Downloaded from orbit.dtu.dk on: May 09, 2024

DTU Library

Pointing by Gaze, Head, and Foot in a Head-Mounted Display

Minakata, Katsumi; Hansen, John Paulin; Bækgaard, Per; MacKenzie, I. Scott; Rajanna, Vijay

Published in:
Proceeding of ETRA 2019

Link to article, DOI:
10.1145/3317956.3318150

Publication date:
2019

Document Version
Early version, also known as pre-print

Link back to DTU Orbit

Citation (APA):
Minakata, K., Hansen, J. P., Bækgaard, P., MacKenzie, I. S., & Rajanna, V. (2019). Pointing by Gaze, Head, and Foot in a Head-Mounted Display. In *Proceeding of ETRA 2019* Article 69 Association for Computing Machinery. https://doi.org/10.1145/3317956.3318150

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Hands-free Pointing by Foot, Head, and Gaze in a Head-Mounted Display

Anonymous

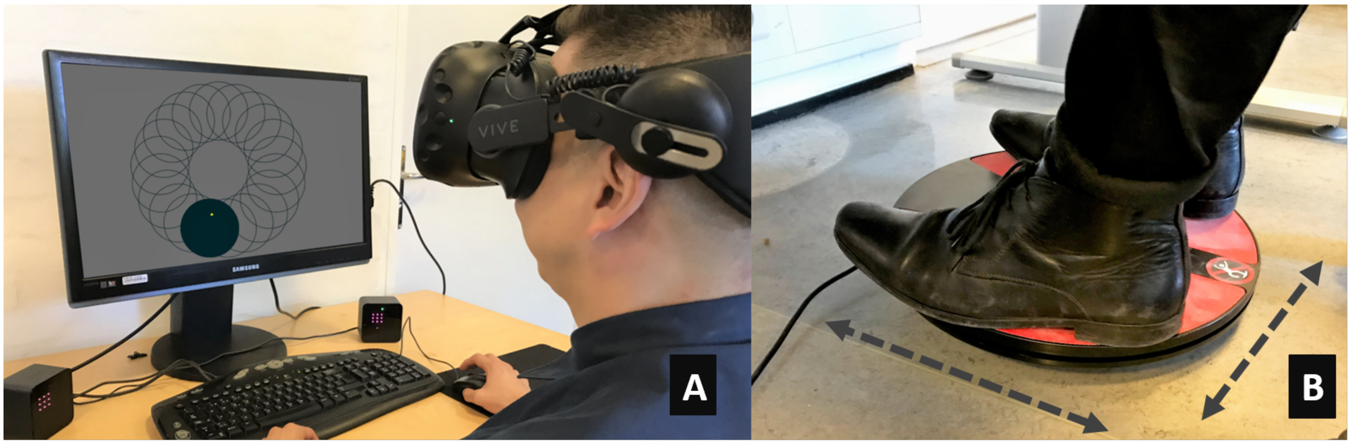

Figure 1: A. Participant with an HTC VIVE headset (HMD) equipped with the Pupil Lab's binocular gaze tracking unit (used for gaze input). The Fitts' multi-directional selection task that the participant is performing inside the headset is also shown on the monitor. The headset also tracks orientation of the head (used for head input). B. Feet on a 3DRudder foot mouse with 360° of movement control (used for foot input). Tilting the 3DRudder moves the cursor in that direction.

ABSTRACT

Hands-free interaction is important when using head-mounted displays to support manual work or when motor disability prevents finger and hand control. This paper presents a Fitts' law experiment with 27 participants comparing time to activate, throughput, target width, and pupil dilation when pointing by feet, head, and gaze in a VR headset. Mouse input provides a baseline. Gaze was slower than the other pointing methods, especially in the lower visual field. Throughput for foot and gaze were lower than mouse and head pointing and their effective target widths were also higher than both head and mouse. Pupil dilations were smaller for gaze than foot, even though subjects rated gaze as the most mentally demanding method.

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.
CHI '19, May 2019, Glasgow, UK
© 2018 Association for Computing Machinery.
ACM ISBN 978-x-xxxx-xxxx-x/YY/MM. . . $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

CCS CONCEPTS

• **Human-centered computing → Pointing devices**;

KEYWORDS

Fitts' law, ISO 9241-9, foot interaction, gaze interaction, head interaction, dwell activation, head mounted displays

ACM Reference format:
Anonymous. 2018. Hands-free Pointing by Foot, Head, and Gaze in a Head-Mounted Display. In *Proceedings of ACM Conference, Glasgow, UK, May 2019 (CHI '19),* 12 pages.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Pointing within the realm of virtual reality (VR) is a relevant area for HCI research [43]. Head- and gaze-tracking are now available in VR headsets, thus presenting a number of questions for HCI researchers: How big should targets be if aimed at with head motions? How fast can people activate a button in VR by using gaze? Motivated by recent developments in Augmented Reality (AR) headsets like Microsoft's Hololens and MagicLeap One, this paper focuses on hands-free interaction methods for head-mounted displays (HMD).

In scenarios with situation-induced impairments, a user's hands are engaged in other tasks, and hence unavailable for pointing or text entry [11, 24, 25]. For example, a surgeon in need of medical image information during surgery, cannot use his/her hands to interact with the images, since they are occupied and sterilized [8, 16]. Wearing an AR HMD, the surgeon may use gaze pointing to control the images. We expect to see HMDs supporting work tasks in the future, and therefore we need a better understanding of the size of the UI elements and potential hands-free input modalities that work best for HMDs.

According to a 2016 disability report, 7.1% of the United States population have an ambulatory disability [3] that restricts the movements of the limbs and makes it harder or even impossible to work on a computer using a mouse or keyboard. Also, it is estimated that by 2050 there will be nearly 3.6 million people living with the loss of a limb [48]. Individuals without full hand or finger control rely on alternative input, for instance gaze, to accomplish simple tasks on a computer or to communicate with others [7, 32]. For such individuals, the need for hands-free alternatives when interacting in HMDs is crucial. Accessability to VR is particularly relevant, since this medium provides opportunities to experience places and events that are not accessible in real life. Also, VR models of products, buildings, and surroundings can be used to evaluate their accessability by individuals in a wheelchair [6, 15].

When situations change there may be a need to shift between alternative input methods. A foot control may work well when seated [46], but when laying in a bed head control might be a better option. On the other hand, head control may not work well when the user's body is not stable, e.g., while driving in a car. Gaze interaction is a potential candidate in such situations. These discussions leave us with the question of which among these three hands-free input methods (gaze, head, foot) is the best choice when interacting with an HMD.

For half a century, Fitts' law tasks have been used to quantitatively evaluate a range of pointing methods and devices, such as the mouse, stylus, hand controllers, track pads, head pointing, and gaze pointing [44]. The use of a standard procedure makes it possible to compare performance across pointing methods and across studies. However, only a few studies have applied this procedure to HMDs (e.g.[37]).

A Fitts' law point-select task is easy to do. It can be be performed by children and computer-naive individuals, which is not the case for typing tasks. Therefore, such tasks hold potential as a standard assessment method for testing alternative input methods for their effectiveness with individuals with motor challenges. What should the target size be for an individual to master pointing? How many selections can he or she do per minute? The last issue relates to communication speed while typing. Hansen et al. [14] observed that HMDs work well for some people with motor- and cognitive challenges. Immersive headsets eliminate external stimuli that may otherwise distract the user. Motivated by these observations, we seek to establish a baseline of performance indicators on a non-disabled user group.

