

# Using fNIRS for Prefrontal-Asymmetry Neurofeedback: Methods and Challenges

Gabor Aranyi, Marc Cavazza<sup>(✉)</sup>, and Fred Charles

School of Computing, Teesside University, Borough Road,  
Middlesbrough TS1 3BA, UK

{g.aranyi,m.o.cavazza,f.charles}@tees.ac.uk

**Abstract.** Functional near-infrared spectroscopy (fNIRS) has become increasingly accessible in recent years, which allows this relatively low-cost and portable brain sensing modality for the application of brain-computer interfaces (BCI). Although there is a growing body of research on fNIRS-based BCI utilising users' covert psychophysiological activity, there is comparably less research on active BCI, where users engage in thinking strategies with the explicit intention of controlling the behaviour of an interactive system. We draw on four empirical studies, where participants received real-time neurofeedback (NF) of left-asymmetric increase in activation in their dorsolateral prefrontal cortex (DL-PFC), which has previously been identified as a correlate of approach-related motivational tendencies. We discuss methodological considerations and challenges, and provide recommendations about brain-signal selection and integration, NF protocol design, post-hoc and real-time applications of NF success criteria, continuous visual feedback, and individualised feedback based on the variations of the brain-signal in a reference condition.

**Keywords:** Prefrontal asymmetry · Functional near-infrared spectroscopy · (Affective) brain-computer interfaces · Neurofeedback

## 1 Introduction

The development of affective brain-computer interfaces (BCI) is a relatively recent trend [11], which has the potential to support new interactive systems, human-robot interaction [9], and cultural [14] or entertainment applications. One of the main challenges faced by affective BCI, especially if compared with BCI based on motor areas, is to identify a clear mapping between a target affective state and a BCI signal. This could be obtained from knowledge about the neural localisation of affective states: however, such knowledge is particularly elusive, as suggested by a recent review [15].

However, research in psychophysiology has identified a possible neural correlate for a specific affective dimension, known as approach/withdrawal [4], in the form of asymmetric activity in the prefrontal cortex (PFC). Approach/withdrawal behaves as a high-level affective dimension, and has been shown to play an important role in motivational processes, reward expectation, risk taking and

depression. It has originally been explored through EEG studies, which have defined asymmetry scores that can characterise this asymmetry. These behave as individual traits but are also subject to dynamic variations: furthermore, they can be controlled through neurofeedback (NF), as originally demonstrated by Rosenfeld and colleagues [19]. The potential use of prefrontal asymmetry to support affective BCI has been discussed in various reviews [17], although without reporting specific implementations.

Early work on using prefrontal asymmetry for BCI was based mostly on EEG signals. Wehbe et al. [27] reported passive measurement of EEG prefrontal asymmetry during computer gameplay; however, they claimed to be using it as a measure of arousal rather than approach. Karran et al. [14] explored the role of the PFC in subjects' aesthetic experiences. Our previous work explored EEG-based NF in the alpha band, with simultaneous fMRI analysis over NF epochs [12]. It confirmed that the affective strategies through which users controlled PFC alpha asymmetry corresponded to asymmetric activity in prefrontal regions (across areas BA9 and BA10), with no differences observed in pre-motor areas. These experiments are difficult to interpret any further, due to the small number of subjects, and the finding that the supine position is known to impair the ability of subjects to properly express approach [4]. In further experiments carried out in laboratory conditions, subjects achieved success rates of up to 73% with minimal training [7]. However, signal quality and stability during NF epochs remains an issue, and a limiting factor.

We posit that functional near-infrared spectroscopy (fNIRS) can provide an alternative, offering better signal quality and better resistance to motion artefacts, while also improving spatial resolution for the target brain areas. This is also supported by the finding that areas relevant to approach/withdrawal include the dorsolateral prefrontal cortex (DL-PFC) [25], whose localisation is accessible to fNIRS. Sitaram et al. [23] were amongst the first to suggest that signals based on metabolic activity could be equally suited to BCI than electrical signals. We were also inspired by recent experiments by Zotev et al. [28], which reported PFC-NF with both EEG and real-time fMRI. Finally, Naseer and Hong [18] have also reviewed recent uses of fNIRS for NF.

In this paper, we discuss methodological aspects of deploying fNIRS to implement BCI based on prefrontal asymmetry, under a NF paradigm. These are based on several experiments, one published [5] and the others accepted for publication or under review, during which we explored various settings for controlling approach, in affective contexts as diverse as empathy, anger or motivation. Rather than reproducing these studies here, we shall concentrate on specific elements of methodology, some common to all studies, such as optode selection and signal definition, and some more specific, such as the definition of baseline, control tasks compared to NF epochs, and calculations of statistical significance, both online and post hoc.

