A Fitts' Law Study of Pupil Dilations in a Head-Mounted Display

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

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

Katsumi Minakata
Technical University of Denmark
Kgs. Lyngby, Denmark
katmin@dtu.dk

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

ABSTRACT

Head-mounted displays offer full control over lighting conditions. When equipped with eye tracking technology, they are well suited for experiments investigating pupil dilation in response to cognitive tasks, emotional stimuli, and motor task complexity, particularly for studies that would otherwise have required the use of a chinrest, since the eye cameras are fixed with respect to the head. This paper analyses pupil dilations for 13 out of 27 participants completing a Fitts' law task using a virtual reality headset with built-in eye tracking. The largest pupil dilation occurred for the condition subjectively rated as requiring the most physical and mental effort. Fitts' index of difficulty had no significant effect on pupil dilation, suggesting differences in motor task complexity may not affect pupil dilation.

CCS CONCEPTS

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

KEYWORDS

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

1 INTRODUCTION

Virtual reality (VR) and augmented reality (AR) headsets like the HTC VIVE, Microsoft Hololens, MagicLeap one, Qualcomms 845 VRDK reference, and the Oculus varifocal Half-Dome prototype are increasingly available, and are used in settings such as product development and manufacturing [Choi et al. 2015], education [Freina

and Ott 2015] and health care [Chirico et al. 2016; Khor et al. 2016; Sherman and Craig 2018; Valmaggia et al. 2016]. Some, like the FOVE headset, also include eye tracking, opening up the possibility of tracking pupillary reactions, potentially offering insight into the mental effort or other effort invested by the user.

This paper uses a well-known Fitts' law task [Soukoreff and MacKenzie 2004] over four input conditions (mouse, head, foot, and gaze) to investigate the corresponding pupillary dilation (i.e., pupil diameter) and the relation to perceived mental effort.

It is an open question if pupils dilate when simple visual-motor tasks increase in difficulty (e.g., Fletcher et al. [2017]; Jiang et al. [2014, 2015]). Do pupils dilate more when the index of difficulty (ID) in a Fitts' law task increases? Is there a difference between pupil dilation for the mouse, foot, head, and gaze input conditions? If so, do these differences align with the user's subjective experience?

2 RELATED WORK

Numerous studies found that the pupils dilate when cognitive load increases. Hess and Polt [1964] originally suggested using pupil dilation as an index of mental activity during multiplication. Kahneman and Beatty [1966] confirmed this finding in a separate study, which gave rise to pupillometry as a topic of research [Beatty and Lucero-Wagoner 2000; Laeng et al. 2012; Stanners et al. 1979] and creating interest among HCI researchers to include pupil measures in user performance assessments (e.g., Iqbal et al. [2004]).

Fitts' law [Soukoreff and MacKenzie 2004] is a frequently used model used in HCI, and was originally developed for one-dimensional tasks and since been extended and standardized [Fitts 1954; ISO 2000; MacKenzie and Buxton 1992]. It has been investigated elsewhere in VR settings [Teather and Stuerzlinger 2011] and using head mounted displays [Hansen et al. 2018; Lubos et al. 2014; Minakata et al. 2019; Qian and Teather 2017].

Richer and Beatty [1985] 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. [2014, 2015] 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. [2017] also used a Fitts' law movement task to manipulate motor response precision. Contrary to previous findings, increased precision demand was 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 indicates increased workload.

3 METHOD

3.1 Participants

Twenty-seven participants were recruited from the local university on a voluntary basis. Participants were initially screened for color blindness.

3.2 Apparatus

A HTC VIVE HMD, with a resolution of $2*1080 \times 1200$ pixels and a nominal field of view of 110°, was mounted with a Pupil Labs binocular eye-tracking system. The HMD renders at 90 Hz and the eye-tracking data were collected at 120 Hz. During the calibration, the background was brown-black (rgb[35,23,10]) and the tracking target was a black-and-white bullseye with 100% contrast. During the task, the background was the same color as the calibration; 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.

As manual input devices, a conventional Logitech corded mouse and a foot-mouse by 3DRudder were used.

A Fitts' law 2D implementation in Unity[1] was used as the experimental software.

3.3 Procedure

Before commencing the Fitts' law experiment, participants first did a calibration of the eye tracking unit. Pupillary data collected during this period later served as a baseline. Mouse input was used first, followed by head, foot, and gaze in different orders according to a Latin square. A target was activated when the cursor dwelled inside the target for 300 ms, based on the work of Majaranta et al. [2009].