Future headsets with eye sensors can provide continuous data on changes in pupil size. Numerous studies have found that pupils dilate when cognitive load increases. Hess and Polt [17] originally suggested that pupil dilation could be used as an index of mental activity during multiplication. Kahneman and Beatty [23] confirmed this finding in a separate study, which gave rise to pupillometry as a discipline [1, 27, 45] and created an interest among HCI researchers to include pupil measures in user performance assessments (e.g., [18]). It is an open question if pupils dilate when simple visual-motor tasks increase in difficulty (e.g., [10, 21, 22]). Do pupils dilate more when the level of difficulty in a Fitts' law task increases? Is there a differences between pupil dilation for the mouse, foot, head, and gaze input conditions? If so, do these differences align with subjective user experiences?

Virtual reality applications, by their very nature, display the world in 3D by positioning objects at different depths. Also, the majority of VR applications introduce motion in the user's field of view to achieve an immersive experience. However, in the current study we did not manipulate depth or introduce motion cues. The targets in our Fitts' law task were positioned at the same depth since our goal is to first establish a neutral baseline of input characteristics without depth cues or background motion. Future studies will then include depth cues and motion as independent variables to study how they interact with target size, target amplitude, and pointing method. Additionally, we only examined one equipment setup for each pointing method and, as a practical consideration, we only included two target amplitudes and two target widths. Finally, we only consider the input methods as mono-modal leaving out any combined used of them.

The main contributions of this paper are (i) a comparison of three hands-free input methods (feet, head, gaze) for HMDs using a standard procedure, (ii) an analysis of pupil data associated with the three input conditions, and (iii) discussion on the potential of a standard test procedure for VR and accessible computing.

2 EVALUATION USING FITTS' LAW

Our evaluation used Fitts' law as per the methodology in the ISO 9241-9 standard for non-keyboard input devices [20]. The most common ISO 9241-9 evaluation procedure uses a two-dimensional task with targets of width W arranged along a layout circle. Selections proceed in a sequence moving across and around the circle (see Figure 2). Each movement covers an amplitude A - the diameter of the layout circle (this is same as the distance between the centers of two opposing targets). The movement time (*MT*, in milliseconds) is recorded for each trial and averaged over the sequence.

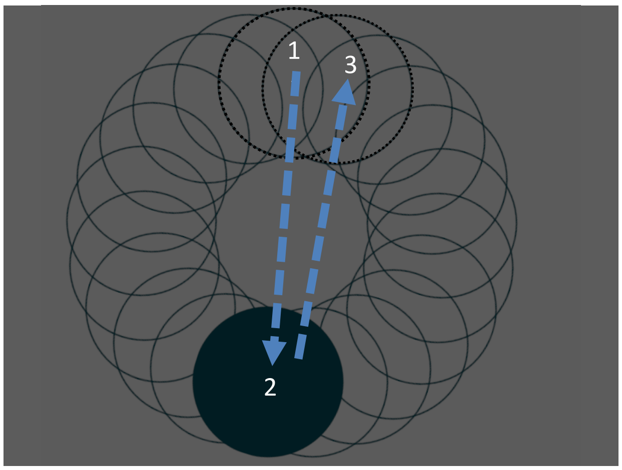

Figure 2: Two-dimensional target selection task in ISO 9241-9. After pointing at circle 1 for a prescribed time (i.e., 300 ms) selection occurs and the next target appears in solid (2). When this target is selected, the next target (3) appears. Arrows, numbers, and circle annotations were not shown in the actual task display.

The difficulty of each trial is quantified using an index of difficulty (*ID*, in bits) and is calculated from *A* and *W* as shown in Equation 1.

$$ID = log_2\left(\frac{A}{W} + 1\right) \quad (1)$$

The main performance measure in ISO 9241-9 is throughput (*TP*, in bits/second or bps) which is calculated over a sequence of trials as the $ID - MT$ ratio, as shown in Equation 2.

$$TP = \frac{ID_e}{MT} \quad (2)$$

The standard specifies calculating throughput using the effective index of difficulty (ID_e). The calculation includes an adjustment for accuracy to reflect the spatial variability in responses as shown in Equation 3.

$$ID_e = log_2\left(\frac{A_e}{W_e} + 1\right) \quad (3)$$

where

$$W_e = 4.133 \times SD_x \quad (4)$$

The term SD_x is the standard deviation in the selection coordinates computed over a sequence of trials. Selections are projected onto the task axis, yielding a single normalized x-coordinate for each trial. The factor 4.133 adjusts the target width for a nominal error rate of 4% under the assumption that the selection coordinates are normally distributed. The effective amplitude (*A*e) is the actual distance traveled along the task axis. (see [28] for additional details).

Throughput is a potentially valuable measure of human performance because it embeds both the speed and accuracy of participant responses. Comparisons between studies are therefore possible, with the proviso that the studies use the same method in calculating throughput.

3 RELATED WORK

The ISO 9241-9 standard for pointing devices has been used to compare alternative methods to control a mouse cursor using a numeric keypad [9] . In a user study with non-disabled participants, throughputs were about 0.5 bits/s among the methods compared. A user from the target community (the first author of the paper, who has the neuromuscular disease Friedreich Ataxia) achieved throughputs around 0.2 bits/s, reflecting a lower performance. In a second study using the same ISO 9241-9 standard, this individual participated in a comparison of click actuation methods for users of a motion-tracking mouse interface [29]. In this case, mean throughputs of about 0.5 bits/s were achieved. Keates et al. [26] found a significant difference in throughput with the mouse between able-bodied users (4.9 bits/s) and motion-impaired users (1.8 bits/s).

The ISO-9241-9 standard has also been used to evaluate the efficiency of gaze and head tracking, but not involving individuals with motor challenges. Zhang and MacKenzie [47] measured a throughput of 2.3 bits/s for long dwell-time selection (750 ms) and 3.1 bits/s and for short dwell-time selection (500 ms). For head tracking systems, De Silva et al. [4] reported a throughput of 2.0 bits/s, and Roig et al. [42] reported 1.3 bits/s when interacting with a tablet by head movements detected by the device's front-facing camera. Five studies of mouse throughput, reported by Soukoreff and MacKenzie [44], found throughputs in the range of 3.7 to 4.9 bits/s.

Foot pointing with HMDs

While foot input has been used for navigational tasks (moving, turning, etc.) in HMDs for gaming [33], point-and-select interactions in HMDs with foot input is yet to be explored. There is prior research on foot input in desktop settings for point-and-select interactions, however [46]. Pearson et al. [36] first demonstrated an input device called "Moles" that functions similar to a mouse. Pakkanen et al. [35] presented a foot-operated trackball that could be used for non-accurate pointing tasks. Dearman et al. [5] demonstrated tapping on a foot pedal as a selection trigger for text entry on a mobile device.

Recently, foot input has been combined with gaze input in a multi-modal interaction setup. Göbel et al. [12] first combined gaze input with foot input for secondary navigation tasks like pan and zoom in zoomable information spaces. Furthermore, gaze input has been combined with foot input, achieved through a wearable device, for precise point-and-click interactions [39] and for text entry [38]. Lastly, Hatscher et al. [16] demonstrated how a physician performing minimally-invasive interventions can use their gaze and input from the foot to interact with medical image data presented on a display. These examples demonstrate the potential and the benefits of combining foot input with hand or gaze inputs for interactions in a desktop setting.