## 2 Brain-Signal Acquisition, Selection and Integration

We used fNIRS to operationalise BCI input based on asymmetric functional activation in the DL-PFC. Although the spatial resolution of fNIRS falls short that of fMRI and is limited to scanning the outer cortex, it has a number of advantages, such as lower susceptibility to motion artefacts and lower cost, that make it appropriate for application in BCI [3]. We followed the recommendation of Solovey et al. [24] for the use of fNIRS in HCI settings. We used an fNIR400 Optical Brain Imaging Station by Biopac Systems, with a 16-channel sensor with fixed 2.5cm source-detector separation (see [20] for channel locations). Data were collected with 2Hz sampling rate. This fNIRS device measures intensity changes in two wavelengths (730nm and 850nm) over time to calculate the change in oxygenated (HbO) and deoxygenated (HbR) haemoglobin concentration (in units of  $\mu\text{Mol/L}$ ) using the modified Beer-Lambert Law [1]. In order to provide real-time feedback based on brain activity measured by fNIRS, there is a need to select a single metric: HbO, HbR, or HbT (total haemoglobin; the sum of HbO and HbR). We conducted a pilot study to inform this decision.

The pilot study used a no-feedback paradigm including an approach task (watching pictures of delicious food under the instructions to imagine reaching out for the food and eating it) and a withdrawal task (watching pictures of spiders under the instructions to imagine escaping from the situation)<sup>1</sup>. We compared the metrics of HbO, HbR, and HbT to assess how well they are able to discriminate between the tasks. Based on literature co-authored by the developer of the fNIRS system used in our experiments (e.g., Ruocco et al. [20]), on literature applying HbO to affect-related manipulation in the DL-PFC [26], and to approach/withdrawal-related experimental manipulation [16], and based on our pilot study, we elected to use HbO for real-time application; we based post-hoc analyses on the same metric for consistency.

The haemodynamic response measured by fNIRS takes several seconds [3]. We took two approaches to accommodate for this approximately 7s delay: we either (a) simply removed the first 14 data points (corresponding to 7s sampled at 2Hz) of each epoch on each channel, or we (b) also included the 14 data points after the completion of the epoch (i.e., windowing; see Sarkheil et al. [22] for a similar approach).

The complexity of measured changes in blood oxygenation associated with the differential functional activation of the DL-PFC needs to be reduced to afford effective BCI input. This consists in deriving a single asymmetry metric from the continuous flow of oxygenation data from the input channels [8]. We averaged HbO values over the four leftmost and four rightmost channels (located over the left and right DL-PFC, respectively), then subtracted average Right from average Left. This metric reflects the inter-hemispheric difference in HbO change

---

<sup>1</sup> Pilot subjects ( $N = 4$ ) confirmed positive attitude towards the food items and negative towards spiders. We additionally included an approach condition involving the same spider pictures with instructions to imagine approaching the spiders in protective clothing and swatting them.

in micromolar units ( $\mu\text{Mol/L}$ ). Note that this measure is relative to a baseline [2], and more importantly, it lacks an absolute zero point, as opposed to, for example, alpha-power asymmetry in EEG-based NF [6]. This has important practical consequences in defining and quantifying NF success. For example, as this operationalisation of asymmetry yields interval-level data, a ratio of task/no-task signals for defining and quantifying success (e.g., [22]) cannot be applied.

### 3 Protocol Design Considerations

There is a growing body of research on fNIRS-based NF [18]; however, fMRI-based NF research can also effectively inform fNIRS study design due to the comparability of the haemodynamic signal measured by the two neuroimaging modalities (see [3] for a comparison). We sought inspiration from fMRI-based studies [22, 28] to inform study design, because of their relevance to the target mental activity (affective regulation), experimental task (up-regulation of activity in a target area using thinking strategies), and feedback operationalisation.

Protocols for experiments and interaction design for active BCI need to be tailored for supporting a feedback strategy and applying a success criterion, depending on the tasks the participants are required to carry out for interacting with the system. The length of individual epochs (short time periods with a specific task) and blocks (a sequence of epochs), and that of the entire protocol (number of blocks), needs to be manageable for participants, while it also needs to provide a sufficient quantity of data for the purpose of research and application. These considerations place constraints on how much data is to be collected, and how data collection can be structured in a way that it is most manageable for participants.

With regards to the length of each block and the overall number of blocks, the two main considerations are (a) the participants' ability to maintain focus on the mental activity, and (b) the amount of data necessary to support the use of the success criterion. To address (a), we asked participants to provide subjective difficulty ratings of each task involved in the experiments, and we also conducted post-use interviews to gather qualitative data about their interaction with the system. Regarding (b), when using a statistical success criterion, discussed in detail in the following section, we advise conducting a power analysis to determine the number of observations required within each epoch to detect an effect with a given magnitude (some data may need to be discarded in the filtering process).

These considerations are inherently related to how feedback is provided and NF success is defined. In the following sections, we illustrate the approach we took to addressing these challenges through two sets of experiments.

### 4 Criterion for Neurofeedback Success

We applied a statistical criterion to determine NF success. Specifically, we characterised NF success as a statistically significant increase in left-asymmetry during a NF epoch, compared to either (a) a baseline or (b) a reference epoch.

