

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

## Decoding overt shifts of attention in depth through pupillary and cortical frequency tagging

**This is a pre print version of the following article:**

*Original Citation:*

*Availability:*

This version is available <http://hdl.handle.net/2318/1807166> since 2021-09-30T00:28:25Z

*Published version:*

DOI:10.1088/1741-2552/ab8e8f

*Terms of use:*

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

# Decoding overt shifts of attention in depth through pupillary and cortical frequency tagging

Niccolò Zovetti<sup>1,2§</sup>, Tatiana Baroni<sup>1§</sup>, Silvestro Roatta<sup>3</sup>, Claudio de'Sperati<sup>1,4</sup>

<sup>1</sup> Laboratory of Action, Perception and Cognition, Faculty of Psychology, Vita-Salute San Raffaele University, Milan, Italy

<sup>2</sup> Section of Psychiatry, Department of Neurosciences, Biomedicine and Movement Sciences, University of Verona, Verona, Italy

<sup>3</sup> Laboratory of Integrative Physiology, Department of Neuroscience, University of Turin, corso Raffaello 30, 10125 Torino, Italy

<sup>4</sup> Experimental Psychology Unit, Division of Neuroscience, IRCCS San Raffaele, Milan, Italy

§ These authors contributed equally to the work. E-mail: [desperati.claudio@univr.it](mailto:desperati.claudio@univr.it)

## Abstract

We have recently developed a prototype of a novel human-computer interface for assistive communication based on the voluntary shift of attention (gaze) from a far target to a near target associated with a decrease of pupil size (Pupillary Accommodative Response, PAR), an automatic vegetative response that can be easily recorded. In 18 healthy volunteers, here we further investigate the possibility of decoding attention shifts in depth by exploiting the evoked oscillatory responses of the pupil (Pupillary Oscillatory Response, POR, recorded through a low-cost device) and visual cortex (Steady-State Visual Evoked Potentials, SSVEP, recorded from 4 scalp electrodes). With a simple binary communication protocol (focusing on a far target meaning "No", focusing on the near target meaning "Yes"), we aimed at discriminating when observer's overt attention (gaze) shifted from the near to the near target, which were flickering at different frequencies. By applying a binary linear classifier (Support Vector Machine, SVM, with leave-one-out cross validation) to POR and SSVEP signals, we found that, with only twenty trials and no subjects' training, the offline median decoding accuracy was 75% and 80% with POR and SSVEP signals, respectively. When the two signals were combined together, accuracy reached 83%. The number of observers for whom accuracy was higher than 70% was 11/18, 12/18 and 14/18 with POR, SVVEP and combined features, respectively. A signal detection analysis confirmed these results. The present findings suggest that exploiting frequency tagging with pupillary or cortical responses during an attention shift in the depth plane, either separately or combined together, could be a viable method to communicate with Complete Locked-In Syndrome (CLIS) patients when oculomotor control is unreliable and traditional assistive communication, even based on PAR, is unsuccessful.

## 1. Introduction

The Locked-In Syndrome (LIS) is a condition in which a patient is aware but unable to move or communicate through standard methods due to an almost total and debilitating paralysis. This invalidating condition can arise from various diseases and complications such as poisoning, brainstem strokes, circulatory system diseases, overdoses, brain stem specific lesions and motor neuron diseases such as ALS (Pasqualotto et al., 2012). LIS can be difficult to diagnose and in its later and most severe stages is often referred to as Complete Locked-in syndrome (CLIS), a condition that mimics loss of consciousness, impairs breathing and makes communication virtually impossible (Laureys et al., 2005) (Smith and Delargy, 2005, Leon- Carrion et al., 2002)(Leon-Carrion et al., 2002). The distinction between LIS and CLIS is important because while in the former residual voluntary control over few muscles is spared (e.g., eye movement, eye blinks, twitches with the lip) in the latter motor control is completely lost (Birbaumer, 2006). Despite recent scientific progress, current therapies are far from being able to restore muscular control or communication capabilities in these patients, and are often aimed at improving survival (Zarei et al., 2015). For these reasons, Brain-Computer Interfaces (BCI) are extremely important, because through invasive (e.g., implantable electrodes) or non-invasive (e.g., electroencephalogram) methods these devices are developed specifically with the goal of restoring patients' communication capabilities (Holz et al., 2015).

During the last twenty years, many BCIs have been developed and tested on healthy subjects and patients, often showing promising results. Most of these prototypes seem to be suited only for LIS patients, that is, individuals where at least a minimal muscular or ocular control is spared (Hinterberger et al., 2003) (Kübler et al., 2009) (Vansteensel et al., 2016). Despite the severity of the CLIS condition where both oculomotor and skeletomotor control are lost, recent studies have examined possible BCI applications for completely paralyzed patients (i.e., CLIS patients) using non-invasive techniques (Lesenfants et al., 2014a) (Allison et al., 2008) (Chaudhary et al., 2016) (Chaudhary et al., 2017) (Marchetti and Priftis, 2015, Lesenfants et al., 2014b)(Lesenfants et al., 2014). However, these techniques often require cognitive load or a long, tedious and tiresome training that could affect the usability of the BCI (Pfurtscheller et al., 2003) (Barbosa et al., 2016) (Onishi et al., 2017). BCIs developed for CLIS patients are called independent BCIs because are independent of muscular activity. They are usually based on the real-time analysis of specific neural or biological signal such as the Steady-State Visual Evoked Potentials (SSVEPs), which are oscillatory cortical responses elicited by flickering visual stimuli (Allison et al., 2008) and other electroencephalographic signals such as the P300 (Onishi et al., 2017), and many others (e.g., (Chaudhary et al., 2017)).