Each pointing method had four levels of difficulty, comprised of target amplitudes of 80 and 120 pixels combined with target widths of 50 and 75 pixels, resulting in layout circle centres approximately 2.5° and 3.8° from the centre and target widths of approximately 3.1° and 4.7°, respectively. For each of the four levels, 21 trials were performed, presenting the initial target at the 12 o'clock position and then moving across and back to a neighbouring target in a continuous clockwise rotation until all targets were activated, as per the ISO 9241-9 procedure [ISO 2000]. See Figure 1.

If the participant failed to activate more than 20% of the targets, the sequence was repeated. At the end of each sequence, a small rest was allowed.

Participants were subsequently asked to rate the level of mental and physical demand for each control method. Responses were on a scale from 1 (least demanding) to 10 (most demanding).
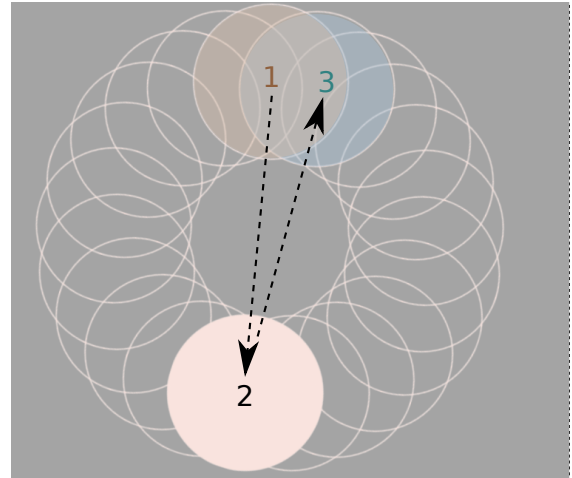

Figure 1: For Fitts' law task in 2D, targets are presented in the order shown, and continuing similarly clockwise.

3.4 Design

The experiment was 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 (80 pixels, 120 pixels)
- Target width (50 pixels, 75 pixels)

The dependent variable was

- Pupil dilation

4 RESULTS

4.1 Pupil Dilation

The pupil diameter, or dilation, 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), 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. It was, however, used to establish an approximate mean pupil diameter in mm corresponding to reported pixel values (see below).

All data frames with confidence less than or equal to 0.6 were discarded (as recommended by the vendor). This results in 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 0.82. Participants were then included only if (i) the Pearson R between eyes 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 data for 13 participants; the remaining participants were rejected because less than 25% valid eye tracking data was recorded from one or more of the test sequences. Of the rejected participants, the majority still had a high Pearson R, with a mean of 0.75, whereas the included participants had a mean of 0.89.

---

[1]based on an original at http://www.yorku.ca/mack/GoFitts/ which has been ported to Unity at https://github.com/GazeIT-DTU/FittsLawUnity
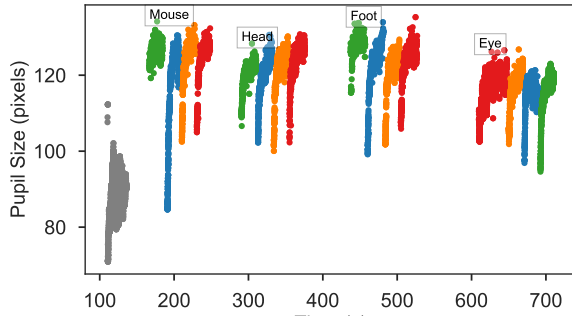
Figure 2: Typical pupil size data (average both eyes in pixels, for participant 8) as it varies with blocks (annotated), over the duration of the entire experiment (time in seconds). Grey is used for data recorded while calibrating, whereas the colors green, blue, orange, and red indicate the four conditions of target amplitude and target width (80, 75), (80, 50), (120, 75) and (120, 50), left to right.

Figure 2 shows typical eye tracking data. Most participants show a pattern of dilated pupils during the task versus calibration, even if both use the same background colour.

For the subsequent pupillary analysis, the average of left and right eye pupil dilation 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 dilation was established[2]. This value was subtracted from all pupil dilation measurements, independently for each participant, resulting in a pupil dilation vs. calibration (also in pixels). In addition, the pupil dilation vs. block mean was calculated by subtracting the mean of each block of sequences of trials from all pupil dilation readings within the particular block.

