

# A Fitts' Law Study of Click and Dwell Interaction by Gaze, Head and Mouse with a Head-Mounted Display

John Paulin Hansen

Technical University of Denmark, Kgs. Lyngby, Denmark  
jpha@dtu.dk

I. Scott MacKenzie

York University, Toronto, Ontario, Canada  
mack@cse.yorku.ca

Vijay Rajanna

Texas A&M University, College Station, Texas, USA  
vijay.drajanna@gmail.com

Per Bækgaard

Technical University of Denmark, Kgs. Lyngby, Denmark  
pgba@dtu.dk

## ABSTRACT

Gaze and head tracking, or pointing, in head-mounted displays enables new input modalities for point-select tasks. We conducted a Fitts' law experiment with 41 subjects comparing head pointing and gaze pointing using a 300 ms dwell ( $n = 22$ ) or click ( $n = 19$ ) activation, with mouse input providing a baseline for both conditions. Gaze and head pointing were equally fast but slower than the mouse; dwell activation was faster than click activation. Throughput was highest for the mouse (3.24 bits/s), followed by head pointing (2.47 bits/s) and gaze pointing (2.13 bits/s). Dwell activation was faster than click activation. The effective target width for gaze ( $\approx 94$  pixels; about  $3^\circ$ ) was larger than for head and mouse ( $\approx 72$  pixels; about  $2.5^\circ$ ). Subjective feedback rated the physical workload less for gaze pointing than head pointing.

## CCS CONCEPTS

- Human-centered computing → Pointing devices;

## KEYWORDS

Fitts' law; ISO 9241-9; gaze interaction; head interaction; dwell activation; head mounted displays.

### ACM Reference format:

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 *Proceedings of Workshop on Communication by Gaze Interaction, Warsaw, Poland, June 14–17, 2018 (COGAIN '18)*, 5 pages.

<https://doi.org/10.1145/3206343.3206344>

## 1 INTRODUCTION

Several developers of eye-wear digital devices, including FaceBook, Google and Apple, have recently acquired companies specialising in gaze tracking, with the result that gaze and head tracking

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.

COGAIN '18, June 14–17, 2018, Warsaw, Poland

© 2018 Association for Computing Machinery.

ACM ISBN 978-1-4503-5790-6/18/06...\$15.00

<https://doi.org/10.1145/3206343.3206344>

are now both integrated in commodity headsets for virtual reality (VR). There are mainly two reasons for this: (1) substantial processing-power in headsets may be saved using gaze-contingent rendering where the full image is only shown at the current fixation point [Murphy and Duchowski 2001; Reingold et al. 2003], and (2) gaze may serve as a hands-free pointer for effortless interaction with head-mounted displays (HMD) [Jalaliniya et al. 2015]. Dwell selection has become the preferred selection method for "eyes only" input. This eliminates the need for a hand controller, reducing cost and complexity of the systems. However, if dwell selection is less efficient than click selection, then a physical trigger may be required. People with motor disabilities could use an HMD for interaction when a remote tracking setup is not possible, for instance while lying in bed or sitting in a car. If HMDs become a successful product with head- and gaze-tracking, they may offer a low-cost alternative to high-end gaze communication systems.

This paper aims to inform designers of eye-wear applications what to expect from the gaze- and head-interaction options appearing in HMDs. There are number of questions we address. Is dwell selection a viable alternative to click selection? How large should targets be for accurate selection? How fast is target activation? Is it better to use head interaction instead of gaze interaction? Do people experience physical or mental strain with these new pointing methods? The main limitation is that our study only addresses interaction where the pointer symbol (i.e., cursor) is visible at all times, thereby allowing the user to compensate for inaccuracies in gaze tracking through additional head movements. The study is also not addressing interaction in VR since we are using the original "flat" (i.e., 2D) version of the Fitts' law test. Finally, we only use one type of headset.