The analysis of pupil behaviour, which is regulated by the autonomic nervous system, also offers possible cues for developing independent human-computer interfaces. Stoll and colleagues (Stoll et al., 2013) exploited the increase in pupil size associated to mental effort, through which they could reliably establish binary communication with three LIS patients and one patient in minimally conscious state. The task consisted of performing a mental calculus in one of two temporal intervals, according to whether the intended response was a “Yes” or a “No” (2-alternative forced choice). A different approach was taken by Mathôt and colleagues (Mathôt et al., 2016), who exploited the pupil response associated to attention allocation (Mathôt et al., 2013) (Binda et al., 2013). In this case, healthy observers were trained to mentally focus on bright or dark targets to produce corresponding pupil size changes, which in turn were used to control a speller.

Very recently, we proposed a novel approach to decode a yes/no response through pupil behaviour associated to gaze shifts in depth, a component of the triadic accommodative response – which, in addition to pupillary constriction, includes lens curvature changes and vergence eye movements (Von

Noorden and Campos, 1996). Based on this reflex, whenever we overtly focus visuo-spatial attention between objects placed at different depth planes, the pupil changes accordingly, namely, it constricts in response to a far-to-near shift and dilates in response to a near-to-far shift. Importantly, the shift can be voluntary and does not necessarily require an eye movement, provided that it is performed monocularly between two targets aligned with the eye's sight line. This approach based on shifting attention in depth is very natural and does not require any special training. By associating a “Yes” response to focusing attention to a near target and a “No” response to focusing attention to a far target, in a sample of healthy observers we achieved 100% of correct responses at 10 bits/min and 96% at 15 bits/min (Ponzio et al., 2019).

Even though the autonomic pupillary control is largely spared in ALS (Baltadzhieva et al., 2005), in LIS and especially CLIS patients the pupil accommodative response may not be always evident (Roatta et al., 2019). Moreover, the pupil light response may be more sensitive to covert shifts of attention than the pupil accommodative response (Mathôt, 2018). Therefore, we sought to further investigate the approach based on attention shift in depth but exploiting a frequency tagging method, and integrating pupillary and cortical signals. To this end, we measured the pupil (Pupil Oscillatory Response, POR) and the cortical (Steady-State- Visual-Evoked-Potentials, SSVEPS) responses to flickering visual stimuli – i.e. responses to light intensity modulation. As to the former (POR), we relied on the capability of the pupil to change its size in response to oscillatory changes in stimulus luminance: pupil size oscillates at the stimulus frequency and these evoked oscillations are enhanced when covert attention is directed to the oscillating stimulus (Naber et al., 2013). By placing two targets in front of an observer, one near and one far, each with a distinct luminance oscillation frequency, it is expected that focusing attention to one or the other target would result in the pupil oscillating more intensely at one or the other frequency, respectively (pupil frequency tagging). As to the latter (SSVEPs), we relied on cortical frequency tagging associated to gaze shifts in depth (Cotrina et al., 2017) (Cotrina et al., 2015). These authors developed a BCI where the evoked cortical oscillations reflected the luminance oscillation frequency of the far or near target, depending on which target the observer's gaze focused to.

In short, we aimed at decoding attention shifts from a far to a near target by relying on multiple signal sources, namely, cortical and pupillary evoked oscillations, and frequency tagging method. Preliminary results have been presented in abstract form (de'Sperati et al., 2019).

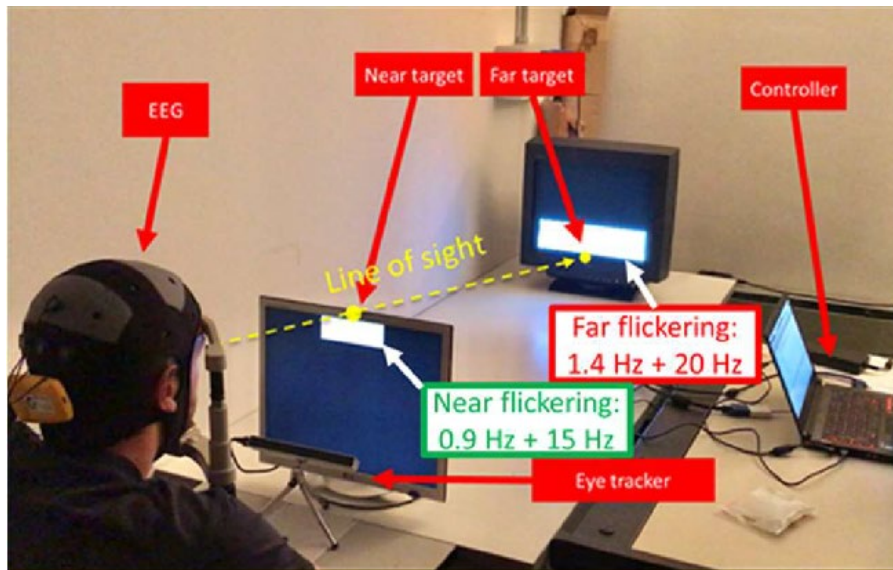
## **2. Materials and Methods**

### **2.1. Participants.**

A group of twenty volunteers were recruited for this experiment. They had normal or corrected-to-normal vision. Due to technical issues, two participants were excluded from the analysis, thus the final sample was composed of 18 participants (10 females and 8 males, mean age: 26.61). None of the participants or their first-degree relatives had a history of neurological diseases. All participants received instructions about the protocol, and gave their informed consent. The study was approved by the “Comitato di Bioetica d'Ateneo”, University of Turin.

