Correspondence should be addressed to s.mathot@cogsci.nl

The Mind-Writing Pupil: A Human-Computer Interface Based on Decoding of Covert Attention through Pupillometry

Sebastiaan Mathôt[1], Jean-Baptiste Melmi[1], Lotje van der Linden[1], and Stefan Van der Stigchel[2]

[1]Aix-Marseille University, CNRS, LPC UMR 7290, Marseille, France

[2]Dept. of Experimental Psychology, Helmholtz Institute, Utrecht University, The Netherlands

Abstract

We present a new human-computer interface that is based on decoding of attention through pupillometry. Our method builds on the recent finding that covert visual attention affects the pupillary light response: Your pupil constricts when you covertly (without looking at it) attend to a bright, compared to a dark, stimulus. In our method, participants covertly attend to one of several letters with oscillating brightness. Pupil size reflects the brightness of the selected letter, which allows us–with high accuracy and in real time–to determine which letter the participant intends to select. The performance of our method is comparable to the best covert-attention brain-computer interfaces to date, and has several advantages: no movement other than pupil-size change is required; no physical contact is required (i.e. no electrodes); it is easy to use; and it is reliable. Potential applications include: communication with totally locked-in patients, training of sustained attention, and ultra-secure password input.

Keywords: human-computer interface, brain-computer interface, pupillometry, covert visual attention

The Mind-Writing Pupil: A Human-Computer Interface Based on Decoding of Covert Attention through Pupillometry

A brain-computer interface (BCI) translates thought into action. BCIs provide new ways to interact with computers; importantly, they can restore the power to act and communicate in locked-in patients with little or no motor control (Birbaumer, 2006; Wolpaw, Birbaumer, McFarland, Pfurtscheller, & Vaughan, 2002).

There are many types of BCIs (reviewed in Donoghue, 2008; Nicolas-Alonso & Gomez-Gil, 2012), which differ in the neural signal that they use (e.g., neural spikes or electroencephalography [EEG]), the way that neural activity is processed (e.g., through a classifier or by measuring overall activity in specific brain areas), and the actions that they perform (e.g. controlling a robotic limb, or writing text). Here we present a new method, which uses pupil size, rather than brain activity, as the controlling signal. Our method is related to two existing methods: the P300 speller (Farwell & Donchin, 1988), which is functionally similar to our method but relies on a different controlling signal; and a recent pupillometry-based method (Stoll et al., 2013), which is functionally different from our method but relies on the same controlling signal.

P300 spellers are among the most successful BCIs (reviewed in Fazel-Rezai et al., 2012). They exploit the fact that rare visual stimuli elicit positive deflections in the EEG signal about 300 ms after their appearance. This event-related-potential (ERP) component is called the P300, and is largest for stimuli that are overtly (while looking at them) or covertly (without looking at them) attended (e.g., Treder & Blankertz, 2010). In a classic P300 speller, the participant sees a grid of letters. One letter, or sometimes one full column or row of letters, is highlighted at a time. A P300 is elicited each time that a letter is highlighted. The participant selects a letter by attending to it, usually by looking at it directly (i.e. overt attention), which leads to an increased P300 when that letter is highlighted. In its simplest form, the letter that, when highlighted, elicits the highest P300 is selected; however, most P300 spellers now use sophisticated classification techniques, and pool information from multiple ERP components (Krusienski et al., 2006). But the principle remains the same.

The first P300 spellers relied heavily on direct fixation: When participants did not move their eyes, but attended covertly to the letters, accuracy was only around 60%. Thus, four out of ten times the system selected another letter than the user had intended (Brunner et al., 2010; Treder & Blankertz, 2010). This was problematic for real-world applications, because BCIs are mostly useful if they work without any overt (eye) movement; otherwise, movement-based methods, such as eye trackers (e.g., Majaranta, MacKenzie, Aula, & Räihä, 2006) or the famous cheek-movement system used by Stephen Hawking, are much more efficient. However, modern P300 spellers no longer require (eye) movement, and reach impressive accuracy based on covert attention alone (Acqualagna & Blankertz, 2013; Treder, Schmidt, & Blankertz, 2011; reviewed in Riccio, Mattia, Simione, Olivetti, & Cincotti, 2012). For example, a recent system that uses sequentially presented stimuli reached 97.1% selection accuracy with 1.35 characters per minute (Treder et al., 2011) . Expressed as information-transfer rate (ITR), which is a common measure for evaluating BCI performance (Yuan et al., 2013), this corresponds to 6.18 bits/min.

However, P300 spellers have several practical disadvantages. First, they require high-quality EEG-recording equipment, which is expensive. Low-cost EEG systems are becoming available, but, for the moment, are less reliable than more expensive systems (e.g., Duvinage et al., 2013). Second, EEG electrodes must be carefully applied to the head. This is a tedious procedure that must be regularly redone, because current EEG systems are not designed for permanent use, and recording quality degrades over time. Electrodes also cause physical discomfort. Third, most P300 spellers require a calibration phase during which a classifier is trained on a person's EEG signature. Again, this is tedious, and must be redone regularly to avoid performance degradation. These are hurdles for real-world applications (see also Birbaumer, 2006).

Recently, a very different method, based on pupillometry, was developed and tested with partly locked-in patients (Stoll et al., 2013). This method exploits the pupillary dilation (enlargement) that accompanies effortful mental activities, such as arithmetic (reviewed in Laeng, Sirois, & Gredebäck, 2012). Participants were first asked a yes/ no question, and then sequentially shown two response options ('yes' followed by 'no', or vice versa; each option was shown once). A calculation was shown together with each response option (e.g. '29 x 49'). Participants were instructed to perform the calculation only during the interval of the intended response. The selection algorithm was simple: The response that elicited the strongest pupillary dilation (i.e. when the participant was calculating) was selected. Healthy participants reached around 90% selection accuracy with around 3.6 selections per minute (ITR = 1.93 bits/min; ITR is low because yes/ no selections carry little information). Typical locked-in patients reached around 70% accuracy with around 2.7 selections per minute (ITR = 1.42 bits/min). This method is much less efficient than a P300 speller; but it requires only a pupillometer (e.g. a remote camera), and does not require extensive preparation or calibration. These are advantages for real-world applications.

Here we present an entirely new human-computer interface (HCI) that combines the performance of a P300 speller with the usability of pupillometry. Our system builds on the recent discovery that the pupil constricts (shrinks) when you covertly attend to a bright stimulus, compared to a dark stimulus (Binda, Pereverzeva, & Murray, 2013; Mathôt, Dalmaijer, Grainger, & Van der Stigchel, 2014; Mathôt, van der Linden, Grainger, & Vitu, 2013; Naber, Alvarez, & Nakayama, 2013) . That is, unlike traditionally assumed, you do not need to look directly at a bright stimulus to elicit a pupillary light response; a covert shift of attention is sufficient (reviewed in Mathôt & Van der Stigchel, 2015). Our method exploits this by presenting multiple letters within circles that oscillate between brightness and darkness. The participant selects a letter by covertly attending to it, without making any overt (eye) movement. The size of the pupil oscillates along with the brightness of the attended letter. This allows us to determine, reliably and in real time, which stimulus the participant intends to select.