The grand mean of the pupil diameter was 115 pixels, corresponding to an approximate 3D modelled pupil diameter of 6.3 mm. The grand mean of the pupil dilation vs. calibration was 30 pixels, corresponding to an approximate 3D modelled pupil dilation of 1.9 mm.

The mean pupil dilations vs. calibration were 28 pixels (gaze), 30 pixels (head-position), 30 pixels (mouse), and 32 pixels (foot-mouse). See Figure 3.

The effect of pointing method on pupil dilation was statistically significant ($F(3, 180) = 7.26$, $p = .0001$, $\eta^2 = .04$, partial $\eta^2 = .11$). A post-hoc Wilcoxon Signed Rank test with Bonferroni correction showed that the foot pointing method was significantly different from gaze ($W = 180$, $p < .0001$) and head ($W = 240$, $p = .0003$). The effects of target amplitude and target width on pupil dilation were not statistically significant ($p > .05$).

The mean value of the pupil dilation by trial within each sequence varies over time, Figure 4, and correspondingly over angle, Figure 5. The pupil dilation over all participants and conditions increased, particularly over the first four trials, from an initial value of 107 pixels to a mean value over the last 17 trials of 116 pixels.

---

[2]This was done to achieve baseline independence of the Fitts' law tasks, nevertheless with identical screen backgrounds to minimize luminance induced pupillary changes.
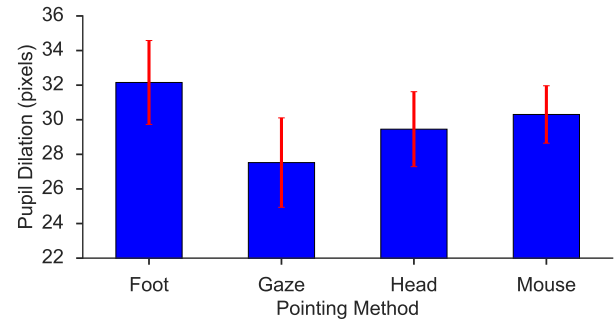

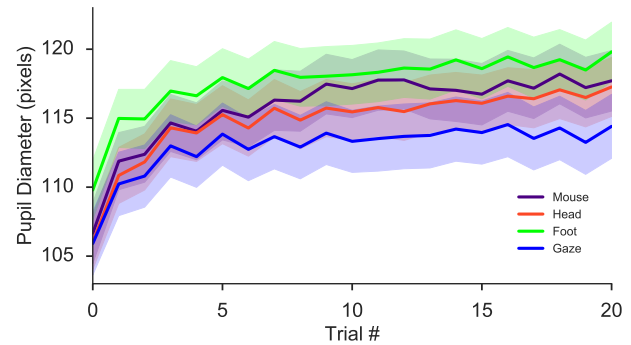**Figure 3: Pupil dilation vs. calibration (pixels) by pointing method. Error bars denote ±1 *SE*.**

Figure 4: Mean pupil diameter (pixels) by trial index for each pointing method. The shaded area denotes ±1 *SE*.

The pupil dilation aggregated over all blocks and participants for the four combinations of target amplitude and width is shown in Figure 6. Recall that the effects of target amplitude and target width on pupil dilation were not significant. The highest value is for the target amplitude and width of (80, 50) at 0.81 pixels whereas the lowest value is for (120, 75) at -0.79. Both have the same corresponding index of difficulty of 1.38 bits.

4.2 Subjective Ratings

Participants rated foot-pointing (mean rank = 3.19) as the most mentally demanding, followed by gaze-pointing (mean rank = 2.92), then head-pointing (mean rank = 2.15); finally, mouse-pointing (mean rank = 1.73) was rated the least mentally demanding. A Friedman test on the mental workload ratings was significant ($\chi^2(3) = 13.49$, $p = .004$). Foot-pointing was significantly more mental demanding than head- and mouse-pointing ($Z = -2.82$, $p = .01$; $Z = -2.67$, $p = .01$). No other differences were significant regarding the pointing-method ratings relating to mental workload ($p > .10$).