## 2.2. Stimuli and task.

The experiments took place in a moderately darkened room. Observers were seated in front of two computer screens, placed at a distance of approximately 40 cm and 180 cm, with the head placed on a chin rest. Two yellow pieces of play-doh stitched to the screen borders served as targets. The two targets were aligned with the cyclopean line of sight, thus one was at the bottom center of the furthest screen and the other at the top center of the nearest screen, in such a way to be both visible and perceptually almost aligned (Figure 1).



**Figure 1.** The experimental setup. The two yellow spots represent the targets, which were stitched to the screen borders and aligned to the cyclopean line of sight, while the white patches are the near and far flickering stimuli presented on the display. Observer held the gaze on the far target to answer “no” and shifted the gaze on the near target to answer “Yes”. Pupil size, eye movements and EEG were concurrently recorded.

A white patch (the near stimulus) was presented on the near display, flickering at 0.9 Hz and 15 Hz, and another white patch was presented on the far display (the far stimulus), flickering at 1.4 Hz and 20 Hz. Flickering was obtained by sinusoidally modulating the patch luminance, thus each stimulus contained a sum of two frequencies. Note that in this experiment we integrated POR and SSVEPs, thus both the near and the far stimulus had to flicker at two distinct frequencies simultaneously, one in a relatively low- frequency range, suited to elicit pupil oscillations (Naber et al., 2013) and one in a relatively high-frequency range, suited to elicit EEG oscillations (SSVEPs (Herrmann, 2001)). In Table 1 are reported the oscillation frequencies. The two stimuli were adjusted in order to approximatively appear as having identical size and brightness.

A trial started when both stimuli started to flicker, and lasted 15 seconds. One second before trial onset, an auditory cue reminded the observer to bring the gaze to the far target. After 7.5 seconds from trial onset, an auditory cue was presented, which could be “Yes” or “No”, indicating to shift the

gaze from the far to the near target or to keep holding the gaze on the far target, respectively. There were 10 “Yes” trials and ten “No” trials, randomly interspersed, with a 10-s inter-stimulus interval. The experiment lasted about 20 minutes, plus the time to prepare for EEG recording (see below).

	<i>POR</i>	<i>SSVEP</i>
<i>Far stimulus</i>	1.4 Hz	20 Hz
<i>Near stimulus</i>	0.9 Hz	15 Hz

**Table 1.** Flickering frequencies of the two stimuli that elicited pupillary (POR) and cortical (SSVEP) oscillatory responses. Each stimulus always flickered at both the POR and the SSVEP frequencies.

### 2.3. Signal recording and data analyses.

An eye tracker was placed at the base of the near target and recorded both eye movements and pupil size binocularly (The Eye Tribe, sampling frequency: 60 Hz). Scalp electrical activity was recorded through an Enobio device (Neuroelectronics; sampling frequency, 500 Hz). The EEG traces were recorded using 4 electrodes positioned on O1, O2, Oz, Pz (extended 10-20 system) and referenced to the right ear lobe. Electrodes were gel-based passive plates (Ag/AgCl coated; impedance < 5 kΩ), and were placed on the scalp by means of an EEG cap. Data were acquired and saved for off-line analysis.

To quantify both pupillary and cortical oscillations we used the Minimum Energy Combination method (Friman et al., 2007), which extracted an index (T-index) from the EEG signals of all electrodes as well as from the pupil size signal. The Minimum Energy Combination method amplifies the frequency response and reduces the impact of the unrelated activity and noise, thus the T-index evaluated the strength of the evoked oscillatory responses with respect to the noise (for further details, see (Gregori Grgič et al., 2016) (Friman et al., 2007)) We considered the first harmonic of each frequency. We computed the mean T-index value over two temporal windows, one before the cue (1.0 - 7.5 s after trial onset) and one after the cue (9.0 - 15.0 s after trial onset). The difference provided a measure of the change in pupil and cortical evoked oscillation strength associated to the gaze shift in depth in the “Yes” trials. By contrast, in the “No” trials, which were associated to gaze holding, the change was expected to be negligible.

In order to automatically discriminate between “Yes” and “No” trials, we trained a Support Vector Machine (SVM) to perform a two-class classification. We used the difference between near and far oscillation frequencies of each response (POR and SSVEP) as input feature. The mean T-index values were computed over the two previously- described time windows before and after the cue. The training set comprised all trials from N-1 observers, and the test set comprised all trials from the observer removed from the training set. The procedure was repeated for each observer (leave-one-out cross-validation). Classification performance was quantified subject-wise in terms of accuracy

(percentage of correct predictions) and sensitivity ( $d'$ , with correction for extreme proportions (Hautus, 1995)) for the three possible feature selections (POR, SSVEP and POR+SSVEP).