Methods

Preregistration

This experiment was preregistered on Jan 21, 2015 (https://osf.io/yvaqs/). Whenever a deviation from registration occurred, it is indicated in the sections below.

Materials and availability

Participant data, experimental software, and analysis scripts are available from: https://github.com/smathot/mind-writing-pupil. This repository also includes a ready-to-use package for using our HCI with supported systems (currently tested with EyeLink and EyeTribe eye trackers, and Windows and Linux operating systems). A screencast of our method is available on-line: https://youtu.be/cGfkD2opTz4

Participants

Ten naive participants from the community of Aix-Marseille Université were recruited (normal or corrected vision; 7 women; age range: 20-25). Participants received €90 for their participation (deviation from preregistration: We originally planned to pay €60). Participants provided written informed consent prior to the experiment. The study was conducted with approval of the ethics committee of Aix-Marseille Université (Ref.: 2014-12-03-09), and conformed to the Declaration of Helsinki (7th rev.).

Software and apparatus

Eye position and pupil size were recorded monocularly with an EyeLink 1000 (SR Research, Mississauga, ON, Canada), a video-based eye tracker sampling at 1000 Hz. The right eye was recorded, unless the left eye provided a better signal. Stimuli were presented on a 21" ViewSonic p227f CRT monitor (1280 x 1024 px, 85 Hz) running Ubuntu Linux 14.04. Testing took place in a dimly lit room. The experiment was implemented with OpenSesame (Mathôt, Schreij, & Theeuwes,

2012) using the PsychoPy back-end (Peirce, 2007) for display control and PyGaze (Dalmaijer, Mathôt, & Van der Stigchel, 2014) for eye tracking.

General stimuli and procedure

Before each block, a nine-point eye-tracker calibration was performed. At the start of each trial, an automatic single-point recalibration (drift correction) was performed. The display consisted of a green central fixation dot (r = 0.2°) on a gray background (13.0 cd/m^2). Items were presented in a circular configuration at an eccentricity of 9.2° (Figure 1). Items consisted of colored letters against a circular background (r = 3.1°). When only two items were presented, each item was accompanied by a mirror-symmetric placeholder (see Figure 1a; this configuration was chosen because pilot experiments showed it to be the most effective of several tested configurations). The backgrounds alternated between brightness (97.0 cd/m^2) and darkness (5.1 cd/m^2) in cycles of 1.25 s (0.8 Hz). Each cycle consisted of a smooth brightness transition of 0.5 s, followed by 0.75 s of constant brightness (Figure 1c).

The participant attended covertly to the target stimulus, while keeping gaze on the central fixation dot. The target was either indicated by a cue (Phase 1-3) or chosen by the participant (Phase 4). The cue was both visual (e.g., the letter ‘A’ shown on the display) and auditory (e.g., a synthesized French voice saying *Sélectionnez A*). The participant could replay the auditory cue at any moment by pressing the space bar. The trial ended when a selection was made (Figure 1b, see Selection algorithm).

The display was designed to make the selection procedure as intuitive as possible. First, the size of the letters indicated how close they were to being selected; that is, a letter increased in size until it was selected (Figure 1d). This type of sensory feedback is believed to increase BCI/ HCI performance (e.g., Astrand, Wardak, & Ben Hamed, 2014). Second, after a letter had been selected, it smoothly moved towards the display center. This animation increased the participants’ sensation of *grabbing* letters with their mind’s eye.

PeerJ PrePrints | https://dx.doi.org/10.7287/peerj.preprints.1361v1 | CC-BY 4.0 Open Access | rec: 10 Sep 2015, publ: 10 Sep 2015