Since statistical significance depends on the sample size, or in the present case, the number of observations in a time-series we refer to as an epoch, we also calculated different effect-size measures ( $r$  and Cohen's  $d$ ) to characterise the magnitude of NF success. We also implemented real-time, automated determination of NF success as well as post-hoc testing. Additionally, we explored if applying a non-parametric approach (bootstrapping) delivers practical benefits. We discuss the advantages and disadvantages of these approaches in the following sections through two sets of experiments. We describe the experimental protocols designed to support the application of the statistical success criterion, with variations along the following properties:

- Success evaluated against zero asymmetry or asymmetry during a reference epoch;
- Success evaluated real-time or post-hoc;
- Treatment of delay (trimming or windowing);
- Threshold characterisation and feedback mapping (fixed or personalised).

#### 4.1 Evaluating NF Success Against Baseline

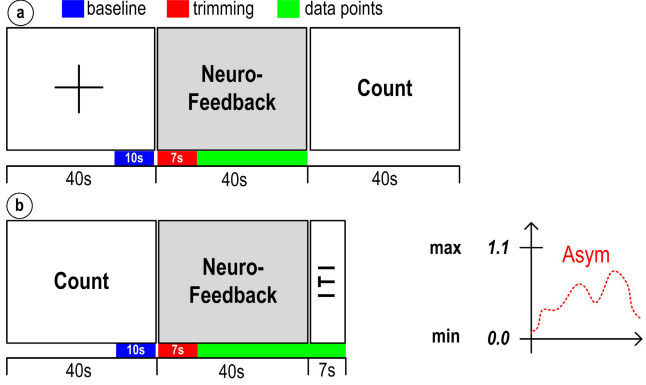
With the fNIRS system we applied, baseline is measured over 10s, against which the asymmetry scores collected during a NF epoch can be compared, without the need for a reference epoch. Since this baseline measurement consists in collecting light-intensity data on each channel as a reference for calculating oxygenation changes [1], the asymmetry metric we derive from the channels is zero for the baseline. Therefore, we determined NF success using the baseline criterion by performing a one-sample t-test on the asymmetry scores collected during the NF epoch against the test value zero. Performing this one-sample t-test [10] is computationally simple and can be implemented real-time by calculating the  $t$  value upon the completion of the NF epoch by dividing the mean of observed asymmetry values by the estimate of the standard error (Fig. 1a), which is then compared to a critical value to determine NF success.

We conducted two experiments using this success criterion (Fig. 2). In Experiment 1a, success was determined real-time, based on unfiltered data, using a parametric criterion, delay was treated by trimming, while threshold and mapping for the feedback channel were fixed (i.e. the same for each subject and each experimental trial). By comparison, in Experiment 1b, success was determined post-hoc, based on filtered data, using a distribution-free criterion, delay was treated by windowing, while threshold and mapping were also fixed.

Experiment 1a used 33s long NF epochs that contained 66 observations (2Hz sampling frequency), therefore we applied the  $t$  critical value for  $p = .05$  (two-tailed) with 65 degrees of freedom ( $df$ ) for each block:  $t_{crit}(65) = 2.00$ . The experimental software logged asymmetry values during the NF epoch, calculated the  $t$  value, and if it was larger than 2, the block was deemed successful. The experimental software did not test the parametric assumption of normality; however, post-hoc analyses using bootstrapping resampling method resulted in accepting the same epochs as successful.

$$\begin{aligned}
\text{(a)} \quad t &= \frac{\bar{x} - \mu_0}{SD/\sqrt{n}} & \text{(b)} \quad r &= \sqrt{\frac{t^2}{t^2 + df}} & \text{(c)} \quad d &= \frac{\overline{NF} - \overline{Ref}}{\sqrt{\frac{(n_{NF} - 1)SD_{NF}^2 + (n_{Ref} - 1)SD_{Ref}^2}{n_{NF} + n_{Ref} - 2}}}
\end{aligned}$$

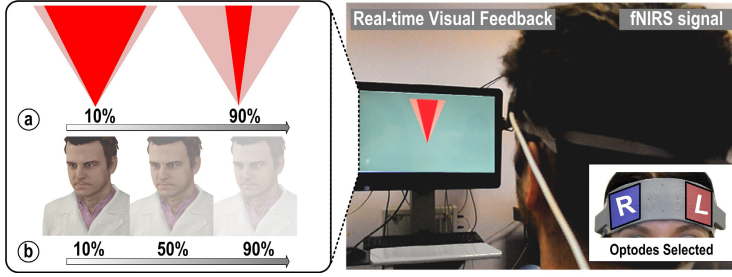
**Fig. 1.** Equations used in real-time implementation of NF success, where  $\bar{x}$  is the mean of observed values, the test value  $\mu_0$  is zero,  $SD$  is the standard deviation,  $n$  is the number of observed values, and  $df$  is degrees of freedom.  $\overline{NF}$  and  $\overline{Ref}$  are the mean of asymmetry values during NF and the reference epoch, respectively.



