

# Feasibility of a Device for Gaze Interaction by Visually-Evoked Brain Signals

Baosheng James Hou  
Technical University of Denmark  
Lyngby, Denmark

I. Scott MacKenzie  
York University  
Toronto, Canada

Cihan Uyanik  
Technical University of Denmark  
Lyngby, Denmark

Jacopo M. Araujo  
Technical University of Denmark  
Lyngby, Denmark

John Paulin Hansen  
Technical University of Denmark  
Lyngby, Denmark

Per Bækgaard  
Technical University of Denmark  
Lyngby, Denmark

Sadasivan Puthusserypady  
Technical University of Denmark  
Lyngby, Denmark

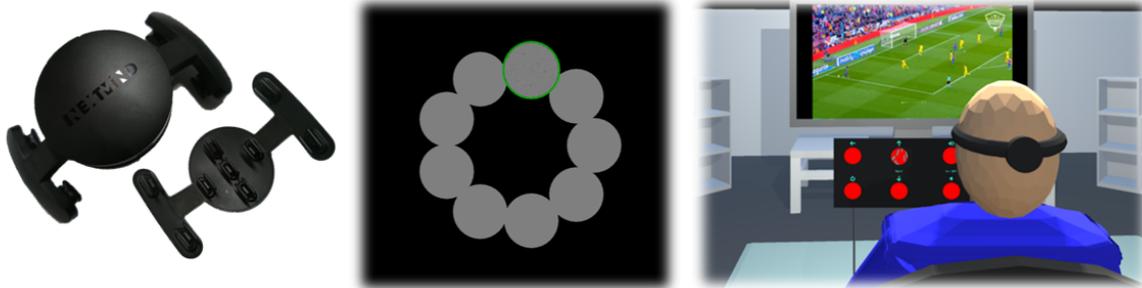


Figure 1: Left, the NextMind sensor unit (front and back view). Middle, the Fitts' law task interface. Right, a simulated use-case with wheelchair steering and smart home-control of a TV.

## ABSTRACT

A dry-electrode head-mounted sensor for visually-evoked electroencephalogram (EEG) signals has been introduced to the gamer market, and provides wireless, low-cost tracking of a user's gaze fixation on target areas in real-time. Unlike traditional EEG sensors, this new device is easy to set up for non-professionals. We conducted a Fitts' law study ( $N = 6$ ) and found the mean throughput ( $TP$ ) to be 0.82 bits/s. The sensor yielded robust performance with error rates below 1%. The overall median activation time ( $AT$ ) was 2.35 s with a minuscule difference between one or nine concurrent targets. We discuss whether the method might supplement camera-based gaze interaction, for example, in gaze typing or wheelchair control, and note some limitations, such as a slow  $AT$ , the difficulty of calibration with thick hair, and the limit of 10 concurrent targets.

## CCS CONCEPTS

• Human-centered computing → HCI design and evaluation methods.

## KEYWORDS

Fitts' law, brain-computer interface (BCI), SSVEP, gaze interaction, Hybrid BCI, accessibility, augmentative and alternative communication,

## ACM Reference Format:

Baosheng James Hou, Cihan Uyanik, Per Bækgaard, I. Scott MacKenzie, Jacopo M. Araujo, Sadasivan Puthusserypady, and John Paulin Hansen. 2022. Feasibility of a Device for Gaze Interaction by Visually-Evoked Brain Signals. In *2022 Symposium on Eye Tracking Research and Applications (ETRA '22)*, June 8–11, 2022, Seattle, WA, USA. ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/3517031.3529232>

## 1 INTRODUCTION

Hitherto, most gaze interactive applications have been driven by camera-based eye tracking, including those found in communication systems for people with motor disabilities. Like camera-based eye tracking, brain-computer interface (BCI) may provide information on where a user is looking, and thereby enable gaze interaction. Thus, BCI could serve as a viable alternative to common video-oculography (VOG) and infrared (IR) light technologies. VOG tracking has a number of issues when applied for gaze interaction:

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

ETRA '22, June 8–11, 2022, Seattle, WA, USA

© 2022 Association for Computing Machinery.

ACM ISBN 978-1-4503-9252-5/22/06...\$15.00

<https://doi.org/10.1145/3517031.3529232>

It is sensitive to ambient light, especially sunlight, to eye makeup, to vibrations of the camera, to gross movements of the user, and it has a limited field of operation – the head-box – wherein the eyes can be tracked accurately. Some people report that the IR light makes their eyes feel “dry” after prolonged use. A number of these practical limitations may be overcome with BCI systems, for instance, the sensitivity to reflections from ambient light or dry eyes. BCI has been applied in several studies, for example, augmented and alternative communication (AAC) [Speier et al. 2013] and control of wheelchairs or robots [Leeb et al. 2013]. The main drawbacks of using BCI outside the laboratory are that the user is required to wear an electroencephalogram (EEG) cap mounted with electrodes and wires. The electrodes can lose their contacts with the scalp if the conductive gel dries out, or if the user’s head moves. Also, some BCI methods require lengthy calibration. Most of these systems are expensive and the set-up and calibration process is too complicated for the average user or caregiver. So there is a technology gap for BCI systems for use by non-experts outside the lab. Thus, it is relevant to study commercial devices for the gaming market.