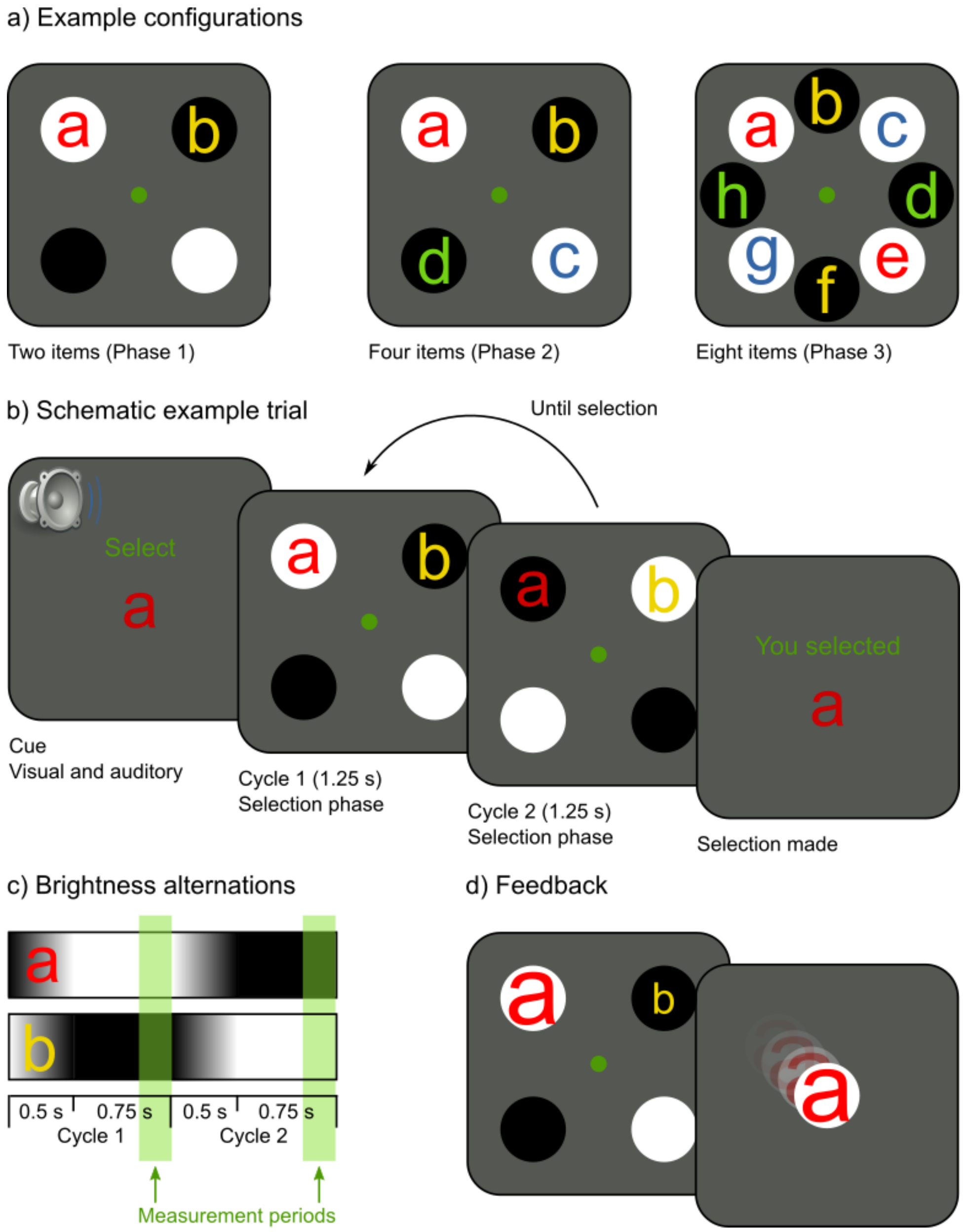

Figure 1. *a) Participants selected one of two (Phase 1), four (Phase 2), or eight (Phase 3) simultaneously presented stimuli. b) The target stimulus was indicated by a cue. This example shows a correct selection, because the selected stimulus ('a') matches the cue. c) During each cycle, the brightness of the stimulus gradually changed in 0.5 s, and then remained constant for 0.75 s. Pupil size was measured during the last 0.25 s. d) The size of the letters indicated how close they were to being selected. When a letter was selected, it smoothly moved toward the center.*

Control for eye position

In each but the final block of each phase, the experiment was paused when fixation was lost (gaze deviated more than 2.6° from the display center for more than 10 ms), and continued when fixation was re-established. In the final block of each phase, the entire display was locked to gaze position (from now on: gaze-stabilization mode): When the eyes drifted slightly to the left, all stimuli except the central fixation dot would shift slightly to the left as well. This made sure that selection was not driven by small eye movements in the direction of the attended stimulus (cf. Mathôt et al., 2014, 2013).

Selection algorithm

Letters are divided into two groups: bright and dark backgrounds. Each group has a parameter *L* that reflects how likely it is that the attended letter is part of that group. Initially, *L* is 1 for both groups. After each cycle, a proportional pupil-size difference (*PPSD*) is determined (see Pupil-size measurement). For the letter group that has changed from bright to dark, *L* is multiplied by *PPSD* (because we expect the pupil to dilate if the target is part that group). For the letter group that has changed from dark to bright, *L* is divided by *PPSD* (because we expect the pupil to constrict if the target is part that group). Cycling continues until the proportional difference between the *L*s for both groups exceeds a threshold *T* ($L1/L2 > T$ or $L1/L2 < 1/T$), after which the group with the highest *L* is designated as the winner. If groups consist of more than one letter, the losing group is discarded, and the winning group is subdivided into two new bright/ dark groups (See Figure 2). The selection process then starts anew. This continues until the winning group contains only a single letter, after which the final selection is made. The analysis is performed on-line, while the participant performs the task.

A crucial property of this algorithm is that it continues until there is sufficient evidence for reliable selection. Selection can be made faster but less accurate by reducing the threshold *T*, and slower but more accurate by increasing it.

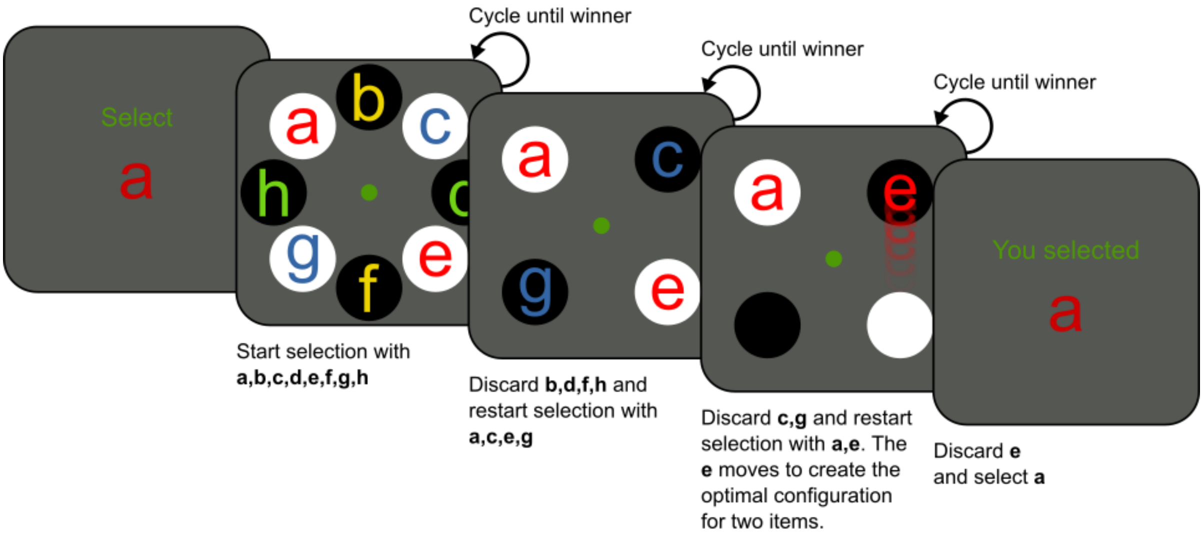

Figure 2. *A schematic example of the selection procedure in the case of eight stimuli. Stimuli are grouped by the brightness of their background. One group is eliminated on each selection, after which the remaining group is subdivided anew. This procedure repeats until a single stimulus remains.*

Pupil-size measurement

The proportional pupil-size difference on cycle *i* (*PPSD(i)*) is defined as:

$$PPSD(i) = \frac{PS(i)}{PS(i-1)}$$

Here, *PS(i)* is the median pupil size during the last 250 ms of cycle *i* (see Figure 1c).

Training program

The training program consisted of four phases. In Phases 1-3, participants were trained to make progressively more complicated selections. In Phase 4, participants wrote a short self-selected sentence using an extension of the technique trained in Phases 1-3. Training took about 10 hours, spread over multiple days.

Phases 1-3: Selecting a predefined stimulus

In Phase 1, participants were trained to select one of two simultaneously presented stimuli. Blocks consisted of 16 selections.

Training was successful when participants reached: 100% accuracy after completing at least 6 bloks; or at least 80% accuracy on block 12. Thus, participants completed between 6 and 12 blocks. If training was unsuccessful, the phase was restarted with a more conservative threshold of 1.5 (default threshold = 1.375). If training then failed again, the experiment was aborted and training was considered unsuccessful for that participant. After training was successfully completed, participants completed a single block in gaze-stabilization mode (see: Control for eye position). Our criteria for success were stringent: Commonly, 70% accuracy is taken as a lower limit for a useful BCI/ HCI (e.g., Astrand et al., 2014; Birbaumer, 2006).

Phases 2 and 3 were identical to Phase 1, except that participants selected one out of four (Phase 2) or eight (Phase 3) stimuli.