Head and eye-gaze pointing with HMDs

Pointing at objects in a VR space is mainly achieved through an external controller and the ray-casting method [34]. HMDs like HTC Vive, Oculus Rift, Samsung Gear VR, and so on, provide the user with a hand-held controller to interact with objects in VR. For those VR systems where controllers need to be tracked within a fixed space, this is a serious limitation of mobility, and also, the user's hands are then occupied in operating the controllers. Qian et al. [37] compared head-only input to eye+head and eye-only inputs in a click-selection task in a FOVE HMD, which is the first commercially available headset with build-in gaze tracking . They used a 3D background decoration and volume surface on targets. The authors found that head-only input resulted in a significantly lower error rate (8%) compared to eye-only (40%) and eye+head inputs (30%). Consequently, head-only input achieved the highest throughput (2.4 bits/s) compared to eye-only (1.7 bits/s) and eye+head inputs (1.7 bits/s). Task completion time and subjective ratings were also in favour of the head-only input method.

Hansen et al. [13] compared pointing with head to pointing with gaze and mouse inputs in a Fitts' law experiment with the same FOVE HMD, but within a neutral 2D (i.e., "flat") scene. The authors compared dwell (300 ms) and mouse click as selection methods. Overall, throughput was highest for the mouse (3.2 bits/s), followed by head pointing (2.5 bits/s) and gaze pointing (2.1 bits/s). Blattgerste et al. [2] investigated if using head or gaze pointing is an efficient pointing method in HMDs. The authors tested pointing and selection on two interfaces: a virtual keyboard and a selection menu typical in VR applications. It was found that aiming with gaze significantly outperforms aiming with the head in terms of time-on-task and head movement. When interacting with the virtual keyboard, pointing with the eye reduced the time-on-task by 32% when compared to pointing with the head. Similarly, when working on the menu, gaze input reduced the aiming time by 12% compared to head input.

Pupil dilation and pointing

Richer and Beatty [41] were the first to study the relationship between motor task complexity and pupil dilation. When more fingers where involved in performing a sequence of key presses, the amplitude of pupil dilation increased. Jiang et al. [21, 22] conducted a simple continuous aiming task where a tooltip is placed on targets with various sizes and amplitudes, resembling a micro-surgery task. The results showed that higher task difficulty, measured in terms of *ID*, evoked higher peak pupil dilation and longer peak duration. Fletcher et al. [10] also used a Fitts' law movement task to manipulate motor response precision. Contrary to previous findings, increased precision demands were associated with *reduced* pupil diameter during response preparation and execution. The authors suggest that for discrete tasks dominated by precision demands, a decrease in pupil diameter is an indicator of increased workload.

Summary

The ISO 9241-9 standard procedure has been applied in several studies of input devices, a couple of them involving a user with motor disabilities. Foot-assisted point-and-click interaction has been examined in several studies, but none of them involving HMDs. Three studies compared head and gaze input in HMDs, two using a FOVE HMD found head to be superior to gaze, while the last study, applying a high-precision gaze tracker, found gaze to outperform head with rather large margins.

Pupil dilations has been found to increase with *ID*, but recently Fletcher et al. [10] suggested that a distinction between task complexity and precision may be needed, since they found a significant decrease in pupil diameter when precision demands increased.

4 METHOD

Participants

Twenty-seven participants were recruited from a university on a voluntary basis. The mean age of the participants was 25 years, (*SD* = 5 yrs); 18 male, 12 female. The mean inter-pupillary distance was 60.3, (*SD* = 3.15 mm). Most (45%) had tried HMDs several times before, and 28% only one time before. Some (45%) had previously tried gaze interaction. None had tried the foot-mouse. All participants had normal or corrected-to-normal vision.

Apparatus

An HTC VIVE HMD was used. The HMD has a resolution of 2160 × 1200 px, renders at a maximum of 90 fps, and has a field of view of 110° visual angle. A Pupil Labs binocular eye-tracking add-on system for the HTC VIVE was installed and collected data at 120 Hz, it has a gaze-accuracy and precision of < 1° and .08° visual angle, respectively. The background was brown-black (rgb[35,23,10]), the target circles were violet-blue (rgb[29,11,40]); when the targets were entered/selected they turned green-blue (rgb[0,30,36]), and the cursor was violet-red (rgb[52,0,0]). These colors were chosen because they are equiluminant and would not differentially influence the degree of dilation of participants' pupils during eye tracking. Participants were screened for color blindness, as a result. The HMD weighs 520 grams and has IR-based position tracking plus IMU-based orientation tracking. A Logitech corded M500 mouse was used for the manual input. A foot-mouse was used as an input device for foot pointing; it is capable of controlling a cursor on a monitor via foot movements. The foot mouse is made by 3DRudder and is a "foot powered VR and gaming motion controller" with the capability of 360° of movement, see Figure 1.

Software to run the experiment was a Unity implementation of the 2D Fitts' law software developed by MacKenzie, known as FittsTaskTwo[1]. The Unity version[2] includes the same features and display as the original; that is, with spherical targets presented on a flat 2D-plane, cf. Figure 2.

Procedure

The participants were greeted upon arrival and were asked to sign a consent form after they were given a short explanation of the experiment. Next, the participants' interpupillary distance was measured and the HTC VIVE headset was adjusted accordingly. Then, the participants were screened for color blindness with an online version of the Ishihara test [19].

The participants completed the baseline mouse condition as their first pointing method condition. The other three conditions were head-position, foot-mouse, and gaze-pointing. The order of these conditions was assigned according to a Latin square. For the mouse condition, participants' dominant hand was guided by the experimenter to the position of the mouse, which was required for moving the cursor. The participants only used the different input devices for pointing. Selection was performed by positioning the cursor inside the target for the prescribed dwell time. Based on the findings of Majaranta et al. [30], a dwell time setting of 300 ms was chosen.

For each pointing method, four levels of index of difficulty (*ID*) were tested, composed of two target widths (50 pixels, 75 pixels) and two target amplitudes (160 pixels, 240 pixels). The target widths spanned visual angles of about 2° and 4°, respectively.

Spatial hysteresis was set to 2.0. When targets were entered, their size doubled , while visually remaining constant.

[1]available at http://www.yorku.ca/mack/FittsLawSoftware/ [last accessed: last accessed - Sept. 13, 2018]

[2]available at https://github.com/GazeIT-DTU/FittsLawUnity [last accessed: last accessed - Sept. 13, 2018]

For each of the four *ID*s, 21 targets were presented for selection. As per the ISO 9241-9 procedure, the targets were highlighted one-by-one in the same order for all levels, starting with the top position (12 o'clock). When this target was selected, a target at the opposite side would be highlighted (approximately 6 o'clock), then when activated a target at 1 o'clock was highlighted and so on, moving clockwise, see Figure 2. The first target at 12 o'clock is not included in the data analysis in order to minimize the impact from initial reaction time.

The target layout was locked in world space and would not move with head motion. The pointer (i.e., cursor) was visible at all times and appeared as a red dot. For the mouse pointing condition, the cursor was the mouse position on screen. For gaze pointing, this was the location the participants' looked at on the monitor, as defined by the intersection of the two gaze vectors from the centre of both eyes on the target plane. For head pointing, the cursor was the central point of the headset projected directly forward.

Failing to activate 20% of the targets in a 21-target sequence triggered a repeat of that sequence. Sequences were separated, allowing participants a short rest break as desired. Additionally, they had time to rest for a couple of minutes when preparing for the next pointing method.

Before testing the gaze pointing method, participants performed a gaze calibration procedure. This consisted of a bull's eye target that moved to one of six locations in the field of view. Participants were asked to fixate on the target each time it moved to one of the six locations. This sequence was repeated twice (i.e., once for calibration and once for verification of the calibration). The verification sequence was further conducted before the last pointing method condition and was collected in order to determine the calibration stability and quality from the beginning to the end of the experiment. Completing the full experiment took approximately 30 minutes for each participant.