Visually-evoked potentials can be measured from the visual cortical areas of the brain. This is commonly done by the steady-state visual evoked potentials (SSVEP) based method. Different frequency potentials are evoked by looking at an interface with objects (i.e., buttons) flashing at different frequencies. Consequently, SSVEP may provide the information needed for a user to actively control an interface by gaze. A general user-requirement will be that the activation time is acceptable (i.e., compatible with other VOG gaze tracking systems), that several selection options can be present simultaneously, and that the required target sizes are suitable for standard monitors. Also, the success rate should be high enough for unambiguous and robust selections, even when several targets are present at the same time.

Our paper presents a number of initial tests of a new BCI sensor system, namely the “NextMind” that provides gaze interaction. It is mounted on the visual cortex area of users with a head-band and transmits wireless EEG data to a PC or laptop. The device was marketed in 2020 at \$300 for computer games and virtual reality (VR) applications, but is now discontinued<sup>1</sup>. Unfortunately, the NextMind device does not provide the access to the acquired data streams that research instruments do; so, we are unable to provide a full account of its operation or potential. Slow-motion video observations of the flashing targets suggest that the NextMind device determines which target the user is looking at by only having one target flashing at a time, with a frequency of approximately 3 Hz for time-interlocked flashes.

We consider NextMind a first example of what may become a new generation of user-friendly gaze tracking devices based on brain signals. Thus, the main contribution of this paper is not to provide specifications for this particular product but to use this first example to reflect on human-system performance, accessibility, and feasibility of BCI devices for gaze communication. On a general level, one might argue that the traditional eye-position data from image analysis, electrooculogram (EOG) or other means are subordinate, epiphenomenal signals of perception, compared to data on

which of a number of targets that actually gets attended to in the visual cortex.

This paper is structured as follows: A Fitts’ Law method (see section 3) compares the activation time ( $AT$ ), hit-rate, and throughput ( $TP$ ) between task displays with three target widths ( $W$ ) and three movement amplitudes ( $A$ ) between the targets. In Experiment 1, we only show one target at a time. In Experiment 2, we present nine targets simultaneously. A nine-target condition is representative for a normal user interface with several selection options. BCI methods may allow for estimations of objects attended, even when the user’s head is positioned close to the screen. We therefore tested the  $W \times A$  conditions twice: With the user’s head 50 cm from the monitor, which is common for VOG experiments, and also 20 cm from the monitor, which is outside the recommended head-box for most remote-tracking VOG systems. Viewing distances are of interest since we consider a smart-home use-case for people with motor challenges, where the control monitor may be placed either at a wheelchair or located at the electronic device, potentially distributed on both. This use-case is addressed in section 5. Finally, in section 6, we discuss the pros and cons of a BCI system relative to VOG systems.

## 2 BRAIN-COMPUTER INTERFACE

A BCI system enables the control of external devices using brain activity as an input variable. The potential and development of hybrid gaze-BCI systems for multimodal human-computer interaction (HCI), where one can “point with your eye and click with your mind” has been discussed for several decades [Velichkovsky and Hansen 1996]. Zander et al. [2010] investigated gaze with motor-imagery (MI) BCI in a search-and-selection task, and found that although slower, the hybrid method could be more accurate than short dwell-time, especially in difficult tasks. Yong et al. [2011] implemented a gaze-BCI system for text-entry, where gaze was the pointing method and an attempted hand extension was the selection method. Dong et al. [2015] implemented a gaze-MI BCI to control a cursor, MI BCI was used to steer the cursor, integrated with gaze information to aid classification. They found an improved accuracy and task completion speed compared to BCI alone. A number of hybrid systems that combines VOG eye tracking and BCI has been developed. Petrushin et al. [2018] describe steering a tele-robot by eye gaze, but only when engaged by a MI BCI command. This allows the user to explore the scene from the tele-robot camera without moving it. MI BCI commands have also been combined with eye tracking for communication [Vilimek and Zander 2009], robot arm control [Frisoli et al. 2012], and 3D selection [Lee et al. 2010]. Hou et al. [2020] evaluated a hybrid gaze and MI BCI system in a Fitts’ law experiment. Eye tracking was used for pointing and different MIs were studied for selection. The best performance was found for the condition where users were given the option to choose which MI to use, with a mean throughput of 0.59 bits/s, the fastest selection time of 2650 ms, and an error rate of 14.6%.