Phase 4: Free writing

In Phase 4, participants trained to write text by selecting characters and control symbols (‘backspace’: a leftward arrow; ‘space’: a low bar; and ‘accept’: a square) on a virtual keyboard. The participant initially selected one of eight symbol groups. This group then unfolded, after which the participant selected one symbol. Structurally, selecting a symbol was therefore identical to a one-of-eight selection (Phase 3) followed by a one-of-four selection (Phase 2), or, in the case of ‘accept’ and ‘backspace’, a one-of-two selection (Phase 1). This procedure is similar to the Hex-o-Spell P300-based human-computer interface (Blankertz et al., 2006).

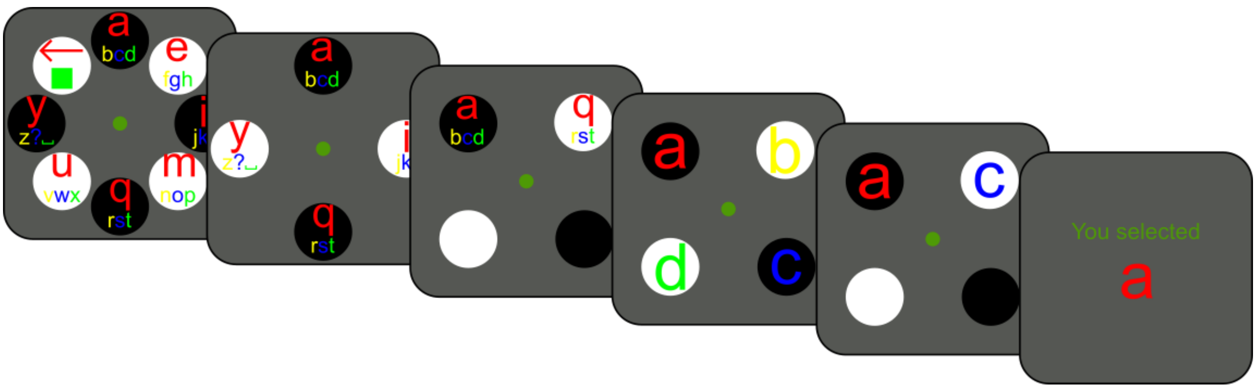

Figure 3. *The symbol-selection procedure used for free writing. Initially, there are eight groups of characters and control symbols (‘backspace’, ‘space’, and ‘accept’). When one group has been selected (here ‘abcd’), it unfolds into four individual symbols (here ‘a’, ‘b’, ‘c’, and ‘d’), after which a final selection is made (here ‘a’).*

First, participants were given a print-out of the virtual keyboard to familiarize themselves with its layout. Next, they practiced by writing the French word “ecrire” (without accent). Practice was completed when the word was written successfully, with a maximum of three attempts. Next, participants chose a short sentence (deviation from preregistration: several participants wanted to write a long sentence, and we therefore abandoned our initial maximum of 15 characters).

Participants were given two opportunities to write this sentence. Writing was considered successful when the final sentence matched the specified sentence. The use of backspace to correct mistakes during text input was allowed.

Results

Phases 1-3: Selecting a predefined stimulus

Criteria and statistical analyses

No participants or selections were excluded from the analysis. Two blocks (32 selections) were lost due to technical problems. Two participants chose not to finish the experiment, and were replaced. In total, 257 blocks (4,112 selections) were included in the analysis.

We analyzed accuracy using Generalized Linear Mixed-Effects Models (GLMER) with correctness (binomial) as dependent variable. We analyzed response times using Linear Mixed-Effects Models with response time as dependent variable. We included by-participant random intercepts and slopes (i.e. maximal random effects), unless this model failed to converge, in which case we included only random intercepts. Fixed effects were considered reliable when $p < .05$; however, we emphasize general patterns over significance of individual results. These analyses were conducted in R (R Core Team, 2014), using the packages `lme4` (Bates, Maechler, Bolker, & Walker, 2015) and `lmerTest` (Kuznetsova, Brockhoff, & Christensen, 2015).

Information-transfer rate (ITR) is a measure of communication efficiency, and depends on both accuracy and speed. ITR was determined using the following formula (Yuan et al., 2013):

$$ITR = \frac{\log_2 N + Acc \log_2 Acc + (1 - Acc) \log_2 \frac{1 - Acc}{N - 1}}{RT / 60}$$

Here, *ITR* is in bits per minute, *N* is the number of response options, *Acc* is proportion correct responses, and *RT* is the response time in seconds.

The response time was the interval between the start of the first selection cycle and the end of the last selection cycle. Mean accuracy, response time, and ITR were first determined per participant, and then averaged to arrive at grand means (i.e. a means-of-means approach).

Pupillary responses

Figure 4a shows how pupil size evolved during one cycle (1.25 s) as a function of whether the attended stimulus changed from bright to dark (blue line) or dark to bright (orange line). This is data from one participant. Each cycle started with a 500 ms transition period, during which the brightness of the stimuli smoothly changed. During transition, pupil size still reflected the pretransition brightness: the pupil was larger if the attended stimulus was dark (orange line) than if it was bright (blue). Next, there was an adaptation period of 500 ms. During adaptation, the pupil gradually started to reflect the new brightness of the attended stimulus, as reflected by the crossover of the blue and orange lines. Finally, there was a measurement period of 250 ms, during which the brightness effect (i.e. the difference between the orange and blue lines) was roughly stable. Median pupil size during this period was used for the analysis; that is, our method exploited the fact that pupil size was larger when a target was dark (blue line) than when it was bright (orange line; see also Methods: Pupil-size measurement).

As shown in in Figure 4b, all participants showed a qualitatively identical pattern. One participant (indicated in red) showed a weak effect; this was the only participant who did not reach our criteria for successful training (see Selection accuracy and speed).