Upon completion of the task, participants filled out a questionnaire. The questionnaire solicited demographic information and ratings of the different pointing methods for mental and physical workload and comfort. In addition, participants were encouraged to submit their impressions of using each method and were asked to rank the different pointing methods from most preferred to least preferred.

Design

The experiment followed a $4 \times 2 \times 2$ within-subjects design with the following independent variables and levels:

- Pointing method (mouse, head-position, foot-mouse, gaze)
- Target amplitude (160 pixels, 240 pixels)
- Target width (50 pixels, 75 pixels)

Pointing method was the primary independent variable. Target amplitude and target width were included to ensure the conditions covered a range of task difficulties.

The dependent variables were time to activate, throughput, and effective target width, calculated according to the standard procedures for ISO 9241-9. Errors were not possible because all targets were activated via dwell selection. In addition, we measured pupil size at 120 Hz throughout the experiment.

For each sequence, 21 trials were performed. There were 4 such sequences in a block, one sequence for each combination of target amplitude and target width. Four such blocks were performed for each participant, one for each pointing method. There were 27 participants in total. In all, 27 Participants × 4 Pointing Methods × 2 Target Amplitudes × 2 Target Widths yielded 9072 trials in total. Trials with an activation time greater than two *SD*s from the mean were deemed outliers and removed. Using this criterion, 128 out of 9072 trials (1%) were removed.

5 RESULTS

Four three-way, repeated-measures ANOVAs were executed; one per dependent variable. Subjective measures were collected after the end of both experiments, which consisted of a ranking of the pointing methods. Mental- and physical-workload, and comfort rating scales were collected, which ranged from 1 as the lowest value to 10 as the highest value. These measures were analyzed with an omnibus Friedman test followed by post hoc Wilcoxon signed-rank tests. All post hoc tests were submitted to a Bonferroni correction to mitigate type-I errors.

Time to Activate

The grand mean for time to activate per trial was 785 ms. The mean for time to activate per trial for mouse pointing was 697 ms, followed by head pointing at 721 ms, then foot pointing at 823 ms, and finally gaze pointing at 898 ms. In terms of target amplitude, the 160-pixel condition (M = 708 ms) yielded lower mean times to activate relative to the 240-pixel condition (M = 861 ms). Regarding the target-width factor, the 75-pixel condition (M = 719 ms) yielded lower mean times to activate relative to the 50-pixel condition (M = 851 ms; see Figure 3).

To check if the mean time to activate differed as a function of target orientation, a polar plot was created. The 0-degree target orientation was excluded from the analysis (see Figure 4). The gaze-pointing conditions exhibited an upper hemifield bias such that the mean time to activate was higher at the lower-target orientations relative to the lateral-target orientations. The other pointing methods yielded no orientation dependence of movement times.

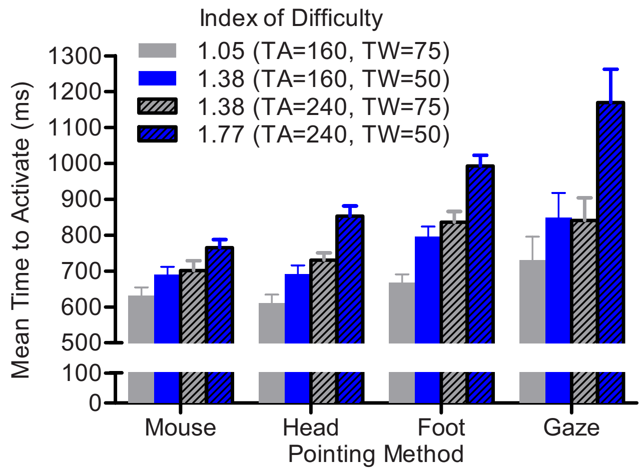

Figure 3: Time to Activate (ms) by pointing method and index of difficulty. Error bars denote one standard error of the mean.

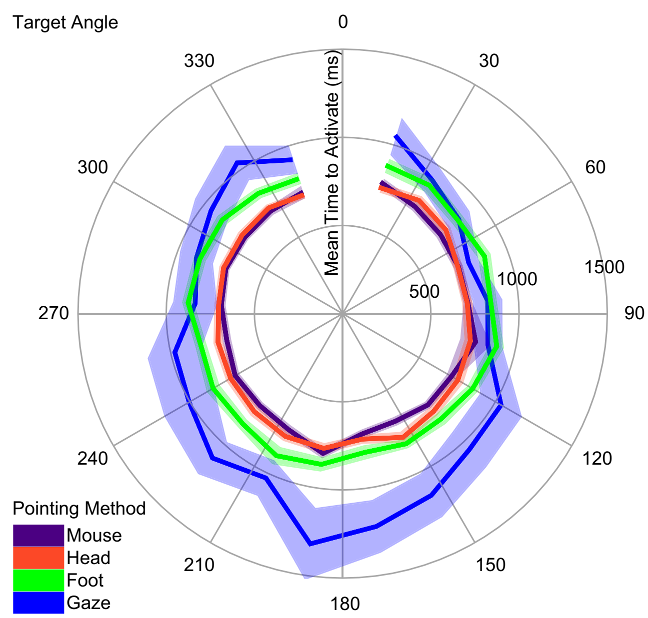

Figure 4: Polar plot showing the distribution of the mean time to activate according to target orientation. One standard error of the mean is depicted with a shaded area around the mean.

The main effects of pointing method, target amplitude, and target width were statistically significant, $F(3, 78) = 8.95$, $p = .00004$, $F(1, 26) = 123.96$, $p < .00001$, $F(1, 26) = 92.24$, $p = .00001$, respectively. The interactions of pointing method by target amplitude, pointing method by target width, target amplitude by target width were statistically significant, $F(3, 78) = 4.57$, $p = .005$, $F(3, 78) = 7.98$, $p = .0001$, $F(3, 78) = 14.43$,

p = .0008, respectively. The third-order interaction between pointing method, target amplitude and target width was statistically significant and qualified all the main effects and lower-order interactions, $F(3, 78) = 6.74$, $p = .0004$.

Throughput

The grand mean for throughput was 3.10 bits/s. The mean throughput for mouse pointing was 3.87 bits/s, followed by head-position pointing at 3.40 bits/s, then foot-mouse pointing at 2.58 bits/s, and finally gaze pointing at 2.55 bits/s. In terms of target amplitude, the 160-pixel condition (M = 3.16 bits/s) yielded a higher mean throughput relative to the 240-pixel condition (M = 3.04 bits/s). Regarding target width, the 75-pixel condition (M = 3.22 bits/s) yielded higher a mean throughput relative to the 50-pixel condition (M = 2.98 bits/s; see Figure 5).

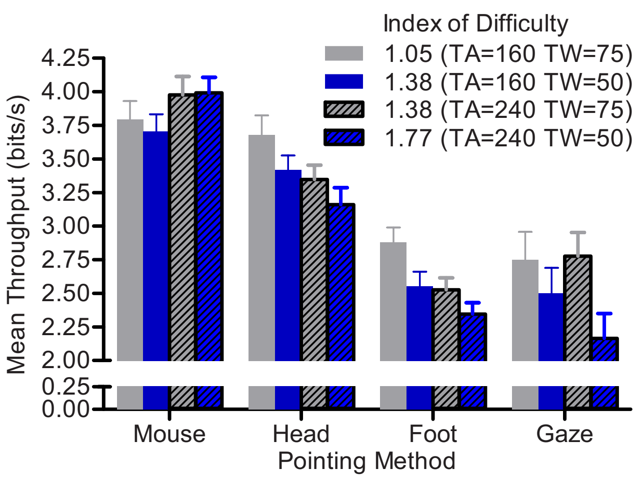

Figure 5: Throughput (bits/s) by pointing method and target amplitude. Error bars denote one standard error of the mean.