NextMind and other SSVEP systems are not hybrid gaze-BCI like those just mentioned, but “pure” gaze-interaction-by-BCI systems, because they do not explicitly require or track fixation, bypassing the need for an eye tracker. A target is activated when the device detects the SSVEP frequency in the EEG matching the frequency

<sup>1</sup>The NextMind company has been acquired by Snap Lab and thus likely to be embedded in their future XR glasses <https://techcrunch.com/2022/03/23/snap-buys-mind-controlled-headband-maker-nextmind/?guccounter=1>

assigned to the target. Therefore the question is not “where the participant looked” but “whether the SSVEP signal is detected”. The activation “click” is not assigned to the location of the fixation, but the entire target. This is attractive for some in-the-wild applications, where performance of a VOG system may be limited by noise from changing ambient light, movement, glasses etc. SSVEP systems requires no training, which is particularly important when introduced to people with neuro-degenerative diseases, who easily become mentally fatigued.

### 3 EXPERIMENT

Fitts’ law experiments are widely used to test and model interactions that involve pointing rapidly to a target and selecting it. It has been used over the decades to evaluate a variety of computer input devices including the mouse, joystick, and even gaze and BCIs [Cuaresma and MacKenzie 2017; Felzer et al. 2016; Hassan et al. 2019; Hou et al. 2020; MacKenzie and Teather 2012; Roig-Maimó et al. 2017; Soukoreff and MacKenzie 2004]. In a Fitts’ law experiment, we investigate how the amplitude between targets ( $A$ ) and the width of the targets ( $W$ ) affect performance metrics, such as  $AT$  and throughput ( $TP$ ). Recently, it has been argued that individual saccades are pre-programmed, ballistic movements [Schuetz et al. 2019] and for this reason Fitts’ Law is not applicable. However, our use of Fitts’ law is to measure  $TP$  over a series of whole trials, wherein each trial includes multiple saccades and moments of fixation. Fitts’  $TP$  has been used reliably in past research in this context (e.g., [Zhang and MacKenzie 2007]).

#### 3.1 Participants

Six participants (mean age 26.2 years; 2 females and 4 males) were recruited from the authors personal network. One of the participants was wearing glasses. Three had never tried BCI before, while two had tried it several times and one participant had tried it once. None reported having had any epileptic issues. No compensation was given to the participants. We acknowledge that the number of participants is low, but the COVID-19 lock-down made it impossible to recruit more.

#### 3.2 Apparatus

The experiments were conducted at two different physical locations with a full set-up at both places (again, due to COVID-19 restrictions). The computers at both locations were Windows laptops with an extra monitor attached. At one location, this was a 24-in Samsung model S24C350H (W 53.2 × H 30 cm). The other set-up used a 23.5-in Samsung C24F396FHU (W 52.1 × H 29.3 cm). The refresh rate was 60 Hz for both monitors. Two identical NextMind sensors were used. The wearable sensor module has a weight of 60 gm and has nine comb-shaped dry electrodes to pick up the EEG signals and send them to the computer (Fig. 1, left). The laptop computers run the NextMind real-time algorithms that transform the EEG signals into commands. The commands were sent to a Unity application that conducts the calibration and from where our Unity version of the GoFitts software<sup>2</sup> was launched, upon a successful calibration.

<sup>2</sup>The original version may be found at <http://www.yorku.ca/mack/FittsLawSoftware/> and the Unity version at <https://github.com/GazeIT-DTU/FittsLawUnity>

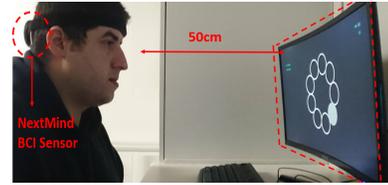


Figure 2: Experiment 1 with 50-cm subject distance

Two versions of the standard circular target layout was made, each with nine target circles. The NextMind Unity kit has a limit of 10 active buttons (“NeuroTags”) on each scene.<sup>3</sup> In the layout for Experiment 1, only one flickering (selectable) target was visible at a time. In the layout for Experiment 2, nine flickering targets were shown. With this interface, the next target was marked with a green circle.

After each successful selection of a target, the system emitted a click-sound and the next target appeared across the layout circle in a clockwise pattern, as common for a 2D Fitts’ law experiment [MacKenzie 2012].

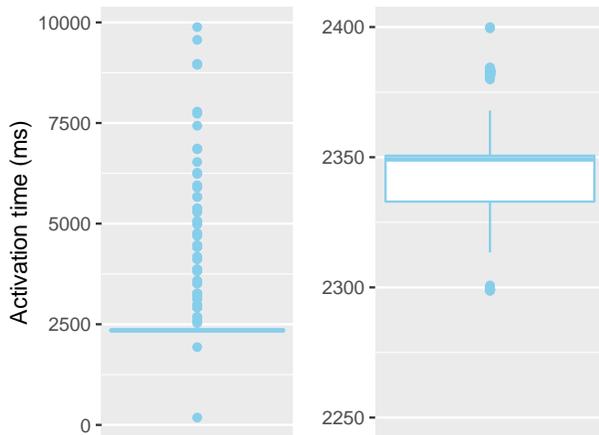
Both Experiment 1 and 2 were performed under two different viewing conditions, mixed within the target widths and amplitudes (see section 3.3). The 50 cm viewing condition of Experiment 1 is shown in Fig. 2. Similar viewing conditions were applied in Experiment 2 (not shown) as well.