**Fig. 2.** Block design for Experiments 1a and 1b, where NF success is evaluated against baseline. Note that in 1b, an inter-trial interval (ITI) is added to the end to allow for windowing; baseline in 1b is measured during the reference task (see text), decreasing block length.

We argue that the computationally more demanding bootstrapping method should be favoured for post-hoc analysis, but in cases where real-time determination of NF success is important, the simple parametric criterion is sufficient. Experiment 1a also used the magnitude of the NF signal in a successful epoch as graded input to a computer system (mapped to the differential weighting of a search algorithm [5]). This was achieved by characterising the magnitude of NF success by calculating the effect-size measure  $r$  (Fig. 1b). Note that this calculation is not computationally demanding; therefore, it can be applied in real-time too. The effect-size measure  $r$  is interpreted the same as the correlation coefficient, it mitigates the difficulty of comparing fNIRS signals across subjects and blocks [21], and since its value is constrained between 0 and 1, it is convenient for mapping to graded input.

Another advantage of calculating an effect-size measure when a statistical criterion is applied to determine NF success is that it allows for evaluating the sensitivity of the set-up to detect changes in the asymmetry signal by quantifying the magnitude of increase in asymmetry the applied statistical criterion can detect. For example, in Experiment 1a, the smallest effect-size associated with a



**Fig. 3.** Examples of continuous visual feedback in (a) Experiment 1a and (b) Experiment 2a.

successful block was  $r = .28$ , which demonstrates that we could reliably detect medium effect-sizes with 33s long NF epochs. Issues related to low power (i.e. asymmetry increases but it is not detected) can be mitigated by increasing the number of observations in an NF epoch, either by increasing its length or the sampling frequency.

Threshold for providing feedback during the NF epoch (i.e., the minimum reinforced signal magnitude) and mapping the magnitude of the left-asymmetry signal to feedback was fixed in both experiments, that is, each subject in each trial received feedback using a set of pre-defined parameters. The visual feedback channel in both experiments was a downward-pointing red triangle symbolising a light beam, which could be narrowed by up-regulating left-asymmetry (Fig. 3a). This feedback was conceptually related to the experimental context, which involved speeding up an algorithmic search process using a BCI [5]<sup>2</sup>. As discussed above, the threshold was zero (i.e., no increase from baseline). We defined the maximum value for feedback empirically in a pilot study ( $1.1\mu\text{Mol/L}$ ), using a similar design under a no-feedback paradigm. Asymmetry values between the threshold and maximum were mapped linearly to the width of the light beam (updated with the same 2Hz frequency of the signal acquisition), which allows for providing continuous feedback. We successfully applied a similar feedback mapping strategy before in EEG-based NF [7].

A disadvantage of comparing asymmetry scores to a simple baseline to determine NF success is related to the difficulty of interpreting the increase in asymmetry during the NF epoch. As mentioned above, we measured a 10s baseline before each NF epoch, which defined the asymmetry as zero for the start of NF. However, the appropriateness of this baseline is predicated on the assumption that the asymmetry signal is at neutral level when the baseline is taken. Should the baseline be measured when there is high left-asymmetry, the reference-point zero at the start of NF would represent a state of already high left-asymmetry, making it difficult to detect left-asymmetry increase during NF.

<sup>2</sup> In short, this experiment used PFC left-asymmetry as an indicator of approach-related motivational tendency, whose value was mapped to speeding up the behaviour of a search algorithm.

To overcome this, we designed data-collection blocks in Experiment 1a in the following way. Each block consisted of three epochs: NF, Count and Rest (Fig. 2a). Baseline for an NF epoch was measured in the last 10s of the preceding Rest epoch, where subjects were instructed to look at a grey screen and relax. An epoch with a mental counting task was included after each NF epoch to distract subjects’ attention from the thinking strategy used during NF and to promote asymmetry converging to baseline before baseline for the next block would be taken. We elected to use a mental counting task (counting backwards from a given number by increments of a given integer), because it is theoretically unrelated to left-asymmetry and it is one of the most commonly used prefrontal activities for fNIRS-based BCI [18].

In Experiment 1b, we modified the block design by excluding the Rest epoch and measuring baseline for the next block in the last 10s of the Count epoch following NF (Fig. 2b). This reduced block length, allowing for including more blocks in the same protocol. Furthermore, including the counting task during baseline measurement is a more strict control of subjects’ mental activity than the rest instructions. Additionally, Experiment 1b did not use the NF signal for graded input to modify the behaviour of a system; therefore we determined NF success post-hoc using bootstrapping (1000 samples, 95% confidence intervals) on filtered data: we applied sliding-window motion artefact detection (SMAR), raw data were low-pass filtered using a finite impulse response (FIR) filter with order 20 and 0.1Hz cut-off frequency [1].