As the target amplitude decreased from 240 to 160 pixels, the mean throughput increased; however, this pattern significantly reversed only for the mouse pointing condition. Similarly, as the target width increased from 50 to 75 pixels, the mean throughput increased; however, this pattern was not significantly different for the mouse pointing condition.

The main effects of pointing method, target amplitude, and target width were statistically significant, $F(3, 78) = 57.913$, $p < .00001$, $F(1, 26) = 6.94$, $p < .014$, $F(1, 26) = 39.74$, $p < .00001$, respectively. These main effects were qualified by significant interactions between pointing method and target amplitude and between pointing method and target width, $F(3, 78) = 3.98$, $p = .011$, and $F(3, 78) = 3.83$, $p = .013$, respectively.

Effective Target Width

The grand mean for effective target width was 50.9 pixels. The mean effective target width for mouse pointing was 35.4 pixels, followed by head-position pointing at 46.1 pixels, then gaze pointing at 60.9 pixels, and finally foot-mouse pointing at 61.4 pixels. In terms of target amplitude, the 160-pixel condition (M = 48.2 pixels) yielded smaller mean effective target width relative to the 240-pixel condition (M = 53.7 pixels). Regarding target width, the 75-pixel condition (M = 55.5 pixels) yielded larger mean effective target width relative to the 50-pixel condition (M = 46.4 pixels; see Figure 6).

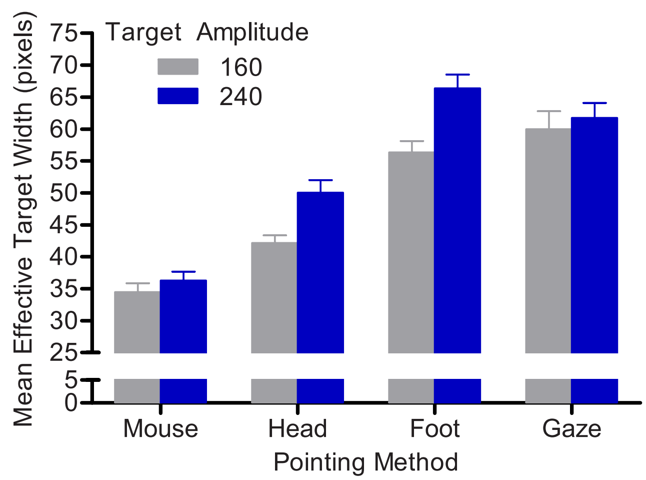

Figure 6: Effective Target Width by pointing method and target amplitude. Error bars denote one standard error of the mean.

The main effects of pointing method, target amplitude, and target width were statistically significant, $F(3, 78) = 42.45$, $p < .00001$, $F(1, 26) = 29.21$, $p < .00001$, $F(1, 26) = 75.24$, $p < .00001$, respectively. However, these main effects were qualified by a significant interaction between pointing method and target amplitude, $F(3, 78) = 4.76$, $p = .004$.

Pupil Dilation

The last dependent variable considered was pupil dilation. The pupil diameter, as seen by the eye tracker camera, was recorded as frames of pixel values for each eye independently. A confidence parameter was also supplied. An estimated 3D modelled pupil diameter (mm), also supplied by the eye tracker software, was not used as it had very low, sometimes negative, correlation between left and right eye (mean value 0.320), indicating the model was not reliably fitting our experimental setup.

All data frames that had a confidence lower than or equal to 0.6 were discarded (as recommended by the vendor). This results in, on average, approximately 55% of the frames being discarded. Next, the Pearson R correlation between left and right eye pupil diameter was calculated for all participants. The mean R across all participants was 0.753. Participants were then included only if (i) the Pearson R between eyes for the participant was larger than 0.5, (ii) the ratio of valid frames to all frames in each test sequence was at least 25%, and (iii) all 16 test sequences were completed with recorded eye tracking data. This left 13 participants.

For the subsequent pupillary analysis, the average of left and right eye pupil diameter was used. Two additional metrics were calculated as follows: When a participant did a calibration at the beginning and possibly one or more verification rounds during the test, a baseline pupil diameter was established. This value was subtracted from all pupil diameter measurements, independently for each participant, resulting in a pupil dilation vs. calibration (also in pixels). In addition, the pupil dilation vs block mean was also calculated by subtracting the mean of each block of sequences of trials from all pupil diameter readings within the particular block.

The grand mean of the pupil diameter was 115.0 pixels, corresponding to an approximate 3D modelled pupil diameter of 6.3 mm. The grand mean of the pupil dilation vs. calibration was 29.9 pixels, corresponding to an approximate 3D modelled pupil dilation of 1.9 mm.

The mean pupil dilation vs. calibration for each of the four pointing methods was 27.5 pixels for the gaze condition, followed by 29.5 pixels for the head condition, 30.3 pixels for the mouse condition, and 32.2 for the foot condition (see Figure 7).

Based on a mixed effects Type III ANOVA analysis with Satterthwaite's method, the pointing method was the only statistically significant effect $F(3, 180) = 7.26$, $p = 0.0001$. A post-hoc Wilcoxon Signed Rank test with Bonferroni correction showed that the foot pointing method was significantly different from gaze $W = 180$, $p < 0.0001$ and head $W = 240$, $p = 0.0003$.

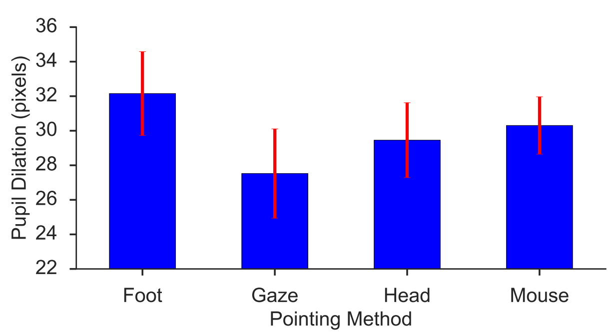

Figure 7: Pupil dilation vs. calibration (pixels) by pointing method. Error bars denote one standard error of the mean.

The mean value of the pupil diameter by trial within each sequence varies over time, as shown in Figure 8. The pupil on average over all participants and conditions increased over the first approximately four trials from an initial value of 107.1 pixels to a mean value over the last 17 trials of 116.2.

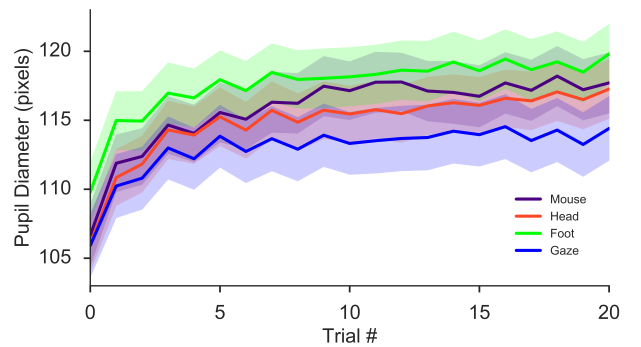

Figure 8: Mean pupil diameter (pixels) by trial index for each pointing method. The shaded area denote one standard error of the mean.