## 2 PRIOR WORK

Ware et al., presented one of the first evaluations of gaze input in 1987 [Ware and Mikaelian 1987]. Gaze was used for pointing, with selection using either dwell, a physical button, or a software (on-screen) button. Zhang and MacKenzie were the first to conduct a Fitts' law evaluation of gaze input according to the ISO 9241-9 standard for evaluating the performance and comfort of non-keyboard computer devices<sup>1</sup> [Zhang and MacKenzie 2007]. Based on two target widths and two amplitudes, the index of difficulties (*IDs*) varied from 1.9 bits to 2.5 bits.

<sup>1</sup><https://www.iso.org/standard/38896.html> [last accessed - June 8, 2018]

Miniotas further explored Fitts' law in gaze interaction by comparing the performance of an eye tracker and a mouse in a simple pointing task [Miniotas 2000]. The experiment included a wide range of IDs by modifying both the target amplitude and width. The author found that the mean time for selection with gaze was 687 ms and with the mouse 258 ms. Hence, contrary to other findings [Sibert and Jacob 2000; Ware and Mikaelian 1987], gaze-based selection did not outperform the mouse. In a recent study, Qian and Teather used a FOVE HMD with built-in gaze tracking to compare head- and eye-based interaction in VR using a Fitts' law task [Qian and Teather 2017]. Gaze pointing had a significantly higher error rate than head pointing. It also had lower movement time and a lower throughput. Participants rated head input better than gaze on 8 of 9 items on user-experience. Only neck fatigue was considered greater for head pointing than gaze pointing. Their study only used click selections, however.

Dwell-time selection (i.e., looking at a target for a set time) is well examined elsewhere (e.g., [Hansen et al. 2003], [Majaranta et al. 2006], and [Majaranta et al. 2009]). Dwell-time is used as a clutch to avoid unintended selection of an object that was just looked at (the so-called "Midas Touch" problem [Jacob 1990]). According to Majaranta et al., expert users can handle dwell times as low as 282 ms with errors being just 0.36% using an on-screen keyboard [Majaranta et al. 2009]. Pointing with the head has been done for more than 20 years (e.g., [Ballard and Stockman 1995]). A recent study [Yu et al. 2017] used the inertial measurement unit (IMU) in an HMD to type by head pointing. Dwell selection (400 ms) was found less efficient than click selection, while head gestures were fastest.

Gaze plus dwell selection in HMDs has not yet been examined. As previously mentioned, dwell selection is the most common technique and supports hands-free interaction with an HMD - which click selection does not. In the study reported below, we applied a dwell time of 300 ms, which is considered feasible for an expert user. We also used the common principle of enlarging the effective target size once it has been entered (so-called spatial hysteresis) and then measuring the effective target width, the spatial variability (i.e., accuracy) in selection.

### 3 METHOD

#### 3.1 Participants

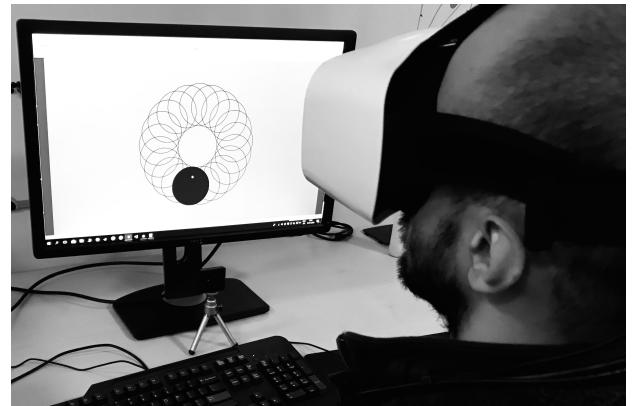
Forty-one participants were recruited on a voluntary basis among visitors at a VR technology exhibition. The mean age was 29 years, ( $SD = 9.7$  years); 31 male, 10 female. A majority (77%) were Danish citizens. Most (91%) had tried HMDs several times before, and 9% only one time before. A few (16%) had previously tried gaze interaction. Vision was normal (uncorrected) for 28 participants, while eight used glasses (which they took off), and five used contact lenses during the experiment.