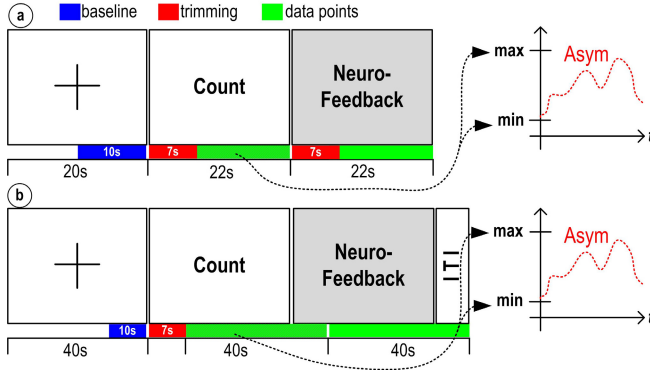
In summary, determining NF success by comparing to baseline allows for a simple block structure, but the baseline may not reflect a meaningful reference point to determine NF success if thought processes are not controlled during baseline measurement. Including a reference task for baseline may alleviate this, but other potential issues remain, for example, the perceptual differences of the stimulus subjects receive during baseline and NF, and having to rely on the same (pre-defined) criteria across blocks and subjects to provide feedback. In the next section, we discuss how including a reference epoch may improve study design.

## 4.2 Evaluating NF Success Against a Reference Epoch

As discussed above, although comparing an NF epoch to baseline for defining success is simple and time efficient, it is still useful to include a reference task either to promote up-regulated left-asymmetry to converge to baseline between blocks (Experiment 1a), or to control thought processes during baseline (Experiment 1b). However, both leave the data collected for the reference task under-analysed. A step forward to better utilising the collected data is to include the reference task in a separate epoch that is directly compared to the NF epoch. We implemented this in two experiments.

In Experiment 2a (Fig. 4a), success was determined real-time, based on data filtered for extreme values, using a parametric criterion, delay was treated by trimming, while threshold and mapping for the feedback channel were individualised (for each subject and each experimental trial). By comparison, in Experiment 2b (Fig. 4b), success was determined post-hoc, on filtered data, using a





**Fig. 4.** Experiments 2a and 2b, where NF success is evaluated against a reference epoch. Note that in 2b, an inter-trial interval (ITI) is added to the end to allow for windowing. Feedback range in the NF epoch is determined using the distribution of asymmetry values during the reference epoch in each block to allow for individualised feedback.

distribution-free criterion; delay was treated by windowing, with individualised threshold and mapping. Both Experiments 2a and 2b applied the same visual stimulus across reference and NF epochs (but the visual stimulus was different across the experiments). The epochs were matched for length. Subjects rated the perceived difficulty of both the counting and NF tasks; statistical analysis revealed no significant difference in subjective difficulty, indicating that the two tasks were adequately matched. Baseline was measured at the start of each block under instructions to rest, but we defined NF success as a statistically significant increase in asymmetry from the reference epoch to the NF epoch.

A notable advantage of this approach to determining NF success is that it provides a control condition within the block; therefore, increase in asymmetry in this case can be readily attributed to change in mental activity. Furthermore, the asymmetry signal does not need to be at neutral level when the baseline is measured, because the success criterion only considers difference between the two epochs matched in length and stimulus.

Experiment 2a applied a real-time success criterion: the experimental software conducted an independent-samples t-test upon the completion of both epochs. Although the asymmetry values were collected from the same subject within the same block (with the same baseline), an independent-samples design is appropriate here, because the subject of analysis is the two population of asymmetry scores. The t value was calculated using unfiltered data; however, the experimental software removed outliers in each epoch (values outside three standard deviations from the mean) for the calculation, which can effectively remove noise resulting from movement artefacts [2]. Since effective epoch-length was 15s (trimmed), which contained at least 29 observations for each epoch

sampled at 2Hz, the software used the  $t$  critical value of 2.05 with 28 degrees of freedom for  $p$  (two-tailed) = .05 as a threshold for success.

Conversely, Experiment 2b applied a post-hoc success criterion, where the  $t$ -test was calculated on filtered data (40s effective epoch length, windowed). In addition to the SMAR and FIR filters described in the previous section, we applied linear detrending on data from each channel [1]. Significance testing was conducted using a distribution-free approach.