#### 3.3 Procedure

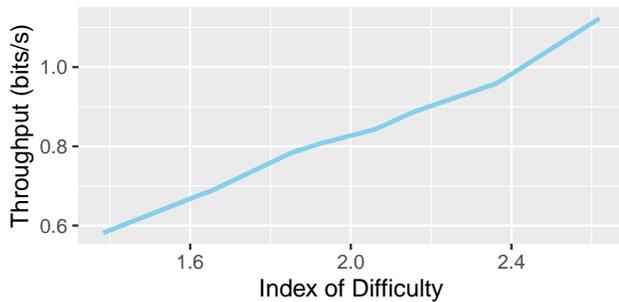
Each participant completed two experiments (i.e. they did both experiment 1 and 2) with the two different stimuli presentations (i.e. with one flickering target and with nine flickering targets). Each experiment contained 18 trials in random order made up of all combinations of the two eye-screen distances (20 and 50 cm), three target amplitudes (11.0, 14.5 and 18.0 cm), and three target widths (3.5, 5.2 and 6.9 cm). The target amplitudes correspond to visual angles of 31.0, 39.8 and 48.5 degrees at 20 cm and 12.6, 16.5 and 20.4 degrees at 50 cm eye-screen distance. The target widths correspond to visual angles of 10.0, 14.8 and 19.6 degrees at 20 cm and 4.0, 6.0 and 7.9 degrees at 50 cm eye-screen distance if the target would have been placed at the centre of view.

The participants were first asked to give their permission that data could be collected in an anonymous format and used for research. After being seated in front of the screen, a questionnaire with their demographic information was filled out. Then the head-strap with the NextMind sensor attached was adjusted to their head until the calibrations software indicated good contact to all nine electrodes. A calibration was done, lasting approximately 30 sec. Then half of the subjects would start with Experiment 1 and the other half with Experiment 2. After each of the experiments, three questions were asked, using a 10-point Likert scale for the reply (1 was “low”, 10 was “high”): “How mentally demanding was the task?”, “How physically demanding was the task?” and “How comfortable do you feel right now?”. At the very end of the experiment, the participants were prompted for general comments on their experience. The experiments lasted between 35 to 60 minutes.

<sup>3</sup><https://www.next-mind.com/documentation/unity-sdk/issues-workarounds/>



**Figure 3: Distribution (boxplot) of the activation time  $AT$  (ms) over all trials. Left shows all included activations; right shows the central part with the median at 2349 ms.**



**Figure 4: Estimated throughput ( $TP$ , bits/s) vs Fitts Law Index of Difficulty ( $ID$ ), over all trials.**

## 4 RESULTS

The  $AT$  of the first target was discarded since the initial gaze point could not be assumed to be diagonally across. The following nine activations were all included for each sequence, with the initial target also being presented at the end of the trial.

This results in  $6 \times 2 \times 2 \times 3 \times 3 \times 9 = 1944$  data points, 972 at each eye-screen distance. Of these, targets that were not activated after 10000 ms were discarded. Using this criterion, 5 targets were discarded at 20 cm and 1 target at 50 cm; the remaining 99.5% and 99.9% of presented targets, respectively, were correctly activated in due time.

The  $AT$  is distributed as shown in Fig. 3. The overall median is 2349 ms, with 91.5% of the values between 2300 and 2400 ms. At 50 cm eye-screen distance, the marginal means is 2552 ms and 2445 ms for experiment 1 and 2 stimuli, respectively. At 20 cm, it is 2493 ms and 2437 cm.

The mean  $TP$  was 0.82 bits/s ( $SE = 0.008$  bits/s) under the assumption that the effective tolerance and target width are equal ( $W_e = W$ ). Across the varying target width and amplitude and eye-screen distances, the lowest throughput was 0.58 bits/s ( $SE = 0.005$

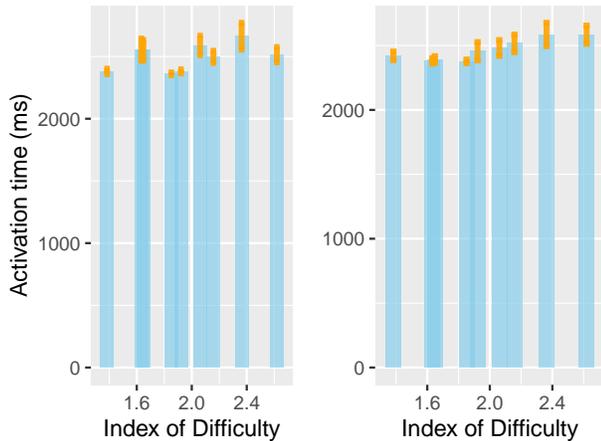
bits/s) at the largest target amplitude and smallest width, and the highest 1.17 bits/s ( $SE = 0.12$  bits/s) at the smallest amplitude and largest width, see Fig. 4.