#### 3.2 Apparatus

A headset from FOVE with built-in gaze tracking was used.<sup>2</sup> The headset has a resolution of  $2560 \times 1440$  px, renders at a maximum of 70 fps, and has a field of view of 100 degrees. The binocular

eye tracking system runs at 120 Hz. Although the manufacturer indicates that tracking accuracy is less than 1 degree of visual angle, in two recent experiments [Rajanna and Hansen 2018], we found the FOVE headset mean accuracy to be around 4 degrees of visual angle. The headset weighs 520 grams and has IR-based position tracking plus IMU-based orientation tracking. A Logitech corded mouse M500 was used for the manual input.

Software to run the experiment was a Unity implementation of the 2D Fitts' law software developed by MacKenzie, known as FittsTaskTwo<sup>3</sup>. The Unity version<sup>4</sup> includes the same features and display as the original; that is, with spherical targets presented on a flat 2D-plane, cf. Figure 1. Based on the findings [Majaranta et al. 2009] a dwell time setting of 300 ms was chosen for the dwell condition.



**Figure 1: Participant wearing a headset with gaze tracking. The monitor shows the Fitts' law task he sees in the headset.**

Using the eyes for selection in an HMD may actually combine head and gaze pointing. When users have difficulty hitting a target by gaze only, they sometimes adjust their head orientation, thereby moving the targets into an area where the gaze tracker may be more accurate. In a remote setup, users are known to do this instinctively when gaze tracking is a bit off-target [Špakov et al. 2014]. We allowed participants to use this strategy by not disengaging head tracking in the gaze-pointing mode.

#### 3.3 Procedure

The participants were asked to sign a consent form and were given a short explanation of the Fitts' law experiment. When seated, the HMD was put on and adjusted for comfort. Then the participants were randomly assigned to do either dwell or click selections. None of the participants did both. Their hand grasped the mouse, which was required for click activation condition. In the dwell condition, the participants only used the mouse for pointing; the mouse button was not used.

Both selection condition groups were exposed to all three pointing methods: gaze pointing, head pointing, and mouse pointing, alternating the order between participants. For each pointing method,

<sup>3</sup>available at <http://www.yorku.ca/mack/FittsLawSoftware/> [last accessed: June 8, 2018]

<sup>4</sup>available at <https://github.com/GazeIT-DTU/FittsLawUnity> [last accessed: June 8, 2018]

four levels of index difficulty *ID* were tested, composed of two target widths (50 pixels, 100 pixels) and two amplitudes (160 pixels, 260 pixels). The target widths span visual angles of about 2° and 4°, respectively. Spatial hysteresis was set to 2.0 for all dwell conditions, meaning that when targets were first entered, they doubled in size, while visually remaining constant. 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 bottom position (6 o'clock). When this target was selected, a target at the opposite side would be highlighted (approximately 12 o'clock), then when activated a target between 6 and 7 o'clock was highlighted and so on, moving clockwise. The first target at 6 o'clock is not included in the data analysis in order to minimize the impact from initial reaction time. The pointer (i.e., cursor) was visible at all times, presented as a yellow dot. For mouse input, the pointer was the mouse position on screen. For eye tracking, this was the gaze position as defined by the intersection of the two gaze vectors from the centre of both eyes on the target plane. For head tracking, the pointer 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. Also during this break, they were given three short Likert-scale (10-point) questions: "How mentally demanding was this task?" "How physically demanding was the task?" and "How comfortable do you feel right now?" Before testing the eye tracking method, participants performed the standard FOVE gaze calibration procedure, and the experimenter visually confirmed that gaze tracking was stable; if not, re-calibration was done. Completing the full experiment took approximately 20 minutes for each participant.

### 3.4 Design

The experiment was a  $3 \times 2 \times 2 \times 2$  design with the following independent variables and levels:

- Pointing method (gaze, head, mouse)
- Selection method (click, dwell)
- Amplitude (160 pixels, 260 pixels)
- Width (50 pixels, 100 pixels)

Pointing method and selection method were the primary independent variables. Amplitude and width were included to ensure the conditions covered a range of task difficulties.