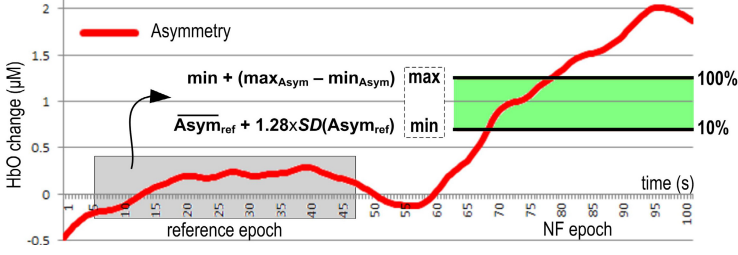
In Experiments 1a and 1b, we used a ‘one-size-fits-all’ model for providing feedback, based on an empirically determined threshold for maximum feedback that was the same of each participant within each block. This approach promotes comparability of asymmetry values across blocks and subjects, but it does not consider individual differences, which can be quite substantial [21]. However, by analysing the distribution of asymmetry scores in a reference epoch, it is possible to devise personalised feedback within each block, taking into consideration the normal fluctuation of the asymmetry signal during a reference task. We recommend using a reference task (e.g., mental counting) for baseline if a simple set-up is preferred or there is an emphasis on collecting a large number of blocks from each participant. Otherwise, it is preferable to use a reference epoch with similar perceptual properties as the NF task, but including a different mental activity; this provides experimental control within a block, and allows for the application of individualised feedback to promote decreasing noise and increasing NF success. We implemented this in Experiments 2a and 2b in the following way.

We defined threshold for providing feedback during the NF epoch based on the asymmetry values collected during the reference epoch within the same block (Fig. 5). The threshold was defined as the mean of asymmetry values during the reference epoch plus 1.28 times their standard deviation<sup>3</sup>. Assuming normally distributed asymmetry values, this threshold would result in reinforcing only the top 10% of asymmetry values in the reference epoch. This approach to determine threshold is consistent with the original one of Rosenfeld et al. [19] for EEG-based frontal-asymmetry NF.

We provided continuous feedback, based on real-time changes in the magnitude of the asymmetry signal. For example, in Experiment 2a, the feedback channel was the image transparency of a virtual character (Fig. 3b), who was previously identified as mischievous, and the experimental subjects could make his image disappear from a virtual scene by expressing anger towards him, thereby up-regulating left-asymmetry [13]. Crossing the threshold during NF was mapped to 10% transparency of the virtual character, while reaching an empirically determined maximum asymmetry was mapped to 100% transparency, effectively removing the virtual character from the scene. The maximum asymmetry value for mapping was defined as the threshold plus the variation range of the asymmetry values during the reference epoch. Visual transparency was mapped linearly between the threshold and maximum value, updated with the same 2Hz frequency of the collection of asymmetry values.

---

<sup>3</sup> Outliers were removed by the experimental software to avoid extreme values, likely reflecting movement artefacts, exerting an undue influence on the threshold.



**Fig. 5.** Calculation of threshold (minimum) and maximum asymmetry values for linear mapping to continuous feedback. Note that the minimum feedback value is set to 10%.

This approach of using a reference epoch to determine feedback range promotes NF success by tailoring the feedback mechanism to the individual subject by considering her own signal variation, with the additional benefit of determining the range of noise in the signal for which no feedback should be provided. Note that in this approach the reference epoch necessarily precedes the NF epoch.

Additionally, we calculated the Cohen’s  $d$  effect-size measure to quantify the magnitude of NF success, which is characterised as the difference between mean asymmetry during the two epochs divided by the pooled standard deviation. This was also calculated real-time in Experiment 2a (Fig. 1c). The  $d$  value reflects the distance between the distribution of asymmetry values between the reference and NF epoch within the same block, which can be readily interpreted. For example, average  $d$  in successful blocks in Experiment 2a was 2.40, which corresponds to an average 23% overlap in asymmetry scores between NF and reference epochs, and there is a 96% chance that an asymmetry value picked randomly from the NF epoch will be larger than a randomly picked asymmetry value from the reference epoch. Calculating these measures can be useful for illustrating the magnitude of asymmetry up-regulation in NF. Although the magnitude of oxygenation changes can differ substantially across subjects and blocks, this approach relies on the distribution of observed asymmetry values within blocks, therefore data collected from different individuals at different times are comparable.

## 5 Summary of Recommendations

We presented two sets of experiments where we applied variations on protocol design to support fNIRS-based PFC-asymmetry NF for active BCI. In this section, we briefly summarise the key points and present recommendations for protocol design. Table 1 illustrates these points through a comparison of two experiments.

*Protocol Length.* NF demands focused attention and concentration (but it is also rewarding and interesting), which leads to fatigue over time: approximately 40s long epochs and 2min per block are manageable, provided there is enough rest between the blocks. A total protocol length of approximately 10min is comfortable, where fatigue towards the end that does not significantly impair the quality

**Table 1.** Summary of the key points illustrated through contrasting two experiments.

	Exp. 1a	Exp. 2b
Reference epoch	No	Yes (counting)
Threshold	0 asymmetry	Dynamic ( $M + 1.28 \cdot SD$ )
Maximum	Fixed (1.1)	Dynamic (min+range)
Statistical test	Parametric	Bootstrapping
Success test	Real-time	Post-hoc
Filtering	No	Yes (FIR, SMAR, detrending)
Delay treatment	Remove 7s	Windowing
Practice	3 blocks	1 block
Number of blocks	6	8
Number of subjects	11	10
Success rate <sup>1</sup>	73%	70%

<sup>1</sup> Indicates the percentage of subjects achieving NF success in at least half of the completed experimental blocks [5,7].

of data (analysis revealed that NF success was not significantly less likely towards the end of the protocol in the studies). A shorter block length is generally preferable, which allows for including more blocks in a protocol. Based on post-use interviews with 42 subjects who participated in our experiments, we advise to determine the length and number of blocks so that data collection fits within approximately 15–20min with instructions, practice, set-up and calibration.