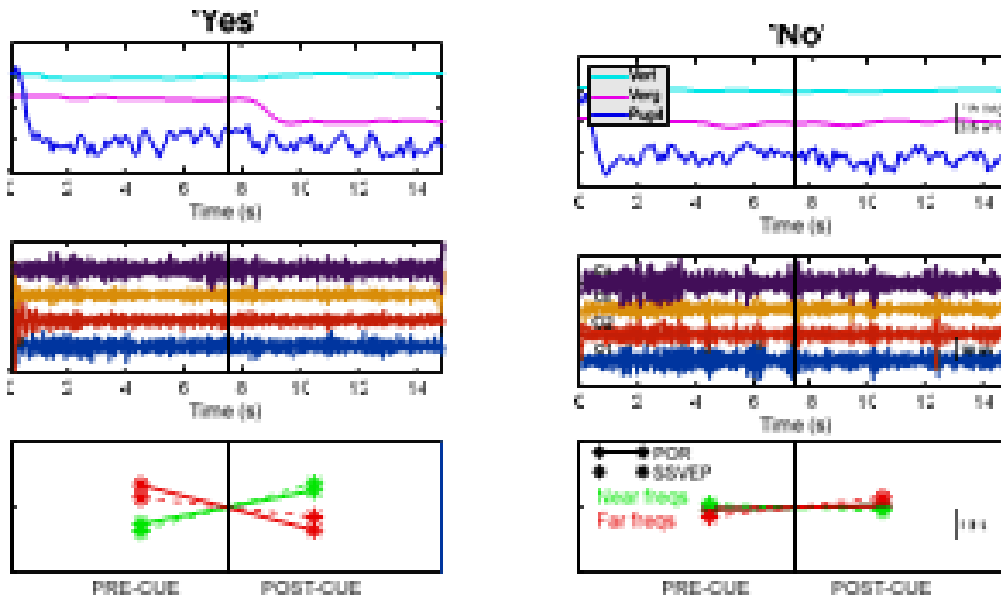
For statistical analyses, one-sample and paired Student's t-test were used, as well as ANOVA for repeated measures. No correction for sphericity was necessary. The distributions conformed to the normality assumption (Shapiro-Wilk test). The Bayes log ratio was also computed (Campbell and Thompson, 2012), which provides the strength of relative evidence for H1 (positive evidence) or H0 (negative evidence).

Stimulus administration, signal acquisition, signal processing, SVM classification and statistical analyses were performed with custom programs written in Matlab R2018b (MathWorks) with Psychtoolbox.

### 3. Results

Figure 2 illustrates examples of individual trials from one participant. An initial strong pupil light response constricting the pupil was always present, corresponding to the onset of the white flickering stimuli. In the “Yes” trial, in which the task was to shift the gaze from the far to the near target at the time of the cue, there was a clear vergence eye movement (convergence) but almost no accommodative pupil constriction (Figure 2A). The lack of a clear pupil accommodative response is likely due to the fact that the pupil was already rather constricted due to the frontal light coming from the flickering stimuli. Importantly, pupil size showed an oscillatory pattern throughout the trial, although at visual inspection it is difficult to tell the passage from a main oscillation at 1.4 Hz, before the cue, to a main oscillation at 0.9 Hz, after the cue (i.e., the flickering frequencies of the far and near stimuli, respectively). We recall that the eyes were always exposed to all four flickering frequencies and thus the pupil responded to a mixture of luminance oscillations (although two frequencies were in the high range and suited for SSVEP), whose relative strength was expected to change after the cue in “Yes” trials but not in the “No” trials.

The strength of pupil oscillation at both the far and the near flickering frequencies was extracted by computing the T-index for each frequency before and after the cue over the two above-mentioned temporal windows (1.0 - 7.5 s after trial onset and 9.0 - 15.0 s after trial onset). The T-index pattern showed a decrease of cortical oscillation strength at the far frequency after the cue, and an increase at the near frequency (Figure 2C). In the “No” trial, in which participants held the gaze on the far target, pupil oscillations did not change appreciably after the cue, and no vergence eye movements or vertical eye movements occurred (Figure 2D). Accordingly, the T-index was almost constant for both far and near frequencies (Figure 2F). SSVEPs showed a very similar behaviour (Figure 2B,C,E,F).