The pupil dilation vs block mean aggregated over all blocks and participants for each of the four combinations of target amplitude and width is shown in Figure 9. The highest value is for the (target amplitude, target width) condition of (160, 50) at 0.81 pixels whereas the lowest value is for (240, 75) at -0.79. Despite the small difference, and the fact that target amplitude and target width did not have a significant effect on the aggregated pupil dilation vs. calibration, a multiple comparison of the pupil dilation vs. block mean using a Wilcoxon Signed Rank test with Bonferroni correction shows that the conditions of (160, 50) and (240, 75) are significantly different, $W = 360$, $p < 0.0164$.

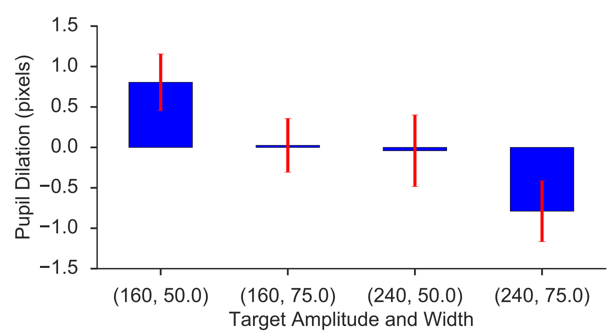

Figure 9: Pupil dilation vs. block mean (pixels) by (target amplitude, target width). Error bars denote one standard error of the mean.

Subjective Ratings

The Friedman test on the pointing method ranking was significant, $\chi^2(3) = 19.52$, $p = .0002$. Participants preferred mouse pointing over the gaze-, head-, and foot- pointing methods, however this difference was only significant between the head-pointing and mouse and the foot- and mouse- pointing methods, $Z = -1.39$, $p = .0049$ and $Z = -2.82$, $p = .00001$, respectively. No other differences were significant regarding the pointing method rankings, ps > .10.

The Friedman test on the mental workload ratings was significant, $\chi^2(3) = 25.04$, $p = .00001$. Participants rated gaze pointing as the most mentally demanding, followed by the foot-mouse, then head-position; finally, the mouse pointing method was rated the least mentally demanding. However, the gaze pointing method was only significantly different from the mouse and head-position pointing methods, $Z = -3.49$, $p = .0005$ and $Z = 3.19$, $p = .0014$. The foot-mouse pointing method was significantly more mental demanding than the head-position and mouse pointing methods, $Z = -3.25$, $p = .0012$ and $Z = -3.53$, $p = .0004$. No other differences were significant regarding the pointing method ratings relating to mental workload, ps > .10.

The Friedman test on the physical workload ratings was significant, $\chi^2(3) = 22.49$, $p = .00005$. Participants rated foot pointing as the most physically demanding, followed by head, then gaze; finally, the mouse pointing method was rated the least physically demanding. The foot pointing method was significantly different from the mouse pointing method, $Z = -4.18$, $p = .00003$. The head pointing method was significantly different from the mouse pointing method, $Z = -3.85$, $p = .0001$. Finally, gaze pointing was rated more physically demanding than the mouse pointing method, $Z = -2.71$, $p = .01$. No other differences were significant regarding the pointing method ratings relating to physical workload, ps > .10.

The Friedman test on the comfort ratings was non-significant, $p > .30$. As a result, no post hoc analyses were executed on the level of comfort for each pointing method device.

User Comments

Participants' responses regarding the different pointing methods were summarized and reviewed for qualitative patterns. Overall, when participants rated the pointing methods in terms of most preferred, mouse pointing was in first place. Comments include, *"It was easy"* ($n = 8$) and *"I have lots of experience with it"* ($n = 6$). Participants that rated gaze pointing as the most preferred said, *"It was the easiest"* ($n = 7$), *"It was quick"* ($n = 2$), and *"It's the least demanding"* ($n = 3$). Participants that rated head-position pointing as their most preferred mentioned that, *"It was easy"* ($n = 3$), and *"It was natural"* ($n = 1$).

When participants rated the pointing methods in terms of their least preferred, they rated the foot-mouse pointing method as the least preferred and said, *"It was mentally demanding"* ($n = 4$) and *"It was too tiresome"* ($n = 5$). Participants that rated head-position pointing as their least preferred said, *"It was uncomfortable"* ($n = 5$) and *"It was annoying"* ($n = 2$). Participants who rated gaze pointing as their least preferred said it had *"bad calibration"* ($n = 2$), that *"the accuracy was off"* ($n = 8$), and that *"it was uncomfortable"* ($n = 3$).

6 DISCUSSION

Throughput values in the current experiment suggest head pointing is an efficient input method, with just 12% less throughput than the mouse. Gaze throughput was 34% lower - and so was foot input. In addition, effective target width and time to activate were better for head than gaze input and users rated gaze more mentally demanding than head input.

Our findings confirm the findings from Qian et al. [37] and Hansen et al. [13] where they reported that head input outperforms gaze. Blattgerste et al. [2], however, reported that gaze was much faster than head input. They suggest their results might be different from Qian et al.'s [37] because they used a more precise gaze tracking system. We note that our work and work by Qian et al. [37] and Hansen et al. [13] compared gaze and head inputs in a Fitts' law evaluation that conforms to ISO 9241-9. However, Blattgerste et al. [2] used point-and-click tasks on a keyboard and a menu, also the field-of-view was varied in the experiment. Since this was not a standard Fitts' law task, the constraints of Fitts' law experiment were not met.

The gaze throughput of the HTC VIVE HMD reported herein was 2.6 bits/s. The throughput of the FOVE HMD tested by [37] and [13], was 1.7 bits/s and 2.1 bits/s, respectively. There may be several reasons for the differences observed. One study [37] used a 3D environment while the other [13] used a standard 2D interface. However, the head throughput for these two studies were almost identical, 2.4 bits/s and 2.5 bits/s, respectively. This suggests that 3D vs. 2D is of importance for gaze but not for head interaction. The difference between the gaze throughput reported herein and the throughput found previously [13] is most likely due to the equipment used, since they had identical 2D set-ups. Future study is needed, for example, comparing 2D vs. 3D interfaces, motion vs. static background and targets in head space vs. targets in world space. For instance, Rajanna et al. [40] reported that motion in the background decreased gaze typing performance on an overlay keyboard in a FOVE HMD. Further study, using a standard test procedure, may clarify the impact of motion and with types of HMDs, for instance VR vs. AR.

The system we tested was particularly slow in the lower hemisphere (see Figure 4), while Hansen et al. [13] reported the FOVE HMD to have the same movement time in all directions. The gaze tracking system in the present experiment may solve this issue with improvements in software or hardware. With the current version of this system, our observations suggest placing UI elements on the horizontal plane or in the upper hemisphere when pointing time is a priority.

Foot input is an option when seated. This input method performed similar to gaze for most of the performance measures, and user ratings did not differ between the two. However, this input method requires an extra device, while both head and gaze tracking are internal to the HMD. Gaze tracking, on the other hand, may require a calibration process that can be difficult for people with cognitive challenges to perform [32].

Pupil results were contrary to our expectations in three ways. First, there was no simple relation between pupil dilation and the index of difficulty, see Figure 9. The only significant difference we found was between the condition with two small targets close together and two larger targets farther away from each other. The first condition yielded the highest pupil dilation, indicative of being the most difficult, while the last condition yielded a smaller pupil dilation, compared to a baseline of the block mean.