*Compensate for Delay in the Haemodynamic Response.* This can be achieved by simply removing initial observations in an epoch or by windowing. Simple removing may be considered a safer solution when it is important to make sure that each data point was collected when feedback was present, but windowing helps to better utilise the collected data by boosting statistical power with increased sample size.

*Using a Reference Task.* We recommend using a reference task (e.g., mental counting) for baseline if a simple set-up is preferred or there is an emphasis on collecting a large number of blocks from each participant. Otherwise, it is preferable to use a reference epoch with similar perceptual properties as the NF task; this provides experimental control within a block, and allows for the application of individualised feedback.

*Using a Statistical Success Criterion.* Rather than relying solely on statistical significance, we advise to quantify NF success (e.g., by calculating effect-size measures). If the NF signal also serves as input and a real-time success criterion is required, a parametric approach is sufficiently robust, but simple data-screening should still be applied (e.g., removing outliers). Otherwise, a post-hoc success criterion should be preferred on filtered data and using a distribution-free method. Conduct a power analysis to inform the necessary length for the NF epoch.

*Continuous, Real-Time Feedback.* Feedback can be provided with the same frequency as the input signal is collected. With an empirically determined threshold and signal range, the feedback can reflect continuous variations of the input

signal, as opposed to a limited set of categories. We recommend using a feedback channel that is conceptually or perceptually related to the NF task. Participants can accommodate delay in the feedback, but they need to be informed in advance to expect some delay. Additionally, we found it useful to instruct participants that some jitter may be also present in the feedback; however, this was not problematic enough to introduce smoothing to reduce fluctuations in the NF signal (e.g., moving average [6,28]).

## References

1. Ayaz, H.: Functional near infrared spectroscopy based brain computer interface. PhD thesis, School of Biomedical Engineering Science & Health Systems, Drexel University (2010)
2. Ayaz, H., Izzetoglu, M., Shewokis, P.A., Onaral, B.: Sliding-window motion artifact rejection for functional near-infrared spectroscopy. In: Engineering in Medicine and Biology Society (EMBC), pp. 6567–6570. IEEE (2010)
3. Bunce, S.C., Izzetoglu, M., Izzetoglu, K., Onaral, B., Pourrezaei, K.: Functional near-infrared spectroscopy. IEEE Engineering in Medicine and Biology Magazine **25**(4), 54–62 (2006)
4. Carver, C.S., Harmon-Jones, E.: Anger and approach: Reply to watson (2009) and to tomarken and zald (2009). Psychological Bulletin **135**(2), 215–217 (2009)
5. Cavazza, M., Aranyi, G., Charles, F.: Brain-computer interfacing to heuristic search: first results. In: Vicente, J.M.F., Álvarez-Sánchez, J.R., de la Paz López, F., Toledo-Moreo, F.J., Adeli, H. (eds.) IWINAC 2015, Part I. LNCS, vol. 9107, pp. 312–321. Springer, Heidelberg (2015)
6. Cavazza, M., Aranyi, G., Charles, F., Porteous, J., Gilroy, S., Jackont, G., Klovatch, I., Raz, G., Keynan, N.J., Cohen, A., et al.: Frontal alpha asymmetry neurofeedback for brain-computer interfaces. In: Brain-Computer Interface Conference (2014)
7. Cavazza, M., Aranyi, G., Charles, F., Porteous, J., Gilroy, S., Klovatch, I., Jackont, G., Soreq, E., Keynan, N.J., Cohen, A., Raz, G., Hendler, T.: Towards empathic neurofeedback for interactive storytelling. In: Finlayson, M.A., Meister, J.C., Bruneau, E.G. (eds.) 2014 Workshop on Computational Models of Narrative of OpenAccess Series in Informatics (OASIS), Dagstuhl, Germany, vol. 41, pp. 42–60. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik (2014)
8. Doi, H., Nishitani, S., Shinohara, K.: NIRS as a tool for assaying emotional function in the prefrontal cortex. Frontiers in Human Neuroscience **7**, Article 770 (2013)
9. Esfahani, E.T., Sundararajan, V.: Using brain-computer interfaces to detect human satisfaction in human-robot interaction. International Journal of Humanoid Robotics **8**(01), 87–101 (2011)
10. Field, A.: Discovering statistics using SPSS. Sage (2009)
11. Garcia-Molina, G., Tsoneva, T., Nijholt, A.: Emotional brain-computer interfaces. International Journal of Autonomous and Adaptive Communications Systems **6**(1), 9–25 (2013)
12. Gilroy, S.W., Porteous, J., Charles, F., Cavazza, M., Soreq, E., Raz, G., Ikar, L., Or-Borichov, A., Ben-Arie, U., Klovatch, I., et al.: A brain-computer interface to a plan-based narrative. In: International Joint Conference on Artificial Intelligence (IJCAI), pp. 1997–2005. AAAI Press (2013)