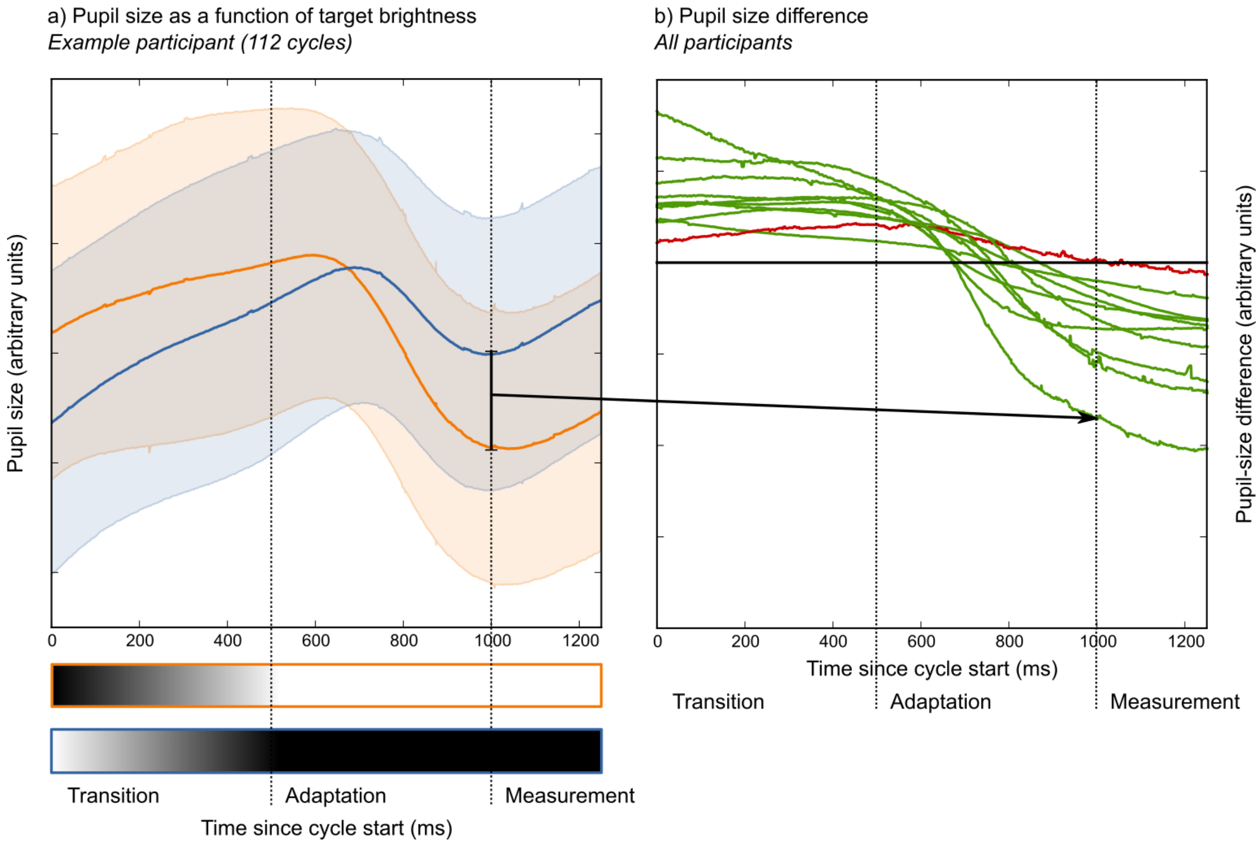

Figure 4. *a) Example data from one participant. Pupil size as a function of whether the target changes from bright to dark (blue line) or from dark to bright (orange line). Shadings indicate standard deviation. b) The pupil size difference (i.e. orange - blue) for all participants. The participant indicated in red did not reach our criteria for successful training. The participant indicated by the arrow corresponds to the example shown in (a). All data is from Phase 1, in which participants selected one out of two stimuli.*

Selection accuracy and speed

The main results are shown in Figure 5, which shows the mean selection accuracy and speed for each participant.

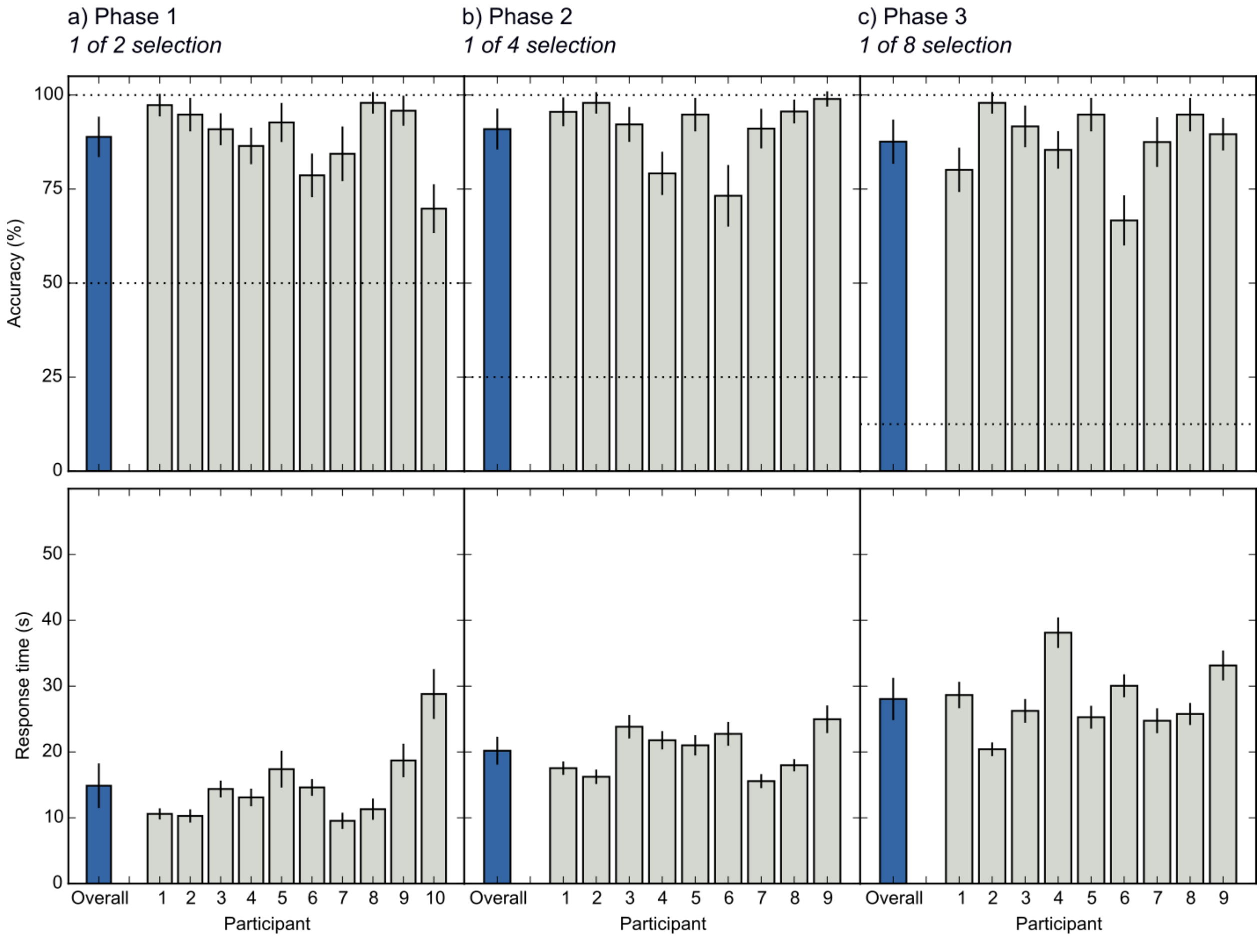

Figure 5. *Selection accuracy (top row) and speed (bottom row) for individual participants (gray bars) and across participants (blue bars). Horizontal dashed lines indicate chance level. a) Results for Phase 1. b) Results for Phase 2. c) Results for Phase 3. Error bars indicate 95% confidence intervals (within-subject where applicable, cf. Cousineau, 2005).*

In Phase 1, the mean accuracy was 88.9% (chance = 50%; *N* = 10), with a mean selection time of 14.9 s. ITR was 2.58 bits/min (Figure 6). Nine out of ten participants met our criteria for successful training (see Methods: Training program). One participant did not meet our criteria for success, and therefore did not participate in subsequent phases (#10 in Figure 5; red line in Figure 4b); however, this participant's accuracy was still 70%, which is often taken as the lower limit for useful HCI performance (Astrand et al., 2014; Birbaumer, 2006). In Phase 2, the mean accuracy was 91.0% (chance = 25%; *N* = 9), with a mean selection time of 20.2 s. The ITR was 4.55 bits/min. All participants met our criteria for successful training. In Phase 3, the mean accuracy was 87.6% (chance = 12.5%; *N* = 9), with a mean selection time of 28.0 s. The ITR was 4.86 bits/min. Again, all participants met our criteria for successful training.