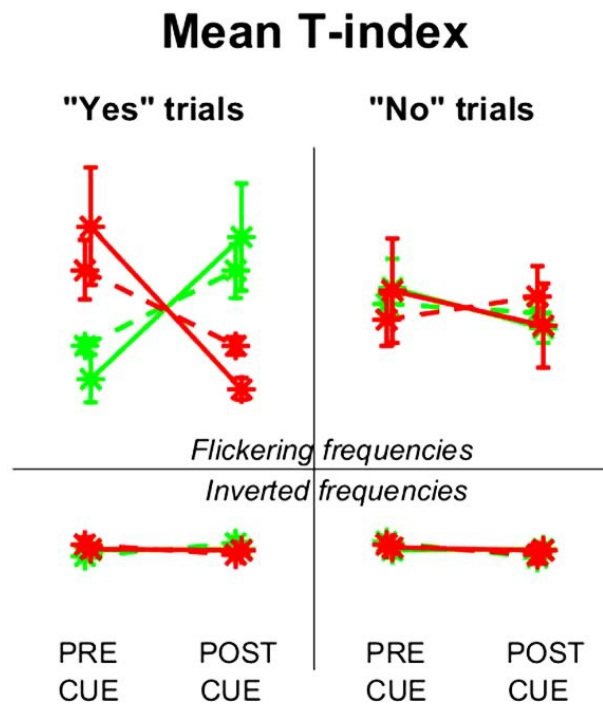


**Figure 2.** Examples of a “Yes” trial (left panels; task: shifting the gaze from the far to the near target) and a “No” trial (right panels; task: holding the gaze on the far target). A, pupil size over time (blue trace) showing an oscillatory behaviour (POR), together with vergence eye movements (magenta trace) showing convergence shortly after the “Yes” cue (vertical line), and vertical eye position (cyan trace) showing no substantial change before and after the cue. B, EEG traces recorded concurrently from 4 channels (O1, O2, Oz, Cz). C, the 1.4 Hz pupil oscillatory component, as indexed by the T-index (arbitrary units, a.u.), decreased after the cue (continuous red line, far target), while the 0.9 Hz component increased (continuous green line, near target), indicating a change in POR behaviour. An analogous change in far/near oscillation strength occurred for SSVEPs (dashed lines), although here the relevant frequencies were 20 Hz and 15 Hz, respectively. For graphical clarity, offset for T-index was set to zero for all curves. A quite different data pattern can be seen in the “No” trial (D-F), where no convergence occurred, and neither POR nor SSVEPs changed sensibly after the cue. Eye position traces were low-pass filtered at 0.3 Hz, pupil size traces were low-pass filtered at 3 Hz, and EEG traces were band-pass filtered at 5-50 Hz. Also shown are the vertical calibration bars.

These observations were confirmed in the group data, where the mean pupil oscillation strength across the 18 subjects before and after the cue was different for the near and the far flickering frequencies (Figure 3, upper plots). A 2 x 2 ANOVA for repeated measures on the T-index difference between near and far frequencies, with factors time interval (pre-cue / post-cue) and trial type (“Yes” / “No” trials) showed a highly significant interaction with very strong positive evidence [ $F(1,17) = 14.444$ ,  $p = 0.001$ , partial  $\eta^2 = 0.459$ , Bayes log ratio = 4.090]. The same analysis was applied to EEG recording, although in this case the relevant frequencies were 20 Hz and 15 Hz, which were associated to the far and near stimuli, respectively. Again, the T-index pattern was clearly different in “Yes” and “No” trials. The interaction time interval x trial type was highly significant with very strong positive evidence [ $F(1,17) = 12.501$ ,  $p = 0.003$ , partial  $\eta^2 = 0.424$ , Bayes log ratio = 3.516], thus confirming the general validity of the frequency tagging approach to distinguish far-to-near attention shifts through either pupillary or cortical signals at the population level. As a control, we switched the POR and SSVEP frequencies in the analysis, thus in fact removing any relevant information about the gaze shift. As expected, the differences between the “Yes” and “No” trials disappeared (Figure 3, lower plots).

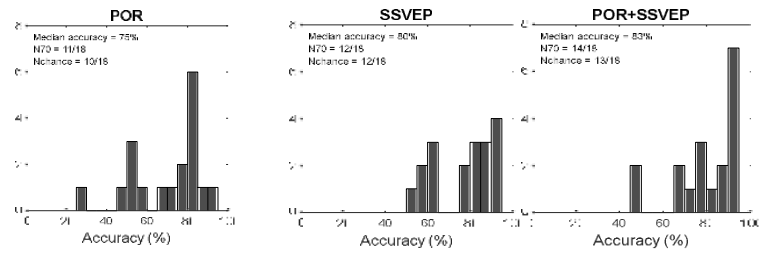


As a next step, we applied binary classification by means of a SVM to decode observers' responses ("Yes" or "No") at the individual level based on the features of pupil and cortical signals. The results are shown in Figure 4, separately for POR, SSVEP and POR+SSVEP features. Classification was quite good when using only POR as feature, yielding a 75% median accuracy – with more than half of the observers (11/18) scoring at least 70% of correct classification (here named N70). When SSVEPs was used as input feature, the classification performance improved, with 80% accuracy and 12/18 N70. By combining the two features together, accuracy reached 83%, with 14/18 N70.



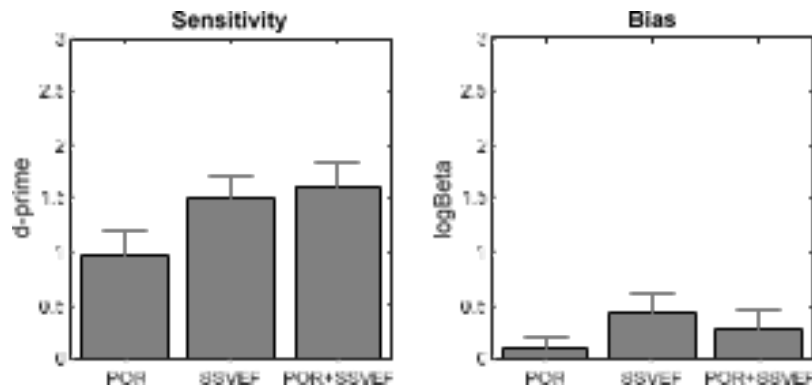
**Figure 3.** Mean POR and SSVEP strength as indexed by the average T-index before and after the cue (PRE CUE and POST CUE) in "Yes" and "No" trials. Also shown is a control condition in which POR and SSVEP frequencies were inverted in the analysis (Inverted frequencies). For graphical clarity, offset was set to zero for all curves. Bars are S.E.M. Colors, markers and line styles as in Figure 2C,F.

Classification performance was confirmed by a signal detection analysis (Figure 5). With either feature selection,  $d'$  was significantly larger than chance (one-sample Student's  $t$  test against zero, always  $p < 0.001$ ), with no significant bias except when SSVEP feature was used (one-sample Student's  $t$  tests on  $\log \beta$  against zero,  $p = 0.015$ ), suggesting the occasional presence of a weak bias in favour of the "No" classification. Moreover, there seemed to be a tendency to improve classification when passing from using POR alone to using SSVEP alone to using POR+SSVEP, although only the comparison between  $d'$  obtained with POR and  $d'$  obtained with the combined features reached statistical significance (Student's paired  $t$  test,  $p = 0.050$ ). No statistically significant differences were found when comparing the bias across the three feature selections.