Secondly, foot input was associated with larger dilations than gaze input, even though the two conditions were similar in most other measures. This might suggest that it is the higher physical effort needed when moving the feet, as compared to the ease of moving the eyes that caused the extra dilation. Another reason might be that pointing with gaze is a natural activity that we do all the time, when directing our visual attention to an area in the environment. Pointing with the feet by use of a balance board is new to most people. Thus it requires extra effort.

Thirdly, Figure 8 shows that for every new sequence encountered, our subjects exhibited an increase in pupil diameter for the first trials in a sequence. This *start-up effect* was found independently of the input method. Twelve of 13 participants examined showed clear signs of a start-up dilation from a lower initial value. It is not likely to be caused by changes in luminescence since the target color appearing at the start were equiluminant with the background seen before the onset of the task sequence.

HMDs may be attractive for a first assessments of gaze interaction since the cost of quality gaze communication systems is higher than an HMD with built-in gaze tracking and we expect that these HMDs soon may be acquired from several vendors. Basic issues of whether a given individual has the capability to perform a calibration or posses the eye motor control needed to maintain a fixation may then be clarified with an off-the-shelf HMD. This raises an important question of how performance measures on commodity hardware generalize to real life task situations when, for example, using a computer or driving a wheelchair. Rajanna et al. [40] found gaze typing in HMDs to be around 9.4 wpm (dwell 550 ms) for first time users, which compares well to performance of novice users of remote gaze trackers for on-screen qwerty-keyboards (5 to 10 wpm for dwell times 1000 to 450 ms) [31] but further research needs to clarify if this is also the case for the Fitts' law task and for individuals with disabilities.

How may pupil data be applied when assessing the pointing capabilities for a given individual? Our data on differences in pupil dilation between the various inputs are not yet conclusive, since it is an open issue whether cognitive vs. physical effort has a dominant effect on the difference between foot and gaze pointing. The start-up effect is consistent for all four input methods. Further research might explore if reduced levels of cortical activity, for instance due to tiredness, medication, or depression, would impact the start-up effect. If so, when a start-up effect is not found, low cortical activity may be taken into consideration when accessing an individual. The pupil diameter measures, however, should be more robust than what we observed in our study, since only 13 of 27 participants provided stable pupil data throughout the experiment.

When the ISO-9241-9 procedure was applied by Felzer et al. [9], the user from the target group had a throughput less than half of the non-disabled participants. Hence, we might not expect performance levels within the average when accessing an individual but rather look for his or her best performance. In addition, it should of course be considered if the highest efficiency was achieved with a device that would otherwise be effective for the user Additional questions remain: Can the HMD be used with glasses? Do caregivers know how to operate the device?

7 CONCLUSION

In conclusion, for the hands-free pointing methods tested in an HMD using ISO-9241-9, we found head input outperforms gaze input on throughput, effective target width, and time to activate, while gaze performed similar to foot input. Pupil dilations were consistently associated with the onset of a task, independent of the pointing method. Further research, including testing with individuals with motor challenges, should clarify the external validity of assessments conducted with the ISO-9241-9 methodology and HMDs.

REFERENCES

[1] Jackson Beatty and Brennis Lucero-Wagoner. 2000. The pupillary system. *Handbook of psychophysiology* 2 (2000), 142–162.

[2] Jonas Blattgerste, Patrick Renner, and Thies Pfeiffer. 2018. Advantages of eye-gaze over head-gaze-based selection in virtual and augmented

reality under varying field of views. In *Proceedings of the Workshop on Communication by Gaze Interaction*. ACM, 1.

[3] Catherine M Capio, Cindy HP Sit, Bruce Abernethy, W Erickson, C Lee, and S von Schrader. 2018. 2016 Disability Status Report: United States. *Ithaca, NY: Cornell University Yang Tan Institute on Employment and Disability(YTI)* (2018).

[4] Gamhewage C De Silva, Michael J Lyons, Shinjiro Kawato, and Nobuji Tetsutani. 2003. Human factors evaluation of a vision-based facial gesture interface. In *Computer Vision and Pattern Recognition Workshop, 2003. CVPRW'03. Conference on*, Vol. 5. IEEE, 52–52.

[5] David Dearman, Amy Karlson, Brian Meyers, and Ben Bederson. 2010. Multi-modal Text Entry and Selection on a Mobile Device. In *Proceedings of Graphics Interface 2010 (GI '10)*. Canadian Information Processing Society, Toronto, Ont., Canada, Canada, 19–26. http://dl.acm.org/citation.cfm?id=1839214.1839219

[6] Giuseppe Di Gironimo, Giovanna Matrone, Andrea Tarallo, Michele Trotta, and Antonio Lanzotti. 2013. A virtual reality approach for usability assessment: case study on a wheelchair-mounted robot manipulator. *Engineering with Computers* 29, 3 (2013), 359–373.

[7] Alistair Edwards. 2018. Accessibility. *The Wiley Handbook of Human Computer Interaction* 2 (2018), 681–695.

[8] Monika Elepfandt and Martin Grund. 2012. Move it there, or not?: the design of voice commands for gaze with speech. In *Proceedings of the 4th workshop on eye gaze in intelligent human machine interaction*. ACM, 12.

[9] Torsten Felzer, Ian Scott MacKenzie, and John Magee. 2016. Comparison of Two Methods to Control the Mouse Using a Keypad. In *International Conference on Computers Helping People with Special Needs*. Springer, 511–518.

[10] Kingsley Fletcher, Andrew Neal, and Gillian Yeo. 2017. The effect of motor task precision on pupil diameter. *Applied ergonomics* 65 (2017), 309–315.

[11] Vicki A Freedman, Emily M Agree, Linda G Martin, and Jennifer C Cornman. 2006. Trends in the use of assistive technology and personal care for late-life disability, 1992–2001. *The Gerontologist* 46, 1 (2006), 124–127.

[12] Fabian Göbel, Konstantin Klamka, Andreas Siegel, Stefan Vogt, Sophie Stellmach, and Raimund Dachselt. 2013. Gaze-supported Foot Interaction in Zoomable Information Spaces (Interactivity). In *Proceedings of the Conference on Human Factors in Computing Systems - Extended Abstracts*. ACM, 4.

[13] John Paulin Hansen, Vijay Rajanna, I. Scott MacKenzie, and Per Bækgaard. 2018. A Fitts' Law Study of Click and Dwell Interaction by Gaze, Head and Mouse with a Head-Mounted Display. In *COGAIN '18: Workshop on Communication by Gaze Interaction, June 14–17, 2018, Warsaw, Poland (COGAIN '18)*. ACM, New York, NY, USA. https://doi.org/10.1145/3206343.3206344

[14] John Paulin Hansen, Astrid Kofoed Trudslev, Sara Amdi Harild, Alexandre Alapetite, and Katsumi Minakata. 2019, submitted. Providing access to VR through a wheelchair. In *Paper submitted as case study to CHI2019*.

[15] C Harrison, Phillipa Dall, PM Grant, MH Granat, TW Maver, and BA Conway. 2000. Development of a wheelchair virtual reality platform for use in evaluating wheelchair access. In *3rd International Conference on Disability, VR and Associated Technologies, Sardinia, Edited by P. Sharkey*.