13. Harmon-Jones, E.: Clarifying the emotive functions of asymmetrical frontal cortical activity. *Psychophysiology* **40**(6), 838–848 (2003)
14. Karran, A.J., Fairclough, S.H., Gilleade, K.: A framework for psychophysiological classification within a cultural heritage context using interest. *ACM Transactions on Computer-Human Interaction (TOCHI)* **21**(6), Article 34 (2015)
15. Lindquist, K.A., Wager, T.D., Kober, H., Bliss-Moreau, E., Barrett, L.F.: The brain basis of emotion: a meta-analytic review. *Behavioral and Brain Sciences* **35**(03), 121–143 (2012)
16. Morinaga, K., Akiyoshi, J., Matsushita, H., Ichioka, S., Tanaka, Y., Tsuru, J., Hanada, H.: Anticipatory anxiety-induced changes in human lateral prefrontal cortex activity. *Biological Psychology* **74**(1), 34–38 (2007)
17. Mühl, C., Heylen, D.: Cross-modal elicitation of affective experience. In: 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, ACII 2009, pp. 1–12. IEEE (2009)
18. Naseer, N., Hong, K.: fNIRS-based brain-computer interfaces: a review. *Frontiers in Human Neuroscience* **9**, Article 3 (2015)
19. Rosenfeld, J.P., Cha, G., Blair, T., Gotlib, I.H.: Operant (biofeedback) control of left-right frontal alpha power differences: Potential neurotherapy for affective disorders. *Biofeedback and Self-Regulation* **20**(3), 241–258 (1995)
20. Ruocco, A.C., Rodrigo, A.H., Lam, J., Di Domenico, S.I., Graves, B., Ayaz, H.: A problem-solving task specialized for functional neuroimaging: validation of the Scarborough adaptation of the Tower of London (S-TOL) using near-infrared spectroscopy. *Frontiers in Human Neuroscience* **8**, Article 185 (2014)
21. Sakatani, K., Takemoto, N., Tsujii, T., Yanagisawa, K., Tsunashima, H.: Nirs-based neurofeedback learning systems for controlling activity of the prefrontal cortex. In: *Oxygen Transport to Tissue XXXV*, pp. 449–454. Springer (2013)
22. Sarkheil, P., Zilverstand, A., Kilian-Hütten, N., Schneider, F., Goebel, R., Mathiak, K.: fMRI feedback enhances emotion regulation as evidenced by a reduced amygdala response. *Behavioural Brain Research* **281**, 326–332 (2015)
23. Sitaram, R., Caria, A., Birbaumer, N.: Hemodynamic brain-computer interfaces for communication and rehabilitation. *Neural Networks* **22**(9), 1320–1328 (2009)
24. Solovey, E.T., Girouard, A., Chauncey, K., Hirshfield, L.M., Sassaroli, A., Zheng, F., Fantini, S., Jacob, R.J.: Using fNIRS brain sensing in realistic HCI settings: experiments and guidelines. In: *The 22nd Annual ACM Symposium on User Interface Software and Technology*, pp. 157–166. ACM (2009)
25. Spielberg, J.M., Miller, G.A., Warren, S.L., Engels, A.S., Crocker, L.D., Banich, M.T., Sutton, B.P., Heller, W.: A brain network instantiating approach and avoidance motivation. *Psychophysiology* **49**(9), 1200–1214 (2012)
26. Tuscan, L.-A., Herbert, J.D., Forman, E.M., Juarascio, A.S., Izzetoglu, M., Schultheis, M.: Exploring frontal asymmetry using functional near-infrared spectroscopy: a preliminary study of the effects of social anxiety during interaction and performance tasks. *Brain Imaging and Behavior* **7**(2), 140–153 (2013)
27. Wehbe, R.R., Kappen, D.L., Rojas, D., Klauser, M., Kapralos, B., Nacke, L.E.: EEG-based assessment of video and in-game learning. In: *CHI 2013 Extended Abstracts on Human Factors in Computing Systems*, pp. 667–672. ACM (2013)
28. Zotev, V., Krueger, F., Phillips, R., Alvarez, R.P., Simmons, W.K., Bellgowan, P., Drevets, W.C., Bodurka, J.: Self-regulation of amygdala activation using real-time fmri neurofeedback. *PloS One* **6**(9), e24522 (2011)

Symbiotic Interaction

4th International Workshop, Symbiotic 2015, Berlin,  
Germany, October 7-8, 2015, Proceedings

Blankertz, B.; Jacucci, G.; Gamberini, L.; Spagnolli, A.;  
Freeman, J. (Eds.)

2015, VIII, 180 p. 57 illus., Softcover

ISBN: 978-3-319-24916-2