We used a mixed design with pointing method assigned within-subjects and selection method assigned between-subjects. For selection method, there were 19 participants in the click group and 22 participants in the dwell group. For pointing method, the levels were assigned in different orders to offset learning effects.

The dependent variables were time to activate, throughput, and effective target width, calculated according to the standard procedures for ISO 9241-9. For click selection, we also logged the errors (selections with the pointer outside the target). Errors were not possible with dwell selection.

For click activation, there were 228 trial sequences (19 Participants  $\times$  3 Pointing Methods  $\times$  2 Amplitudes  $\times$  2 Widths). For dwell,

there were 264 trial sequences ( $22 \times 3 \times 2 \times 2$ ). For each sequence, 21 trials were performed.

Trials with an activation time greater than two SDs from the mean were deemed outliers and removed. Using this criterion, 120 of 5124 trials (2.3%) were removed for click activation and 129 of 5712 trials (also 2.3%) were removed for dwell activation.

## 4 RESULTS AND DISCUSSION

We performed a two-factor mixed model ANOVA with replication on the dependent variables. The results are shown in Table 1. The ef-

**Table 1: ANOVA of the Fitts' evaluation matrix with three dependent variables: time to activate, throughput, and effective target width. The between-subjects factor was selection method and the within-subjects factor was pointing method [Gz-gaze, Hd-head, Ms-Mouse]. Cells highlighted in gray indicate significance at  $p < 0.05$ .**

|                   | Pointing<br>[Gz, Hd, Ms] | Selection<br>[click, dwell] | Interaction        |
|-------------------|--------------------------|-----------------------------|--------------------|
| <b>TimeTo</b>     | $F(2,324) = 135.362$     | $F(1,162) = 5.376$          | $F(2,324) = 1.83$  |
| <b>-Activate</b>  | $p = 0.000$              | $p = 0.022$                 | $p = 0.162$        |
| <b>Through</b>    | $F(2,324) = 125.024$     | $F(1,162) = 2.325$          | $F(2,324) = 1.031$ |
| <b>-put</b>       | $p = 0.000$              | $p = 0.129$                 | $p = 0.358$        |
| <b>Eff-Target</b> | $F(2,324) = 31.803$      | $F(1,162) = 8.191$          | $F(2,324) = 1.292$ |
| <b>-Width</b>     | $p = 0.000$              | $p = 0.005$                 | $p = 0.276$        |

fect of pointing method was statistically significant for all the three dependent variables. The effect of selection method was statistically significant for time to activate and effective target width. No interaction effects were found. The results for each dependent variable are shown in Figure 2 (Time To Activate), Figure 3 (Throughput), and Figure 4 (Effective Target Width).

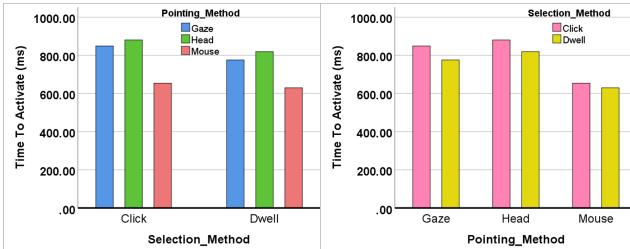
Table 2 gives the results of a post hoc pairwise analysis using the Bonferroni correction. For time to activate, the differences between the mouse and the other methods were statistically significant. For throughput, all pairwise differences were statistically significant for pointing method. The throughput value for the mouse was 3.239 bits/s. This is low compared to other ISO-conforming mouse evaluations, where throughput is typically in the range of 3.7 to 4.9 bits/s [Soukoreff and MacKenzie 2004]. The most likely explanation is that participants in our user study wore an HMD to view the target scene and therefore could not see the mouse or their hand in their peripheral vision. System lag and tracking accuracy are other possible issues.

For effective target width, the differences between gaze and other methods were statistically significant. Lastly, a one-way repeated measures ANOVA revealed that pointing method had a significant effect on error rate for click activation ( $F(2, 150) = 12.49, p < 0.05$ ). Also, from a post hoc analysis with the Bonferroni correction, we found a significant difference between gaze pointing and head pointing ( $p < 0.05$ ), and gaze pointing and the mouse ( $p < 0.05$ ).