For each eye-screen distance separately, a linear mixed-effect model, with a random effect caused by the participant, was used to analyse which of the factors, the target width ( $W$ ) and target amplitude ( $A$ ) have an effect on the mean  $TP$ .  $W$  had an effect both at 50 cm ( $\chi^2(1) = 504, p < .001$ , semi-partial  $R^2 = .402$  by model comparison to a no fixed-effects baseline), and at 20 cm ( $\chi^2(1) = 75.4, p < 0.001$ , semi-partial  $R^2 = .075$ ).  $A$  additionally had an effect both at 50 cm ( $\chi^2(1) = 551, p < .001, R^2 = .658$  for the entire model by model comparison to the previous model with  $W$  as a fixed effect), and at 20 cm ( $\chi^2(1) = 46.5, p < .001, R^2 = .119$ ). As an alternative hypothesis, it was additionally analysed if the Index of Difficulty ( $ID$ ) had an effect on the  $TP$ .  $ID$  did show an effect compared to a no fixed-effects baseline ( $\chi^2(1) = 1068, p < .001$ , semi-partial  $R^2 = .662$ ) at 50 cm distance, as well as at 20 cm distance ( $\chi^2(1) = 125, p < .001$ , semi-partial  $R^2 = .121$ ). Comparing the  $R^2$  of the two models with fixed effects of  $A$  plus  $W$  vs  $ID$  indicates that the latter explains slightly more variance. Further, the Akaike Information Criteria (AIC) for the former model with fixed effects of  $A$  plus  $W$  is  $-1672$  and  $1156$  for 50 cm respectively 20 cm, compared to a model with  $ID$  as a fixed effect of  $-1687$  and  $1152$  respectively, which indicates that the simpler model using  $ID$  may be the preferable model in this case.

At 50 cm, the marginal means of the  $AT$  were 2589 ms ( $SE = 49$  ms), 2476 ms ( $SE = 33$  ms), and 2432 ms ( $SE = 31$  ms) for target widths 4.0, 6.0, and 7.9 degrees. The marginal means were 2507 ms ( $SE = 39$  ms), 2533 ms ( $SE = 47$  ms) and 2457 ms ( $SE = 27$  ms) for target amplitudes 12.6, 16.5 and 20.4 degrees. Factored over the resulting 9 levels of Fitts' law  $ID$ ,  $AT$  varies from a marginal means of 2361 ms ( $SE = 11$  ms) at  $ID = 1.85$  to 2664 ms ( $SE = 108$  ms) at  $ID = 2.36$ . At 20 cm, the marginal means were 2550 ms ( $SE = 44$  ms), 2455 ms ( $SE = 34$  ms), and 2392 ms ( $SE = 14$  ms) at 10.0, 14.8, and 19.6 degrees, respectively. The marginal means were 2431 ms ( $SE = 25$  ms), 2474 ms ( $SE = 39$  ms) and 2491 ms ( $SE = 33$  ms) for target amplitudes 31.0, 39.9 and 48.5 degrees. Factored over the resulting 9 levels of Fitts' law  $ID$ ,  $AT$  varies from a marginal means of 2377 ms ( $SE = 14$  ms) at  $ID = 1.85$  to 2585 ms ( $SE = 91$  ms) at  $ID = 2.36$ .

The marginal means of the  $AT$  factored over the Fitts law  $ID$  is seen in Fig. 5.

As above, for each eye-screen distance separately, a linear mixed-effect model, with a random effect caused by the participant, was used to analyse which of the factors target width ( $W$ ) and target amplitude ( $A$ ) have an effect on activation time ( $AT$ ).  $W$  had an effect both at 50 cm ( $\chi^2(1) = 8.54, p = .003$ , semi-partial  $R^2 = .009$  by model comparison to a no fixed-effects baseline), and at 20 cm ( $\chi^2(1) = 11.78, p = .0006$ , semi-partial  $R^2 = .012$ ). As an alternative hypothesis, it was additionally analysed if the  $ID$  had an effect on the  $AT$ .  $ID$  did show an effect ( $\chi^2(1) = 13.3, p = .0003$ , semi-partial  $R^2 = .013$ ) at 20 cm distance, however no significant effect was found at 50 cm ( $\chi^2(1) = 3.71, p = .054$ ). The effect of  $A$  was not significant in either case.