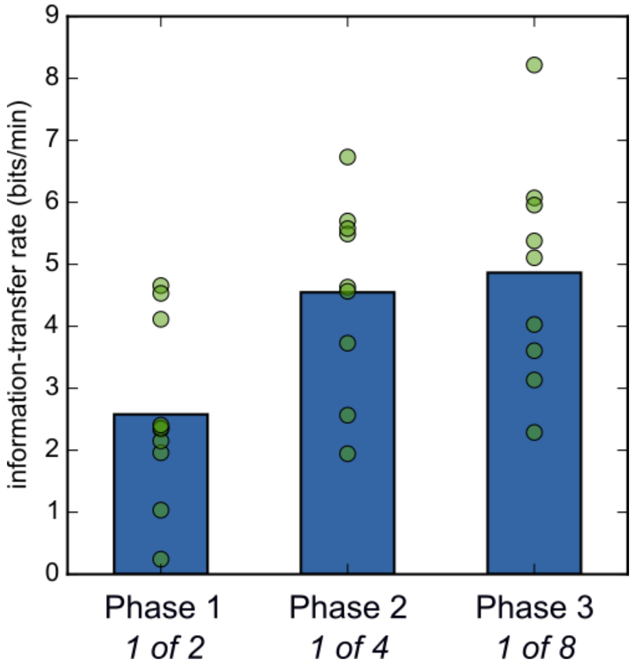

Figure 6. *The information-transfer rate (ITR) in bits per minute. Bars indicate the mean ITR. Dots indicate individual participants.*

Learning

Figure 7 shows how selection accuracy and speed evolve over time. To test whether significant learning occurred, we conducted a GLMER on accuracy with block (continuous) as fixed effect. This was done for each phase separately. There was no notable effect of block (i.e. no learning effect) in any phase: Phase 1 (z = 1.62, p = .104), Phase 2 (z = 1.30, p = .195), Phase 3 (z = 1.48, p = .139). An LMER on response time also did not reveal any notable effect of block: Phase 1 (t = 0.57, p = .565; intercept-only model), Phase 2 (t = 0.37, p = .721), Phase 3 (t = 0.73, p = .488).

Looking at Figure 7a, some learning did appear to occur between blocks 1 and 2 of Phase 1; that is, participants needed a single block of training, before they reached a more-or-less stable level of performance. Importantly, learning effects, if any, were small, and participants were able to use our method right away.

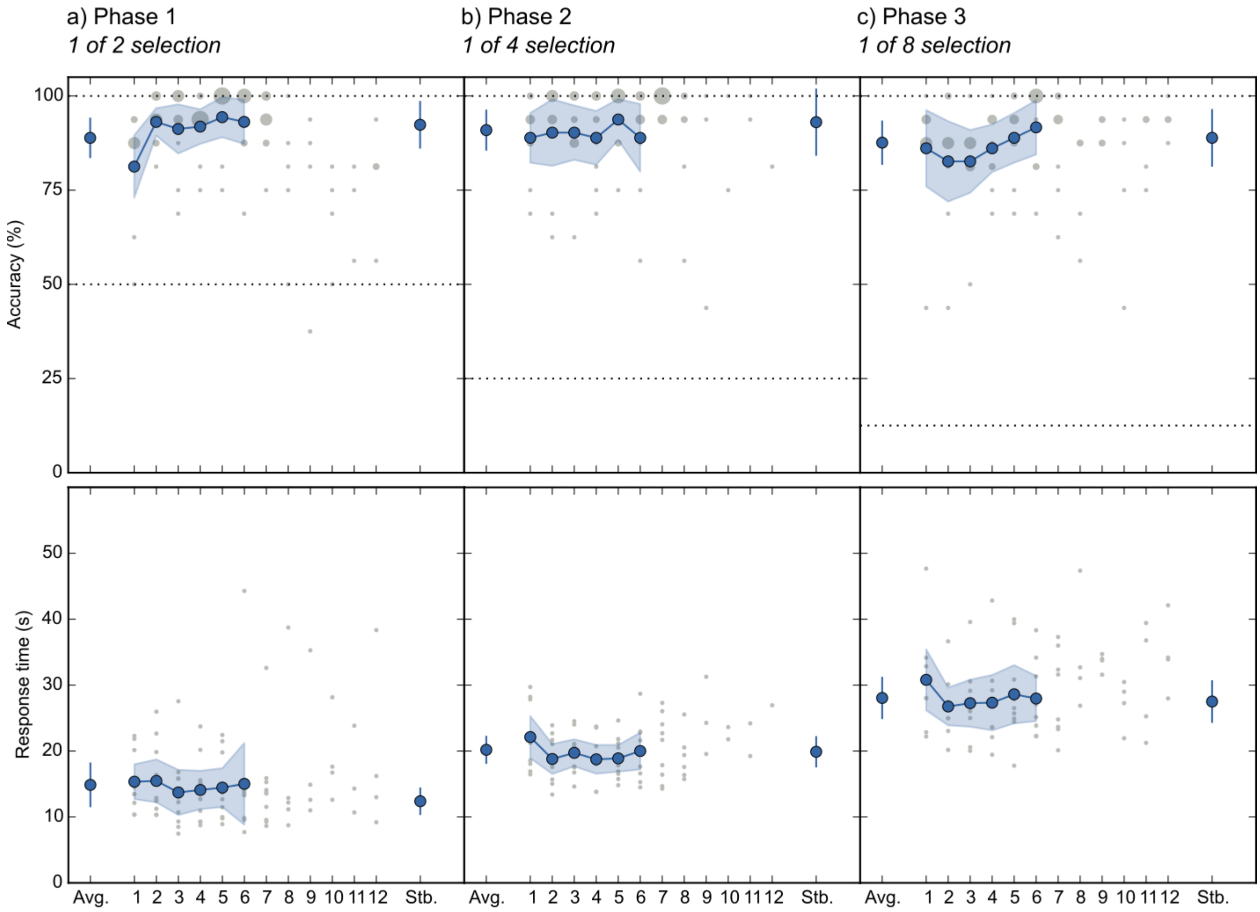

Figure 7. *Selection accuracy (top row) and speed (bottom row) as a function of block number. Blue lines indicate across-participant means during the first six blocks, which were completed by all participants. The size of the gray circles indicates how often a score occurred. Performance during gaze-stabilization blocks is indicated by Stb. Horizontal dotted lines indicate chance level. a) Results for Phase 1. b) Results for Phase 2. c) Results for Phase 3. Error bars indicate 95% within-subject confidence intervals (cf. Cousineau, 2005).*

Gaze independence

A crucial question is whether selection is fully independent of eye position. To test this, we conducted a GLMER on accuracy with gaze stabilization (on/ off) as fixed effect. This revealed no notable effect of gaze stabilization (z = 1.64, p = .102). An LMER on response times also revealed no effect (t = 1.39, p = .174). If anything, performance was slightly better when gaze-stabilization mode was enabled (see also Figure 7 in which gaze-stabilization blocks are marked as 'Stb.').