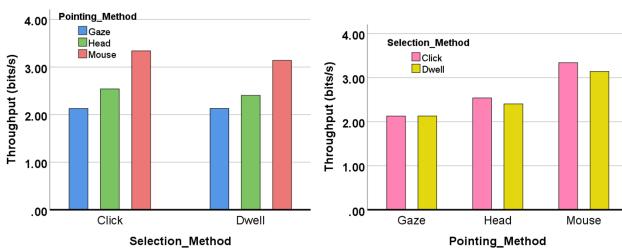
However, there was no significant difference between head pointing and the mouse ( $p > 0.05$ ). The error rates were 1.53% (3.35%) for gaze pointing, 0.26% (0.89%) for head pointing, and 0.06% (0.24%) for the mouse.

**Table 2: ANOVA of Fitts’ evaluation matrix with post hoc analysis [Gz-gaze, Hd-head, Ms-Mouse]. Cells highlighted in gray indicate significance at  $p < 0.05$ .**

| Pointing                | Mean         | SE     | Post hoc Analysis    |
|-------------------------|--------------|--------|----------------------|
| <b>TimeTo Activate</b>  | Gz = 812.190 | 16.471 | (Gz, Hd) $p = 0.053$ |
|                         | Hd = 849.951 | 14.432 | (Gz, Ms) $p = 0.000$ |
|                         | Ms = 641.750 | 9.528  | (Hd, Ms) $p = 0.000$ |
| <b>Throughput</b>       | Gz = 2.127   | 0.068  | (Gz, Hd) $p = 0.000$ |
|                         | Hd = 2.472   | 0.048  | (Gz, Ms) $p = 0.000$ |
|                         | Ms = 3.239   | 0.048  | (Hd, Ms) $p = 0.000$ |
| <b>Eff-Target Width</b> | Gz = 93.765  | 2.733  | (Gz, Hd) $p = 0.000$ |
|                         | Hd = 71.794  | 2.564  | (Gz, Ms) $p = 0.000$ |
|                         | Ms = 72.420  | 2.711  | (Hd, Ms) $p = 1.000$ |

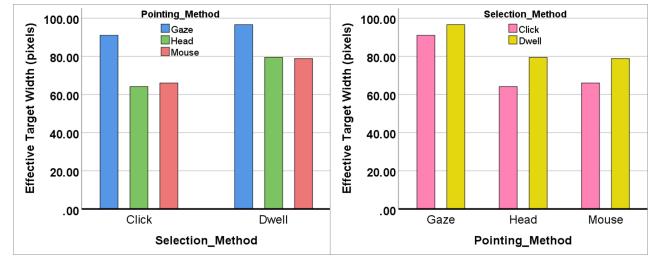


**Figure 2: Time To Activate (ms) by pointing method and selection method.**



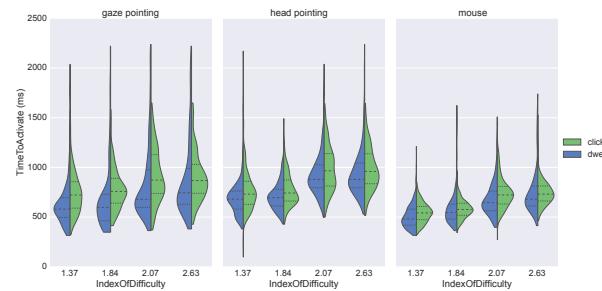
**Figure 3: Throughput (bits/s) by pointing method and selection method.**

The index of difficulty ( $ID$ ) influenced the three input methods according to Fitts’ law: Higher  $ID$  increases the time to activate. Figure 5 shows that dwell time activation was slightly faster than click activation for all pointing methods and  $ID$ s. We also examined



