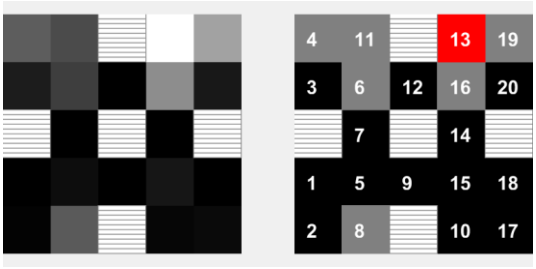


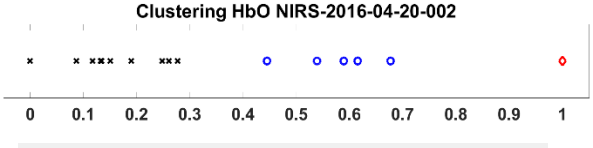

4 11 13 19
3 6 12 16 20
7 14
1 5 9 15 18
2 8 10 17

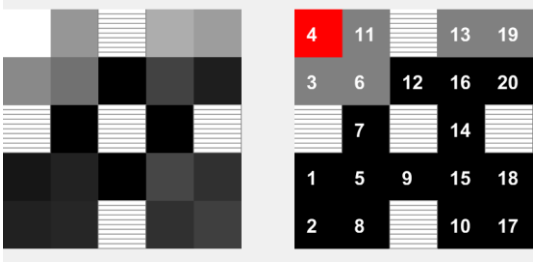

Clustering HbO NIRS-2016-04-19-004
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

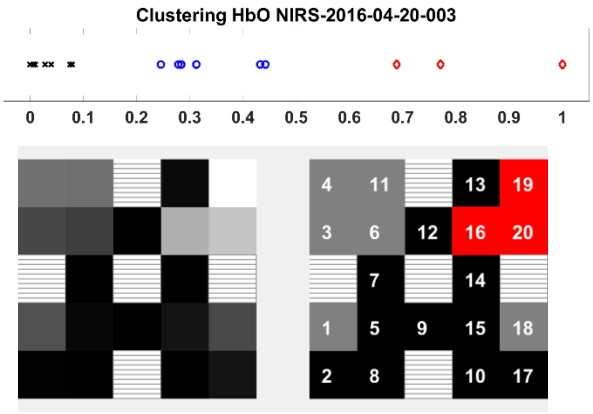

4 11 13 19
3 6 12 16 20
7 14
1 5 9 15 18
2 8 10 17

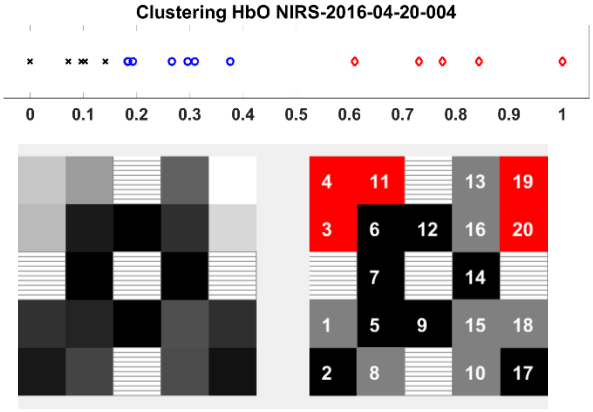

Clustering HbO NIRS-2016-04-20-001
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

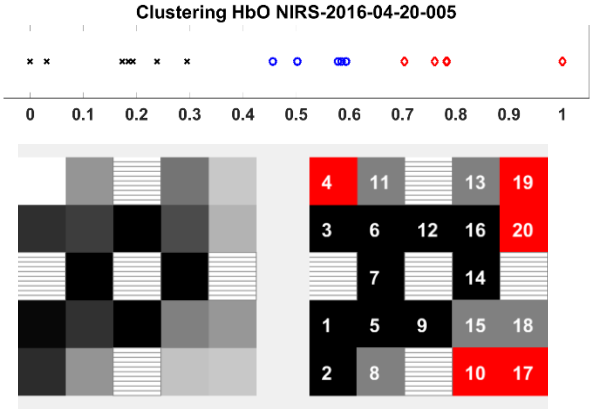

4 11 13 19
3 6 12 16 20
7 14
1 5 9 15 18
2 8 10 17

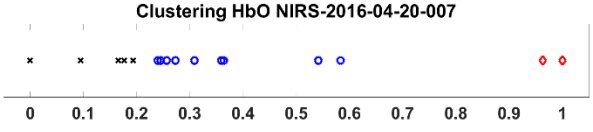

Clustering HbO NIRS-2016-04-20-002
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1


4 11 13 19
3 6 12 16 20
7 14
1 5 9 15 18
2 8 10 17


Clustering HbO NIRS-2016-04-20-003
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
4 11 13 19
3 6 12 16 20
7 14
1 5 9 15 18
2 8 10 17


Clustering HbO NIRS-2016-04-20-004
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
4 11 13 19
3 6 12 16 20
7 14
1 5 9 15 18
2 8 10 17


Clustering HbO NIRS-2016-04-20-005
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
4 11 13 19
3 6 12 16 20
7 14
1 5 9 15 18
2 8 10 17


Clustering HbO NIRS-2016-04-20-007
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

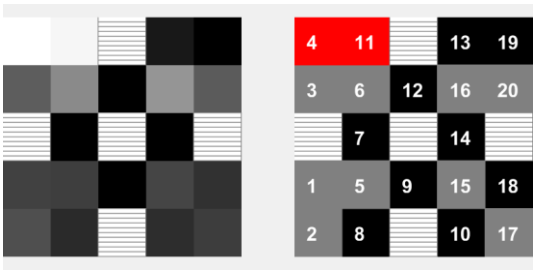


Figure 55. VFT experiment. Results of 10 subjects. Per subject: top graph shows three groups of grayscale values as the results of the two-threshold Otsu method, red values indicate group with highest values; bottom-left image shows the topology of the channels in grayscale; bottom-right image shows the topology of the channels grouped by strength of signal, red cells illustrate locations of strongest signal, gray and black cells illustrate locations of weaker signal by decreasing order, each number represents the corresponding channel index. Cells with stripes indicate non-measured areas.