Crucially, this shows that performance did not depend on small eye movements toward the attended stimuli (cf. Engbert & Kliegl, 2003), which participants could have made when gaze-stabilization was disabled. Our method is fully driven by covert attention.

Phase 4: Free writing

Eight out of nine participants successfully wrote a self-selected sentence (Table 1). The remaining participant made a sentence that was correct except for one typo.

Participants could use a 'backspace' symbol to correct mistakes, and needed to enter an 'accept' symbol to end text input. Therefore, we can distinguish the symbols that were entered (including characters that were later deleted, etc.) from the useful text (the text string that was eventually accepted). In total, participants entered 190 symbols (letters, '?', 'space', 'backspace', and 'accept') for 133 characters of useful text (letters, '?', and 'space'). On average, one symbol took 51.1 s (*SD* = 9.6; including 'backspace' and 'accept'), and one character of functional text took 75.2 s (*SD* = 20.5). The functional ITR was 3.91 bits/min. (A bug in an early version of the software occasionally required participants to enter unnecessary 'backspace' symbols. One sentence was affected by this issue, and was excluded from the analysis above.)

Table 1

Results of Phase 4, during which participants wrote a self-selected sentence. Names have been replaced by asterisks (*).

Response	Translation	Correct
LE CHAT DORT	The cat sleeps	Yes
JE NE SUIS PAS SI RAPIDE QUE CA	I'm not as fast as that	Yes
JE M APELLE ******	My name is ******	Yes
ENFIN TERMINEE	Finally finished	Yes
EXPERIENCE TERMINEE	Experiment finished	Yes
JE VAIS AGRANDIR	I'm going to get bigger	Yes
LE CHIEN BOIT	The dog drinks	Yes
VIVE LE POIIL?	Long live the fur?	Should have been "VIVE LE POIL?"
JE SUIS ****	I am ****	Yes

Discussion

We have introduced a new human-computer interface (HCI) that is based on decoding of covert attention through pupillometry. Participants select a letter by covertly attending to it, without making any overt (eye) movement. Letters are presented within circles of oscillating brightness. Small changes in pupil size reflect the brightness changes of the attended stimulus (reviewed in Mathôt & Van der Stigchel, 2015), and this allows us to determine which stimulus the participant intends to select.

In the experiment reported here, with healthy untrained participants, our method reached a selection accuracy of around 90%, and an information-transfer rate (ITR) of 4.86 bits/min (Figure 6). Out of ten participants, all but one reached our predetermined criteria for successful training; these criteria were exceptionally stringent, and even the unsuccessful participant achieved 70% selection accuracy, which is often taken as sufficient for a useful BCI/ HCI (e.g., Astrand et al., 2014; Birbaumer, 2006). During pilot studies with highly trained participants (authors SM and LvdL), we have systematically reached near-perfect selection and ITRs of over 8 bits/min. (Pilot data is available online; see Methods - Materials and availability.) For comparison, P300 spellers that are based on covert attention (i.e. without eye movements) reach an ITR of around 6 bits/min (based on visual spellers reviewed in Riccio et al., 2012), usually with a combination of trained and untrained participants (e.g., Acqualagna & Blankertz, 2013; Treder et al., 2011). The performance of our method is thus in the same range as that of the best non-invasive covert-attention BCIs to date.

An important advantage of our method, especially when compared to EEG-based methods, is its ease of use. Only a pupillometer, or an eye tracker that records pupil size, is required. For most experiments, we have used a research-grade eye tracker; but we have also successfully used an EyeTribe (The Eye Tribe Aps, Copenhagen, Denmark), a low-cost eye tracker that provides high-quality pupil-size measurements (Dalmaijer, 2014). Our method does not require eye-tracker calibration, nor training of the selection algorithm. Together, these characteristics set our method apart from currently available methods.

An important application of a HCI/ BCI is as a communication channel for completely locked-in patients, that is, patients with complete loss of motor control (Birbaumer, 2006). P300 spellers and pupillometry-based methods have been tested successfully in partly locked-in patients with some remaining motor control (e.g., Marchetti, Piccione, Silvoni, Gamberini, & Priftis, 2013; Stoll et al., 2013). But success with real-world applications has been modest, especially with completely locked-in patients. Important reasons for this limited success are (cf. Birbaumer, 2006): difficulty of use (some methods require extensive training); low selection accuracy; skin problems due to EEG electrodes; low selection speeds; and the need for sustained attention. Our method solves some of these problems by providing ease of use, avoiding physical contact, and providing high selection accuracy. But other challenges remain, notably low selection speed and the need for sustained attention. In addition, it is unclear to what extent the pupillary light response, which our method relies on, remains intact in completely locked-in state (Bauer, Gerstenbrand, & Rumpl, 1979). Therefore, future studies are needed to determine how well our method, or a variation thereof, works in patient groups.

A second application of our method is as an ultra-secure way to enter passwords or PIN codes. Imagine a cash machine that is equipped with a pupillometer. To enter a PIN code, the user would be shown a display similar to that depicted in Figure 1a, and enter digits by covertly attending to them. Based on our results (see Figure 5), entering a four-digit PIN code would take around two minutes. This is slow, but feasible, and could be useful in situations that require high security.

A third application of our method is as a way to train sustained attention. To select a letter, participants must attend to it for some time, which is effortful. Therefore, a game-like variation of our method could be an attractive way to train sustained attention. The main benefit of our method over regular attention-training exercises is direct feedback: The user can be immediately notified when there is a lapse of attention. (In our experiments, feedback was provided by changing the size of the target letter.)

In conclusion, we have presented a new pupillometry-based method to translate thought into letters. Our method is highly accurate and easy to use, and does not require elaborate equipment, preparation, or training. We have highlighted communication with completely locked-in patients, ultra-secure password input, and training of sustained attention as possible applications.

PeerJ PrePrints | https://dx.doi.org/10.7287/peerj.preprints.1361v1 | CC-BY 4.0 Open Access | rec: 10 Sep 2015, publ: 10 Sep 2015

Acknowledgements

The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement n° 622738 awarded to Sebastiaan Mathôt. In addition, this research was funded by a VIDI Grant 452-13-008 from the Netherlands Organization for Scientific Research to Stefan Van der Stigchel.