**Figure 5: Marginal means of the activation time (ms) vs Fitts Law Index of Difficulty (ID) at 50 cm (left) and 20 cm (right). The error bars denote the standard error of the mean.**

#### 4.1 User experiences

The users found Experiment 1 a bit less mentally demanding than Experiment 2, with mean scores being 3.5 and 5.2, respectively. One user commented: “In the 1st experiment the eyes go naturally; the 2nd experiment requires active guidance of eyes”. This refers to the fact that only one target was visible in the first experiment. All users gave a high score to the comfort questions, mean score 8.3 and 8.0, respectively. Even so, some of the users’ comments indicated signs of discomfort when sitting 20 cm from the screen. Also, some of them mentioned that the strap was hurting and left ‘red skin’ on their foreheads. This points to the importance of adjusting the strap very carefully. Finally, the high contrast between the targets and the background caused a problem for one of the participants: “I saw some after-image from the high contrast, sometimes the after-image appeared in the targets of the next sequence. That made it slightly more difficult, but with no noticeable activation time difference to me during the experiment.”

#### 5 SIMULATED USE-CASE

An electric wheelchair simulation environment was developed to test the NextMind BCI sensor.<sup>4</sup> In the simulation, a driver controls the wheelchair by attending to one of the arrows on a control panel at the wheelchair. While attending to the flickering pattern on the control screen, the linear velocity of the wheelchair is adjusted. Eventually, motion begins in the desired direction. To stop movements, the driver simply stops attending to any of the flickering patterns.

When a smart device in the home is approached, for instance a television (TV), a remote control interface for the TV may be launched on the wheelchair control panel. At this stage, the driver can select one of the actions (e.g., Power On/Off, Next Channel, Previous Channel, Volume Up/Down) by focusing on corresponding



**Figure 6: Simulated use scenario with a wheelchair control panel and a NextMind BCI sensor**

patterns (c.f. Fig. 1, right). In order to take the wheelchair control back, the driver will need to focus on a “Back” action pattern.

Flashing control buttons may also be located at the smart-home devices themselves. To get an idea on the range by which the sensor can pick up the flashing from a NeuroTag, three participants made a follow-up check with one target flickering (similar to Experiment 1) at 50, 110 and 140 cm distance from the screen. At 50 cm, just 1 out of 729 activations was unsuccessful. The median AT at this distance was 2335 ms, similar to our findings in Experiment 1. At 110 cm, 16 activations were unsuccessful, and the median AT was 2352 ms. At 140 cm, we noticed a limitation in the operational range: 81 activations could not be done, which suggests the error rate is slightly more than 10% at this distance. Also, the median AT increased slightly to 2641 ms.

#### 6 DISCUSSION

The most remarkable results from the experiments are the very low error rates of less than 1% for all test conditions. Generally error rates in eye gaze experiments are higher, and Hou et al. [Hou et al. 2020] reporting an error rate of 14.6% when combining VOG gaze tracking and MI BCI. Bækgaard et al. [Bækgaard et al. 2019] found error rates for gaze, head, and mouse pointing in a head-mounted display to be 1.53%, 0.26%, and 0.24%, respectively. Soukoreff and MacKenzie [Soukoreff and MacKenzie 2004, Table 4] surveyed nine Fitts’ law studies with error rates ranging from 1.6% to 32%. The low error rate suggest that SSVEP BCI may become a viable alternative to VOG-based eye tracking or other pointing devices if reliability is a major concern, for instance for activation of an alarm.

Xing et al. [2018] produced their own SSVEP sensor system with eight dry electrodes. The accuracy of classification was 93.2% with a one-second AT, providing an average information transfer rate of 92.35 bits/min, corresponding to 1.54 bits/sec. This is better than the mean TP obtained in our Fitts’ law study (0.82 bits/sec). The AT of 2349 ms we observed, but it comes at the cost of an error rate of almost 7%, which would not be acceptable for communication or gaming systems.

<sup>4</sup><https://youtu.be/PqelTCLJ77U>

The high comfort ratings, the short calibration process and the ease of mounting the sensor unit indicate the potential of SSVEP for novice users with special needs, who are likely to have difficulty with light reflections or sensitivity to motion of the head and the camera. Also, the mobility provided by the wireless connection and the low power consumption supports independent living, because when connected to a stationary PC, the user can move around for hours in his or her home without the burden of a computer that needs frequent re-charging. Our simulated use-case with wheelchair and smart-home control illustrates the feasibility of the SSVEP interaction for activities of daily living when coupled with a carefully designed control panel.

There are three major challenges with the SSVEP based system that we tested. The median *AT* of 2349 ms limits productivity when used for tasks that require longer sequences of selections, for instance, eye typing. We observed two users typing five sentences each on a two-step on-screen keyboard. They achieved a mean typing speed of 1.8 wpm, with a maximum of 2.4 wpm.<sup>5</sup> In comparison, Rähkä's review of gaze typing experiments found gaze typing speeds for soft keyboards of 6 wpm when using dwell-time settings of 1000 ms and 18.98 wpm when the dwell-time was between 250 and 420 ms [Rähkä 2015]. Evidently, the long *AT* has a strong negative impact on the efficiency of SSVEP-based communication. Also, in numerous types of games, it is difficult to compete against opponents if *AT*s are more than, say, two seconds.