[16] Benjamin Hatscher, Maria Luz, Lennart E. Nacke, Norbert Elkmann, Veit Müller, and Christian Hansen. 2017. GazeTap: Towards Hands-free Interaction in the Operating Room. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction (ICMI 2017)*. ACM, New York, NY, USA, 243–251. https://doi.org/10.1145/3136755.3136759

[17] Eckhard H Hess and James M Polt. 1964. Pupil Size in Relation to Mental Activity during Simple Problem-Solving. *Science* 143, 3611 (1964), 1190–1192. http://www.jstor.org.proxy.findit.dtu.dk/stable/1712692

[18] Shamsi T. Iqbal, Xianjun Sam Zheng, and Brian P. Bailey. 2004. Task-evoked Pupillary Response to Mental Workload in Human-computer Interaction. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems (CHI EA '04)*. ACM, New York, NY, USA, 1477–1480. https://doi.org/10.1145/985921.986094

[19] Shinobu Ishihara. 1972. *Tests for Colour-blindness*. Kannehara Shuppan Co, Ltd., Tokyo, Kyoto, Japan.

[20] ISO. 2000. *Ergonomic requirements for office work with visual display terminals (VDTs) - Part 9: Requirements for non-keyboard input devices (ISO 9241-9)*. Technical Report Report Number ISO/TC 159/SC4/WG3 N147. International Organisation for Standardisation.

[21] Xianta Jiang, M Stella Atkins, Geoffrey Tien, Bin Zheng, and Roman Bednarik. 2014. Pupil dilations during target-pointing respect Fitts' law. In *Proceedings of the Symposium on Eye Tracking Research and Applications*. ACM, 175–182.

[22] Xianta Jiang, Bin Zheng, Roman Bednarik, and M Stella Atkins. 2015. Pupil responses to continuous aiming movements. *International Journal of Human-Computer Studies* 83 (2015), 1–11.

[23] Daniel Kahneman and Jackson Beatty. 1966. Pupil Diameter and Load on Memory. *Science* 154, 3756 (1966), 1583–1585. http://www.jstor.org.proxy.findit.dtu.dk/stable/1720478

[24] Shaun K Kane, Jeffrey P Bigham, and Jacob O Wobbrock. 2008. Slide rule: making mobile touch screens accessible to blind people using multi-touch interaction techniques. In *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 73–80.

[25] S Keates, PJ Clarkson, and P Robinson. 1998. Developing a methodology for the design of accessible interfaces. In *Proceedings of the 4th ERCIM Workshop*. 1–15.

[26] Simeon Keates, Faustina Hwang, Patrick Langdon, P John Clarkson, and Peter Robinson. 2002. Cursor measures for motion-impaired computer users. In *Proceedings of the fifth international ACM conference on Assistive technologies*. ACM, 135–142.

[27] B. Laeng, S. Sirois, and G. Gredeback. 2012. Pupillometry: A Window to the Preconscious? *Perspectives on Psychological Science* 7, 1 (2012), 18–27. https://doi.org/10.1177/1745691611427305

[28] I. Scott MacKenzie. 2018. Fitts' law. In *Handbook of Human Computer Interaction*, K. L. Norman and J. Kirakowski (Eds.). Wiley, NJ, 349–370.

[29] John Magee, Torsten Felzer, and I Scott MacKenzie. 2015. Camera Mouse+ ClickerAID: Dwell vs. single-muscle click actuation in mouse-replacement interfaces. In *International Conference on Universal Access in Human-Computer Interaction*. Springer, 74–84.

[30] Päivi Majaranta, Ulla-Kaija Ahola, and Oleg Špakov. 2009. Fast Gaze Typing with an Adjustable Dwell Time. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA, 357–360. https://doi.org/10.1145/1518701.1518758

[31] Päivi Majaranta, Ulla-Kaija Ahola, and Oleg Špakov. 2009. Fast Gaze Typing with an Adjustable Dwell Time. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA, 357–360. https://doi.org/10.1145/1518701.1518758

[32] Päivi (Ed. Majaranta. 2011. *Gaze Interaction and Applications of Eye Tracking: Advances in Assistive Technologies: Advances in Assistive Technologies*. IGI Global.

[33] Denys Matthies, Franz Müller, Christoph Anthes, and Dieter Kranzlmüller. 2014. ShoeSoleSense: demonstrating a wearable foot interface for locomotion in virtual environments. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. ACM, 183–184.

[34] Mark R Mine. 1995. Virtual environment interaction techniques. *UNC Chapel Hill CS Dept* (1995).

[35] Toni Pakkanen and Roope Raisamo. 2004. Appropriateness of foot interaction for non-accurate spatial tasks. In *CHI'04 extended abstracts on Human factors in computing systems*. ACM, 1123–1126.

[36] Glenn Pearson and Mark Weiser. 1986. Of moles and men: the design of foot controls for workstations. In *ACM SIGCHI Bulletin*, Vol. 17. ACM, 333–339.

[37] Yuan Yuan Qian and Robert J Teather. 2017. The eyes don’t have it: an empirical comparison of head-based and eye-based selection in virtual reality. In *Proceedings of the 5th Symposium on Spatial User Interaction*. ACM, 91–98.

[38] Vijay Rajanna. 2016. Gaze Typing Through Foot-Operated Wearable Device. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '16)*. ACM, New York, NY, USA, 345–346. https://doi.org/10.1145/2982142.2982145

[39] Vijay Rajanna and Tracy Hammond. 2016. GAWSCHI: Gaze-augmented, Wearable-supplemented Computer-human Interaction. In *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research and Applications (ETRA '16)*. ACM, New York, NY, USA, 233–236. https://doi.org/10.1145/2857491.2857499

[40] Vijay Rajanna and John Paulin Hansen. 2018. Gaze Typing in Virtual Reality: Impact of Keyboard Design, Selection Method, and Motion. In *Proceedings of the Tenth Biennial ACM Symposium on Eye Tracking Research and Applications (ETRA '18)*. ACM, New York, NY, USA. https://doi.org/10.1145/3204493.3204541

[41] Francois Richer and Jackson Beatty. 1985. Pupillary dilations in movement preparation and execution. *Psychophysiology* 22, 2 (1985), 204–207.

[42] Maria Francesca Roig-MaimÃş, I. Scott MacKenzie, Cristina Manresa-Yee, and Javier Varona. 2018. Head-tracking interfaces on mobile devices: Evaluation using FittsâĂŹ law and a new multi-directional corner task for small displays. *International Journal of Human-Computer Studies* 112 (2018), 1 – 15. https://doi.org/10.1016/j.ijhcs.2017.12.003

[43] William R Sherman and Alan B Craig. 2002. *Understanding virtual reality: Interface, application, and design*. Elsevier.

[44] R William Soukoreff and I Scott MacKenzie. 2004. Towards a standard for pointing device evaluation, perspectives on 27 years of FittsâĂŹ law research in HCI. *International journal of human-computer studies* 61, 6 (2004), 751–789.

[45] Robert F Stanners, Michelle Coulter, Allen W Sweet, and Philip Murphy. 1979. The pupillary response as an indicator of arousal and cognition. *Motivation and Emotion* 3, 4 (1979), 319–340. https://doi.org/10.1007/BF00994048

[46] Eduardo Velloso, Dominik Schmidt, Jason Alexander, Hans Gellersen, and Andreas Bulling. 2015. The Feet in Human–Computer Interaction: A Survey of Foot-Based Interaction. *ACM Comput. Surv.* 48, 2, Article 21 (2015), 35 pages. https://doi.org/10.1145/2816455

[47] Xuan Zhang and I Scott MacKenzie. 2007. Evaluating eye tracking with ISO 9241-part 9. In *International Conference on Human-Computer Interaction*. Springer, 779–788.

[48] Kathryn Ziegler-Graham, Ellen J MacKenzie, Patti L Ephraim, Thomas G Travison, and Ron Brookmeyer. 2008. Estimating the prevalence of limb loss in the United States: 2005 to 2050. *Archives of physical medicine and rehabilitation* 89, 3 (2008), 422–429.