References

Acqualagna, L., & Blankertz, B. (2013). Gaze-independent BCI-spelling using rapid serial visual presentation (RSVP). *Clinical Neurophysiology*, *124*(5), 901–908. doi:10.1016/j.clinph.2012.12.050

Astrand, E., Wardak, C., & Ben Hamed, S. (2014). Selective visual attention to drive cognitive brain–machine interfaces: from concepts to neurofeedback and rehabilitation applications. *Frontiers in Systems Neuroscience*, *8*. doi:10.3389/fnsys.2014.00144

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). *lme4: Linear mixed-effects models using Eigen and S4*. Retrieved from http://CRAN.R-project.org/package=lme4

Bauer, G., Gerstenbrand, F., & Rumpl, E. (1979). Varieties of the locked-in syndrome. *Journal of Neurology*, *221*(2), 77–91. doi:10.1007/bf00313105

Binda, P., Pereverzeva, M., & Murray, S. O. (2013). Attention to bright surfaces enhances the pupillary light reflex. *Journal of Neuroscience*, *33*(5), 2199–2204. doi:10.1523/jneurosci.3440-12.2013

Birbaumer, N. (2006). Breaking the silence: Brain–computer interfaces (BCI) for communication and motor control. *Psychophysiology*, *43*(6), 517–532. doi:10.1111/j.1469-8986.2006.00456.x

Blankertz, B., Dornhege, G., Krauledat, M., Schröder, M., Williamson, J., Murray-Smith, R., & Müller, K. (2006). The Berlin Brain-Computer Interface presents the novel mental typewriter Hex-o-Spell. Retrieved from http://eprints.maynoothuniversity.ie/1786/

Brunner, P., Joshi, S., Briskin, S., Wolpaw, J. R., Bischof, H., & Schalk, G. (2010). Does the “P300” speller depend on eye gaze? *Journal of Neural Engineering*, *7*(5), 056013. doi:10.1088/1741-2560/7/5/056013

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson’s method. *Tutorial in Quantitative Methods for Psychology*, *1*(1), 4–45.

Dalmaijer, E. (2014). Is the low-cost EyeTribe eye tracker any good for research? *PeerJ PrePrints*, *2*(e585). doi:10.7287/peerj.preprints.585v1

Dalmaijer, E., Mathôt, S., & Van der Stigchel, S. (2014). PyGaze: An open-source, cross-platform toolbox for minimal-effort programming of eyetracking experiments. *Behavior Research Methods*, *46*(4), 913–921. doi:10.3758/s13428-013-0422-2

Donoghue, J. P. (2008). Bridging the Brain to the World: A Perspective on Neural Interface Systems. *Neuron*, *60*(3), 511–521. doi:10.1016/j.neuron.2008.10.037

Duvinage, M., Castermans, T., Petieau, M., Hoellinger, T., Cheron, G., & Dutoit, T. (2013). Performance of the Emotiv Epoc headset for P300-based applications. *BioMedical Engineering OnLine*, *12*(1), 56. doi:10.1186/1475-925x-12-56

Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research*, *43*(9), 1035–1045. doi:10.1016/s0042-6989(03)00084-1

Farwell, L. A., & Donchin, E. (1988). Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and clinical Neurophysiology*, *70*(6), 510–523.

Fazel-Rezai, R., Allison, B. Z., Guger, C., Sellers, E. W., Kleih, S. C., & KĂźbler, A. (2012). P300 brain computer interface: current challenges and emerging trends. *Frontiers in Neuroengineering*, *5*. doi:10.3389/fneng.2012.00014

Krusienski, D. J., Sellers, E. W., Cabestaing, F., Bayoudh, S., McFarland, D. J., Vaughan, T. M., & Wolpaw, J. R. (2006). A comparison of classification techniques for the P300 Speller. *Journal of Neural Engineering*, *3*(4), 299. doi:10.1088/1741-2560/3/4/007

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). *lmerTest: Tests in Linear Mixed Effects Models*. Retrieved from https://cran.r-project.org/web/packages/lmerTest/index.html

Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science*, *7*(1), 18–27. doi:10.1177/1745691611427305

Majaranta, P., MacKenzie, I. S., Aula, A., & Räihä, K. (2006). Effects of feedback and dwell time on eye typing speed and accuracy. *Universal Access in the Information Society*, *5*(2), 199–208.

Marchetti, M., Piccione, F., Silvoni, S., Gamberini, L., & Priftis, K. (2013). Covert Visuospatial Attention Orienting in a Brain-Computer Interface for Amyotrophic Lateral Sclerosis Patients. *Neurorehabilitation and Neural Repair*, *27*(5), 430–438. doi:10.1177/1545968312471903

Mathôt, S., & Van der Stigchel, S. (2015). New light on the mind's eye: The pupillary light response as active vision. *Current Directions in Psychological Science*.

Mathôt, S., Dalmaijer, E., Grainger, J., & Van der Stigchel, S. (2014). The pupillary light response reflects exogenous attention and inhibition of return. *Journal of Vision*, *14*(14), 7.

doi:10.1167/14.14.7

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324. doi:10.3758/s13428-011-0168-7

Mathôt, S., van der Linden, L., Grainger, J., & Vitu, F. (2013). The pupillary response to light reflects the focus of covert visual attention. *PLoS ONE*, *8*(10), e78168. doi:10.1371/journal.pone.0078168

Naber, M., Alvarez, G. A., & Nakayama, K. (2013). Tracking the allocation of attention using human pupillary oscillations. *Frontiers in Psychology*, *4*(919). doi:10.3389/fpsyg.2013.00919

Nicolas-Alonso, L. F., & Gomez-Gil, J. (2012). Brain computer interfaces, a review. *Sensors*, *12*(2), 1211–1279.

Peirce, J. W. (2007). PsychoPy: Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1-2), 8–13. doi:10.1016/j.jneumeth.2006.11.017

R Core Team. (2014). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Riccio, A., Mattia, D., Simione, L., Olivetti, M., & Cincotti, F. (2012). Eye-gaze independent EEG-based brain–computer interfaces for communication. *Journal of Neural Engineering*, *9*(4), 045001. doi:10.1088/1741-2560/9/4/045001

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*(15), R647–R648. doi:10.1016/j.cub.2013.06.011

Treder, M. S., & Blankertz, B. (2010). Research (C) overt attention and visual speller design in an ERP-based brain-computer interface. *Behavioral & Brain Functions*, *6*.

Treder, M. S., Schmidt, N. M., & Blankertz, B. (2011). Gaze-independent brain–computer interfaces based on covert attention and feature attention. *Journal of Neural Engineering*, *8*(6), 066003. doi:10.1088/1741-2560/8/6/066003

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical Neurophysiology*, *113*(6), 767–791. doi:10.1016/s1388-2457(02)00057-3

Yuan, P., Gao, X., Allison, B., Wang, Y., Bin, G., & Gao, S. (2013). A study of the existing problems of estimating the information transfer rate in online brain–computer interfaces. *Journal of neural engineering*, *10*(2), 026014.