**Figure 4: Effective Target Width by pointing method and selection method.**

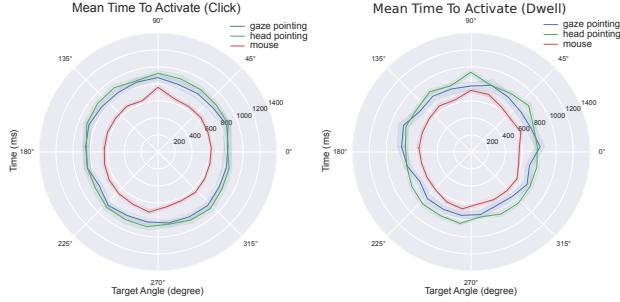
if movement direction had an impact on selection time for both the click and the dwell group. Click performance was very uniform in all directions, except for the very first target which took a bit longer to hit. Dwell activation was consistent in all directions, but showed more variation over target angle than click, cf. Figure 6.



**Figure 5: Violin plot, showing the time to activate the targets (y-axis, ms) vs. Fitts’ index of difficulty (x-axis, bits) split according to the selecting method (left: dwell in blue, right: click in green) and grouped according to pointing method (gaze, head, mouse). The violin shows the kernel density estimate of the underlying distribution. The median value is marked with a dashed line, and the quartile ranges are marked with dotted lines.**

Following the completion of the experiment, participants rated their experience on physical workload, mental workload, and comfort. Responses were on a 10-point Likert scale. These were analyzed using a Friedman non-parametric test. For post hoc analyses, we used a Wilcoxon signed-rank test, which is the non-parametric equivalent of a matched-pairs  $t$ -test. The results are presented in Table 3. The effect of pointing method was statistically significant for the three subjective responses. There exists a difference in the mental workload between the mouse and the other pointing methods; however, when comparing gaze pointing and head pointing, the mental workload is perceived equally. The physical workload differs significantly among the three pointing methods. Also, the participants experienced the highest physical workload with head pointing and the least with the mouse.

In summary, our study provides the following contributions: 1) dwell activation is faster than click selection, providing a mean target activation for both gaze and head of  $\approx 800$  ms (with a 300 ms



**Figure 6:** Polar plots showing the distribution of the time to activate according to target direction. Click conditions are shown on the left, dwell conditions on the right. One standard error of the means is indicated by a coloured shade. Blue is gaze pointing, green is head pointing, and red is mouse-controlled.

**Table 3: Friedman test of subjective responses for mental workload, physical workload, and comfort. Cells highlighted in gray indicate significance at  $p < 0.05$ .**

| Pointing                 | $\chi^2$                       | Post hoc Analysis  |  |  |
|--------------------------|--------------------------------|--|--|--|
| <b>Mental Workload</b>   | $\chi^2 = 21.64$<br>$p < 0.05$ | mean rank: gaze > head > mouse<br>(head, gaze) $p = 0.065 > 0.016$     |  |  |
|                          |                                | (mouse, gaze) $p = 0.000 < 0.016$<br>(head, mouse) $p = 0.001 < 0.016$ |  |  |
| <b>Physical Workload</b> | $\chi^2 = 28.59$<br>$p < 0.05$ | mean rank: head > gaze > mouse<br>(head, gaze) $p = 0.009 < 0.016$     |  |  |
|                          |                                | (mouse, gaze) $p = 0.010 < 0.016$<br>(head, mouse) $p = 0.000 < 0.016$ |  |  |
| <b>Comfort</b>           | $\chi^2 = 17.42$<br>$p < 0.05$ | mean rank: mouse > gaze > head<br>(head, gaze) $p = 0.249 > 0.016$     |  |  |
|                          |                                | (mouse, gaze) $p = 0.022 > 0.016$<br>(head, mouse) $p = 0.001 < 0.016$ |  |  |

dwell time). 2) Target width should be at least 94 pixels ( $\approx 3^\circ$ ) for the gaze method. 3) Comparing gaze and head pointing, error rates, throughput and effective target width are significantly better for head pointing. 4) Gaze pointing is considered less strenuous than head pointing by our participants.