The second main challenge is the limitation of ten selectable targets for each layout. Clearly, this is outside the requirement for 26+ keys for a full on-screen keyboard, which most current VOG-based gaze tracking systems support. There are a number of text entry systems with novel techniques that might work with 10 keys or less. For instance, Hansen et al. [Hansen et al. 2004] reported a typing speed of 6 wpm on the GazeTalk keyboard, which has 10 keys. This suggests that the SSVEP system tested in our study might achieve higher communication speeds than the 1.8 wpm that we observed.

The third main issue we have observed was to get a good electrode contact to the scalp. Of the more than 10 people we have tried the system on, three could not get a good calibration due to thick hair. Among those three was a regular user of VOG-based gaze communication. She noted, though, that had the VOG-gaze tracker used for her daily communication not been an effective solution, she might consider cutting her hair to make a SSVEP system work. A last issue with this system, as noted by [Rupp 2014], is that in some clinical settings, SSVEP may not be the most comfortable, as the occipital electrodes can cause discomfort if the patient has to lay on a pillow.

Besides the low number of participants, our study has additional limitations. We found that larger targets were only slightly faster to select than smaller targets, and there were no significant effects of target amplitudes on *AT*. However, we only used targets with a limited variation of width and therefore it is still unknown the extent to which even smaller targets might impact error rates and *AT*s. And even though we included a short viewing distance of 20 cm in our study, the higher amplitude it necessitated was not extreme enough to challenge the classifier of the NextMind system.

Finally, the differences in *AT* between presenting one or nine targets on the same display was less than 100 ms. It would be interesting to test if, for example, 25 simultaneous targets would degrade *AT* and/or increase the error rate. But, since the NextMind system only supports 10 targets, this was not possible.

## 7 CONCLUSION

The NextMind SSVEP system has a low error rate when used for target selection by gaze. It is relatively slow compared to common VOG-based eye tracking systems and affords a limited number of ten targets per screen. Participants with thick hair are difficult to calibrate. However, compared to other BCI methods, it has a high reliability. The ease of mounting the headband sensor and the wireless connection to a PC gives rise to further applied research within, for instance, control of wheelchairs, exoskeletons, and smart-home devices. In conclusion, we foresee that a new generation of wireless SSVEP gaze tracking systems may replace some of the use-cases from traditional VOG eye tracking technology, when *AT* is not of high importance.

## ACKNOWLEDGMENTS

The research has partly been supported by the EU Horizont 2020 ReHyb project and The Bevica Foundation.

## REFERENCES

- Per Bækgaard, John Paulin Hansen, Katsumi Minakata, and I Scott MacKenzie. 2019. A Fitts' law study of pupil dilations in a head-mounted display. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications – ETRA '19*. ACM, New York, 32:1–32:5.
- Justin Cuaresma and I Scott MacKenzie. 2017. FittsFace: Exploring navigation and selection methods for facial tracking. In *Proceedings of Human-Computer Interaction International - HCII '17 (LNCS 10278)*. Springer, Berlin, 403–416. [https://doi.org/10.1007/978-3-319-58703-5\\_30](https://doi.org/10.1007/978-3-319-58703-5_30)
- Xujiong Dong, Haofei Wang, Zhaokang Chen, and Bertram E Shi. 2015. Hybrid brain computer interface via Bayesian integration of EEG and eye gaze. In *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE, 150–153.
- T Felzer, I Scott MacKenzie, and John Magee. 2016. Comparison of two methods to control the mouse using a keypad. In *Proceedings of the 15th International Conference on Computers Helping People With Special Needs – ICCHP '16 (LNCS 9759)*. Springer, Berlin, 511–518. [https://doi.org/10.1007/978-3-319-41267-2\\_72](https://doi.org/10.1007/978-3-319-41267-2_72)
- Antonio Frisoli, Claudio Loconsole, Daniele Leonardi, Filippo Banno, and Massimo Bergamasco. 2012. A new gaze-BCI-driven control of an upper limb exoskeleton for rehabilitation in real-world tasks. *IEEE Transactions on Systems Man & Cybernetics Part C* 42, 6 (2012), 1169–1179.
- John Paulin Hansen, Kristian Tørring, Anders Sewerin Johansen, Kenji Itoh, and Hirotsuka Aoki. 2004. Gaze typing compared with input by head and hand. In *Proceedings of the 2004 Symposium on Eye Tracking Research & Applications – ETRA '04*. ACM, New York, 131–138.
- Mehedi Hassan, John Magee, and I Scott MacKenzie. 2019. A Fitts' law evaluation of hands-free and hands-on input on a laptop computer. In *Proceedings of the 21st International Conference on Human-Computer Interaction – HCII '19 (LNCS 11572)*. Springer, Berlin, 234–249. <https://doi.org/10.1007/978-3-030-23560-4>
- Baosheng James Hou, Per Bækgaard, I. Scott MacKenzie, John Paulin Hansen, and Sadasivan Puthusserypady. 2020. GIMIS: Gaze input with motor imagery selection. In *ACM Symposium on Eye Tracking Research and Applications – ETRA '20 Adjunct*. ACM, New York, Article 18, 10 pages. <https://doi.org/10.1145/3379157.3388932>
- Eui Chul Lee, Jin Cheol Woo, Jong Hwa Kim, Mincheol Whang, and Kang Ryoung Park. 2010. A brain-computer interface method combined with eye tracking for 3D interaction. *Journal of Neuroscience Methods* 190, 2 (2010), 289–298. <https://doi.org/10.1016/j.jneumeth.2010.05.008>
- Robert Leeb, Serafeim Perdakis, Luca Tonin, Andrea Biasucci, Michele Tavella, Marco Creatura, Alberto Molina, Abdul Al-Khodairy, Tom Carlson, and José dR Millán. 2013. Transferring brain-computer interfaces beyond the laboratory: Successful application control for motor-disabled users. *Artificial Intelligence in Medicine* 59, 2 (2013), 121–132.
- I Scott MacKenzie. 2012. *Human-computer interaction: An empirical research perspective*. Elsevier, Amsterdam.