**Figure 4.** Distribution of individual classification accuracy (N=18) for the three feature selections (POR, SSVEP and combined features). Also reported are the values of median accuracy and N70, i.e., the percentage of observers for whom accuracy was at least 70%.

We also verified the correlation between the classification performance and the change in the mean vertical eye position before and after the cue in the “Yes” trials, measured over the same time intervals during which the T-index was measured. It is possible, in fact, that in the “Yes” trials observers may have erroneously glimpsed to the flickering stimulus, thus spuriously increasing POR and SSVEP strength. The lack of correlation (R square = 1%,  $p = 0.752$ ) ruled out this possibility. Indeed, the difference of vertical eye position before and after the cue in “Yes” trials was negligible ( $M = 0.045$  deg,  $SD = 1.240$ ) and not significantly different from zero [one-sample Student’s t test against zero,  $t(17) = 0.154$ ,  $p = 0.880$ ], confirming the irrelevance of vertical eye movements in accounting for the T-index pattern.



**Figure 5.** Mean sensitivity and bias (N= 18) obtained with the three feature selections (POR, SSVEP and combined features). Bars are S.E.M.

#### 4. Discussion and conclusion

The present study was aimed at extending our previous work on PAR-based assistive communication (Ponzio et al., 2019) (Roatta et al., 2019) by exploring frequency tagging with multiple signals. We combined pupil frequency tagging (Stoll et al., 2013) and accommodation-induced SSVEPs modulation (Cotrina et al., 2015, Cotrina et al., 2017) to show the feasibility of a BCI for binary communication that exploits a simple voluntary response – shifting the gaze from a far to a near target as a proxy for a “Yes” response and holding the gaze on a far target as a proxy for a “No” response. The ultimate goal is to provide a tool to give the chance to those (C)LIS patients who for various reasons are not only unable to control residual eye movements but in whom even the PAR is unreliable (Roatta et al., 2019), of sustaining nonetheless a simple form of communication. For these patients, switching from a PAR-based to a POR/SSVEP-based communication – thus in fact passing from exploiting the pupil accommodative response to exploiting the modulation of the pupil light response – has the advantage that the task remains the same: shifting attention in depth. For many patients, continuing to perform the same task would be a valuable aspect, as the extinction of goal-directed thinking may hamper instrumental learning of new tasks (Kübler and Birbaumer, 2008).

A previous study also used cortical frequency tagging – but not pupil frequency tagging – to decode shifts of attention in depth with the goal of enabling binary communication (Cotrina et al., 2015, Cotrina et al., 2017). In that study, observers shifted the gaze between the two flickering targets, thus that solution did not achieve an independent BCI. By contrast, we took care of separating the targets, which were very small and aligned with the line of sight, from the stimuli, which consisted of two flickering patches located in the upper and lower visual field (Figure 1). With such configuration, observers shifted the gaze from the far to the near target without fixating the flickering stimuli, as shown by the lack of vertical eye movements before and after the cue in “Yes” trials. Thus, our study realized an independent BCI, at least as far as eye movements in the frontal plane are concerned.

The presence of vergence eye movements, however, prevents to qualify our BCI as a fully independent BCI (note that even if we had used monocular condition with one eye patched and with the two targets aligned with the seeing eye, it would have been still insufficient to qualify our approach as independent BCI because of the presence of a vergence eye movement of the fellow, patched eye). Thus, it may be wondered whether the same results would be obtained with (C)LIS patients, who may be in a condition of relying only on covert, not overt, shifts of attention (i.e. without vergence eye movements). Although only testing this approach with patients will provide a definitive answer, we recall that both the pupil light reflex and the cortical responses to luminance stimuli are modulated by covert attention shifts in the frontal plane. In the former case, the capability of the pupil to change size according to whether or not covert attention is deployed to a bright or a dark target (Binda et al., 2013, Mathôt et al., 2013) has been exploited to develop a speller (Mathôt et al., 2016). In the latter case, it is well known that the amplitude of both VEPs and SSVEPs depends on covert attention allocation (Hillyard and Anllo-Vento, 1998, Di Russo et al., 2003, Müller and Hillyard, 2000), to the extent that it has been possible to derive a gradual signal of continuous movements of covert attention, both horizontally (Gregori Grgič et al., 2016) and vertically (Gregori Grgič & de’Sperati, 2016). We expect that the same applies to our condition, that is, when covertly shifting attention in depth rather than in the frontal plane.

We remark again that the reason why we addressed attention shifts in depth in this study (and not, say, horizontal shifts) was to provide continuity with PAR-based assistive communication (Ponzio et al., 2019) (Roatta et al., 2019), which relies on changes of the focal plane. Indeed, an advantage of the PAR-based approach, as compared to other pupil-based assistive communication methods (Stoll

et al., 2013, Mathôt et al., 2013), is that it relies on a very natural “voluntary” control of pupil size and is quite fast. Being based on pupillometry, it is simple to implement and low- cost. By contrast, BCIs relying on scalp potential recordings are more expensive and less easy to implement, especially for everyday use at home. In this regard, we note that while it is true that combining POR and SSVEP features yielded the highest classification performance (83%), yet the performance when using the POR feature alone was nonetheless fairly good (75%), especially considering the 70% gold standard that has been proposed to benchmark these devices (Kübler and Birbaumer, 2008). This means that, depending on budget and usability considerations, it is possible to opt for either a simple system based on POR decoding alone or a more cumbersome but more performant system based on SSVEPs (or POR+SSVEP) decoding. We think that such a double option is a practical, valuable aspect to consider.