Our interpretations are in line with Qian and Teather [Qian and Teather 2017] who concluded that "The eyes don't have it". However, we emphasize that both the present study and their study are based on one type of headset only (i.e. the FOVE), and may well change with improvements in tracking technology. Also, gaze pointing holds promise by tending to be faster than head pointing (although not significant so) and it is rated less physically demanding than head pointing.

## 5 CONCLUSIONS

Overall, gaze pointing is less accurate than head pointing and the mouse, and gaze pointing also has a lower throughput than either alternative considered here. Gaze pointing and head pointing are perceived mentally more demanding than mouse pointing. Head pointing is more physically demanding than gaze pointing.

## ACKNOWLEDGMENTS

We would like to thank the Bevica Foundation for funding this research. Also, thanks to Martin Thomsen, Atanas Slavov, Dr. Tracy Hammond, and Niels Andreassen for their support and guidance.

## REFERENCES

- P. Ballard and G. C. Stockman. 1995. Controlling a computer via facial aspect. *IEEE Transactions on Systems, Man, and Cybernetics* 25, 4 (Apr 1995), 669–677. <https://doi.org/10.1109/21.370199>
- John Paulin Hansen, Anders Sewerin Johansen, Dan Witzner Hansen, Kenji Itoh, and Satoru Mashino. 2003. Command without a click: Dwell time typing by mouse and gaze selections. In *Proceedings of Human-Computer Interaction—INTERACT*. 121–128.
- Robert JK Jacob. 1990. What you look at is what you get: eye movement-based interaction techniques. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 11–18.
- Shahram Jalaliniya, Diako Mardanbegi, and Thomas Pederson. 2015. MAGIC pointing for eyewear computers. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers*. ACM, 155–158.
- 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>
- Päivi Majaranta, I Scott MacKenzie, Anne Aula, and Kari-Jouko Räihä. 2006. Effects of feedback and dwell time on eye typing speed and accuracy. *Universal Access in the Information Society* 5, 2 (2006), 199–208.
- Darius Miniotas. 2000. Application of Fitts' Law to Eye Gaze Interaction. In *CHI '00 Extended Abstracts on Human Factors in Computing Systems (CHI EA '00)*. ACM, New York, NY, USA, 339–340. <https://doi.org/10.1145/633292.633496>
- Hunter Murphy and Andrew T Duchowski. 2001. Gaze-contingent level of detail rendering. *EuroGraphics 2001* (2001).
- 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.
- 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>
- Eyal M Reingold, Lester C Loschky, George W McConkie, and David M Stampe. 2003. Gaze-contingent multiresolutional displays: An integrative review. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 45, 2 (2003), 307–328.
- Linda E. Sibert and Robert J. K. Jacob. 2000. Evaluation of Eye Gaze Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '00)*. ACM, New York, NY, USA, 281–288. <https://doi.org/10.1145/332040.332445>
- 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, 6 (2004), 751–789.
- Oleg Špakov, Poika Isokoski, and Päivi Majaranta. 2014. Look and lean: accurate head-assisted eye pointing. In *Proceedings of the Symposium on Eye Tracking Research and Applications*. ACM, 35–42.
- Colin Ware and Harutune H. Mikaelian. 1987. An Evaluation of an Eye Tracker As a Device for Computer Input. In *Proceedings of the SIGCHI/GI Conference on Human Factors in Computing Systems and Graphics Interface (CHI '87)*. ACM, New York, NY, USA, 183–188. <https://doi.org/10.1145/29933.275627>
- Chun Yu, Yizheng Gu, Zhican Yang, Xin Yi, Hengliang Luo, and Yuanchun Shi. 2017. Tap, Dwell or Gesture?: Exploring Head-Based Text Entry Techniques for HMDs. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 4479–4488. <https://doi.org/10.1145/3025453.3025964>
- Xuan Zhang and I. Scott MacKenzie. 2007. *Evaluating Eye Tracking with ISO 9241 - Part 9*. Springer Berlin Heidelberg, Berlin, Heidelberg, 779–788. [https://doi.org/10.1007/978-3-540-73110-8\\_85](https://doi.org/10.1007/978-3-540-73110-8_85)