<sup>5</sup>[https://youtu.be/RWY7Y-w\\_EYQ](https://youtu.be/RWY7Y-w_EYQ)

- I Scott MacKenzie and Robert J. Teather. 2012. FittsTilt: The application of Fitts' law to tilt-based interaction. In *Proceedings of the 7th Nordic Conference on Human-Computer Interaction – NordiCHI 2012*. ACM, New York, 568–577. <https://doi.org/10.1145/2399016.2399103>
- Alexey Petrushin, Jacopo Tessadori, Giacinto Barresi, and Leonardo S Mattos. 2018. Effect of a click-like feedback on motor imagery in EEG-BCI and eye-tracking hybrid control for telepresence. In *2018 IEEE/ASME International Conference on Advanced Intelligent Mechatronics – AIM '18*. IEEE, New York, 628–633.
- Kari-Jouko Rähkä. 2015. Life in the fast lane: Effect of language and calibration accuracy on the speed of text entry by gaze. In *IFIP Conference on Human-Computer Interaction – INTERACT '15*. Springer, Berlin, 402–417.
- Maria Francesca Roig-Maimó, I Scott MacKenzie, Cristina Manresa, and Javier Varona. 2017. Evaluating Fitts' law performance with a non-ISO task. In *Proceedings of the 18th International Conference of the Spanish Human-Computer Interaction Association*. ACM, New York, 51–58. <https://doi.org/10.1016/j.ijhcs.2017.12.003>
- R. Rupp. 2014. Challenges in clinical applications of brain computer interfaces in individuals with spinal cord injury. *Frontiers in Neuroengineering* 7 (2014), 38. <https://doi.org/10.3389/fneng.2014.00038>
- Immo Schuetz, T Scott Murdison, Kevin J MacKenzie, and Marina Zannoli. 2019. An Explanation of Fitts' Law-like Performance in Gaze-Based Selection Tasks Using a Psychophysics Approach. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.
- R William Soukoreff and I Scott MacKenzie. 2004. Towards a standard for pointing device evaluation: Perspectives on 27 years of Fitts' law research in HCI. *International Journal of Human-Computer Studies* 61 (2004), 751–789. <https://doi.org/10.1016/j.ijhcs.2004.09.001>
- William Speier, Corey Arnold, and Nader Pouratian. 2013. Evaluating true BCI communication rate through mutual information and language models. *PLoS One* 8, 10 (2013), e78432.
- Boris M Velichkovsky and John Paulin Hansen. 1996. New technological windows into mind: There is more in eyes and brains for human-computer interaction. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 496–503.
- Roman Vilimek and Thorsten O Zander. 2009. BC (eye): Combining eye-gaze input with brain-computer interaction. In *International Conference on Universal Access in Human-Computer Interaction – UAHCI '09 (LNCS 5615)*. Springer, New York, 593–602.
- Xiao Xing, Yijun Wang, Weihua Pei, Xuhong Guo, Zhiduo Liu, Fei Wang, Gege Ming, Hongze Zhao, Qiang Gui, and Hongda Chen. 2018. A high-speed SSVEP-based BCI using dry EEG electrodes. *Scientific reports* 8, 1 (2018), 1–10.
- Xinyi Yong, Mehrdad Fatourehchi, Rabab K Ward, and Gary E Birch. 2011. The design of a point-and-click system by integrating a self-paced brain-computer interface with an Eye-tracker. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* 1, 4 (2011), 590–602.
- Thorsten O Zander, Matti Gaertner, Christian Kothe, and Roman Vilimek. 2010. Combining eye gaze input with a brain-computer interface for touchless human-computer interaction. *Intl. Journal of Human-Computer Interaction* 27, 1 (2010), 38–51.
- X. Zhang and I. S. MacKenzie. 2007. Evaluating eye tracking with ISO 9241 – Part 9. In *Proceedings of HCI International 2007 – HCII 2007*. Heidelberg: Springer, 779–788.