### **Acknowledgements**

The authors thank Regina Gregori-Grgič for help in data acquisition in an initial phase of the study, and Luca Mesin for support with EEG recordings. This research was partly funded by Compagnia di San Paolo Foundation, grant CSTO164044.

## References

- ALLISON, B. Z., MCFARLAND, D. J., SCHALK, G., ZHENG, S. D., JACKSON, M. M. & WOLPAW, J. R. 2008. Towards an independent brain–computer interface using steady state visual evoked potentials. *Clinical neurophysiology*, 119, 399- 408.
- BALTADZHIEVA, R., GUREVICH, T. & KORCZYN, A. D. 2005. Autonomic impairment in amyotrophic lateral sclerosis. *Current opinion in neurology*, 18, 487-493.
- BARBOSA, S., PIRES, G. & NUNES, U. 2016. Toward a reliable gaze-independent hybrid BCI combining visual and natural auditory stimuli. *Journal of neuroscience methods*, 261, 47-61.
- BINDA, P., PEREVERZEVA, M. & MURRAY, S. O. 2013. Attention to bright surfaces enhances the pupillary light reflex. *Journal of Neuroscience*, 33, 2199-2204.
- BIRBAUMER, N. 2006. Breaking the silence: brain– computer interfaces (BCI) for communication and motor control. *Psychophysiology*, 43, 517-532.
- CAMPBELL, J. I. D. & THOMPSON, V. A. 2012. MorePower 6.0 for ANOVA with relational confidence intervals and Bayesian analysis. *Behavior research methods*, 44, 1255-1265.
- CHAUDHARY, U., BIRBAUMER, N. & RAMOS-MURGUIALDAY, A. 2016. Brain–computer interfaces for communication and rehabilitation. *Nature Reviews Neurology*, 12, 513.
- CHAUDHARY, U., XIA, B., SILVONI, S., COHEN, L. G. & BIRBAUMER, N. 2017. Brain–computer interface–based communication in the completely locked-in state. *PLoS biology*, 15, e1002593.
- COTRINA, A., BENEVIDES, A. B., CASTILLO-GARCIA, J., BENEVIDES, A. B., ROJAS-VIGO, D., FERREIRA, A. & BASTOS-FILHO, T. F. 2017. A ssvep-bci setup based on depth-of-field. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25, 1047-1057.
- COTRINA, A., BENEVIDES, A. B., CASTILLO-GARCIA, J., FERREIRA, A. & BASTOS FILHO, T. F. 2015. Statistical evaluation of a novel SSVEP-BCI stimulation setup based on depth-of-field. *Research on Biomedical Engineering*, 31, 295-306.
- DE'SPERATI C. GREGORI-GRGIC R., ZOVETTI N., BARONI T., ROATTA, S., 2019. Exploiting pupil and cortical oscillatory responses to flickering stimuli for decoding shifts of attention in depth. *Proc. XX European Conference on Eye Movements*, 18-22 August 2019, Alicante, Spain.
- DI RUSSO, F., MARTÍNEZ, A. & HILLYARD, S. A. 2003. Source Analysis of Event-related Cortical Activity during Visuo-spatial Attention. *Cerebral Cortex*, 13, 486-499.
- FRIMAN, O., VOLOSAYAK, I. & GRASER, A. 2007. Multiple channel detection of steady-state visual evoked potentials for brain-computer interfaces. *IEEE transactions on biomedical engineering*, 54, 742-750.
- GREGORI GRGIČ, R., CALORE, E. & DE'SPERATI, C. 2016. Covert enaction at work: Recording the continuous movements of visuospatial attention to visible or imagined targets by means of Steady-State Visual Evoked Potentials (SSVEPs). *Cortex*, 74, 31-52.

- GREGORI GRGIČ, R. & DE'SPERATI, C. 2016. Covert Attention Tracking: Towards Two-Dimensional Real-Time Recordings. *Journal of Eye Movement Research*, 9(6).
- HAUTUS, M. J. 1995. Corrections for extreme proportions and their biasing effects on estimated values of  $d'$ . *Behavior Research Methods, Instruments, & Computers*, 27, 46-51.
- HERRMANN, C. S. 2001. Human EEG responses to 1– 100 Hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. *Experimental brain research*, 137, 346-353.
- HILLYARD, S. A. & ANLLO-VENTO, L. 1998. Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences*, 95, 781.
- HINTERBERGER, T., KÜBLER, A., KAISER, J., NEUMANN, N. & BIRBAUMER, N. 2003. A brain–computer interface (BCI) for the locked-in: comparison of different EEG classifications for the thought translation device. *Clinical Neurophysiology*, 114, 416-425.
- HOLZ, E. M., BOTREL, L., KAUFMANN, T. & KÜBLER, A. 2015. Long-term independent brain-computer interface home use improves quality of life of a patient in the locked-in state: a case study. *Archives of physical medicine and rehabilitation*, 96, S16-S26.
- KÜBLER, A. & BIRBAUMER, N. 2008. Brain–computer interfaces and communication in paralysis: Extinction of goal directed thinking in completely paralysed patients? *Clinical neurophysiology*, 119, 2658-2666.
- KÜBLER, A., FURDEA, A., HALDER, S., HAMMER, E. M., NIJBOER, F. & KOTCHOUBEY, B. 2009. A brain–computer interface controlled auditory event-related potential (P300) spelling system for locked-in patients. *Annals of the New York Academy of Sciences*, 1157, 90-100.
- LAUREYS, S., PELLAS, F., VAN EECKHOUT, P., GHORBEL, S., SCHNAKERS, C., PERRIN, F., BERRE, J., FAYMONVILLE, M.-E., PANTKE, K.-H. & DAMAS, F. 2005. The locked-in syndrome: what is it like to be conscious but paralyzed and voiceless? *Progress in brain research*, 150, 495-611.
- LEON-CARRION, J., VAN EECKHOUT, P. & DOMINGUEZ-MORALES MDEL, R. 2002. The locked-in syndrome: a syndrome looking for a therapy. *Brain Inj*, 16, 555-69.
- LESENFANTS, D., HABBAL, D., LUGO, Z., LEBEAU, M., HORKI, P., AMICO, E., POKORNY, C., GOMEZ, F., SODDU, A. & MÜLLER-PUTZ, G. 2014a. An independent SSVEP-based brain–computer interface in locked-in syndrome. *Journal of neural engineering*, 11, 035002.
- LESENFANTS, D., HABBAL, D., LUGO, Z., LEBEAU, M., HORKI, P., AMICO, E., POKORNY, C., GOMEZ, F., SODDU, A., MULLER-PUTZ, G., LAUREYS, S.
- & NOIRHOMME, Q. 2014b. An independent SSVEP-based brain-computer interface in locked-in syndrome. *J Neural Eng*, 11, 035002.
- MARCHETTI, M. & PRIFTIS, K. 2015. Brain–computer interfaces in amyotrophic lateral sclerosis: A metanalysis. *Clinical Neurophysiology*, 126, 1255-1263.
- MATHÔT, S. 2018. Pupillometry: Psychology, physiology, and function. *Journal of Cognition*, 1.

- MATHÔT, S., MELMI, J.-B., VAN DER LINDEN, L. & VAN DER STIGCHEL, S. 2016. The mind-writing pupil: a human-computer interface based on decoding of covert attention through pupillometry. *PloS one*, 11, e0148805.
- MATHÔT, S., VAN DER LINDEN, L., GRAINGER, J. & VITU, F. 2013. The Pupillary Light Response Reveals the Focus of Covert Visual Attention. *PLOS ONE*, 8, e78168.
- MÜLLER, M. M. & HILLYARD, S. 2000. Concurrent recording of steady-state and transient event-related potentials as indices of visual-spatial selective attention. *Clinical Neurophysiology*, 111, 1544-1552.
- NABER, M., ALVAREZ, G. A. & NAKAYAMA, K. 2013. Tracking the allocation of attention using human pupillary oscillations. *Frontiers in psychology*, 4, 919.
- ONISHI, A., TAKANO, K., KAWASE, T., ORA, H. & KANSAKU, K. 2017. Affective stimuli for an auditory P300 brain-computer interface. *Frontiers in neuroscience*, 11, 522.
- PASQUALOTTO, E., FEDERICI, S. & BELARDINELLI, M. O. 2012. Toward functioning and usable brain– computer interfaces (BCIs): a literature review. *Disability and Rehabilitation: Assistive Technology*, 7, 89-103.
- PFURTSCHHELLER, G., NEUPER, C., MULLER, G. R., OBERMAIER, B., KRAUSZ, G., SCHLOGL, A., SCHERER, R., GRAIMANN, B., KEINRATH, C. & SKLIRIS, D. 2003. Graz-BCI: state of the art and clinical applications. *IEEE Transactions on neural systems and rehabilitation engineering*, 11, 1-4.
- PONZIO, F., VILLALOBOS, A. E. L., MESIN, L., DE'SPERATI, C. & ROATTA, S. 2019. A human-computer interface based on the “voluntary” pupil accommodative response. *International Journal of Human-Computer Studies*, 126, 53-63.
- ROATTA, S., LORENZO VILLALOBOS A.E., GIULIANO S., MUSSO L., CALVO A., MOGLIA C., DE'SPERATI C. 2019. The pupil accommodative response as a communication tool in ALS patients. *Proc. XX European Conference on Eye Movements*, 18- 22 August 2019, Alicante, Spain.
- SMITH, E. & DELARGY, M. 2005. Locked-in syndrome. *Bmj*, 330, 406-409.
- STOLL, J., CHATELLE, C., CARTER, O., KOCH, C., LAUREYS, S. & EINHÄUSER, W. 2013. Pupil responses allow communication in locked-in syndrome patients. *Current biology*, 23, R647- R648.
- VANSTEENSEL, M. J., PELS, E. G. M., BLEICHNER, M. G., BRANCO, M. P., DENISON, T., FREUDENBURG, Z. V., GOSSELAAR, P., LEINDERS, S., OTTENS, T. H. & VAN DEN BOOM, M. A. 2016. Fully implanted brain–computer interface in a locked-in patient with ALS. *New England Journal of Medicine*, 375, 2060-2066.
- VON NOORDEN, G. K. & CAMPOS, E. C. 1996. *Binocular Vision and Ocular Motility: Theory and Management of Strabismus*; Mosby: St. Louis, 4, 7-442.
- ZAREI, S., CARR, K., REILEY, L., DIAZ, K., GUERRA, O., ALTAMIRANO, P. F., PAGANI, W., LODIN, D., OROZCO, G. & CHINEA, A. 2015. A comprehensive review of amyotrophic lateral sclerosis. *Surgical neurology international*, 6.