Participants rated foot-pointing (mean rank = 3.27) as the most physically demanding, followed by head-pointing (mean rank = 3.19), then gaze-pointing (mean rank = 2.08); finally, mouse-pointing (mean rank = 1.46) was rated the least physically demanding. A Friedman test on the physical workload ratings was significant
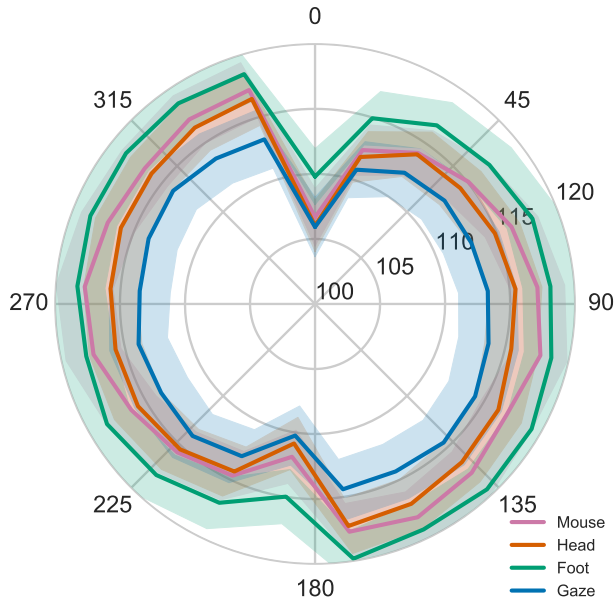
Figure 5: Mean pupil diameter (pixels) by target angle (degree) aggregated over all sequences for each of the four pointing conditions. The shaded area denotes ±1 *SE*.
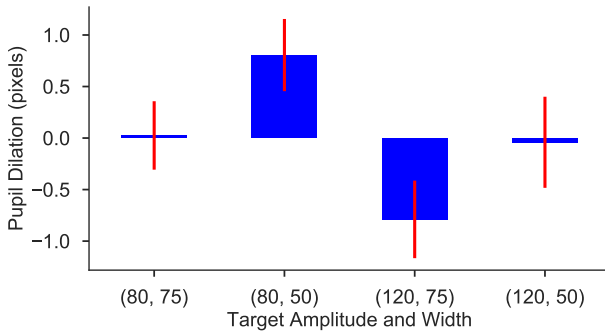

Figure 6: Pupil dilation relative to block mean (pixels) vs. target amplitude and target width (corresponding to Fitts' index of difficulty of 1.05, 1.38, 1.38, and 1.77 bits, left to right). Error bars denote ±1 *SE*.

($\chi^2(3)$ = 19.20, p = .0003). Foot-pointing was significantly more physically demanding than mouse-pointing (Z = -3.21, p = .001). Head-pointing was also significantly more physically demanding than mouse-pointing (Z = -3.09, p = .02). No other differences were significant regarding the pointing-method ratings relating to physical workload (p > .10).

5 DISCUSSION

Pupil results were contrary to our expectations in three ways. First, there was no simple relation between pupil dilation and target amplitude or target width, which yield Fitts' index of difficulty.

Second, foot input was associated with larger dilations than gaze input, possibly suggesting that higher physical or mental effort was needed when moving the feet, as compared to moving the eyes and 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 visual attention in our environment. Pointing with the feet by use of a balance board is new to most people. Thus it requires extra effort.

Third, Figure 4 shows that for every new sequence encountered, participants exhibited an increase in pupil diameter for the first trials in a sequence. This *start-up effect* was independent of the input method. Twelve of 13 participants showed clear signs of a start-up dilation from a lower initial value. This is not attributed to changes in luminance since the target color was equiluminant with the background seen before the onset of the task sequence. We speculate that it is related to the ramp-up phase often seen for task-evoked pupillary responses [Beatty 1982], until seemingly reaching a level suitable to the task. This finding implies there is value in including warm-up trials or leaving out the first trial data when analyzing pupil data for repeated actions.

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 conclusive, since there is an open issue whether mental or 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, impact the start-up effect. If so, when a start-up effect is not found, low cortical activity may be considered when assessing 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. Furthermore, a more robust 3D eye model is needed to more accurately facilitate experiments where areas of interest cannot be centered in the field of view.

6 CONCLUSION

In conclusion, pupil dilations were consistently associated with the onset of a task, independent of the pointing method. Pupil dilations depended on pointing method, but not on target amplitude or target width, thus not on Fitts' index of difficulty.

REFERENCES

Jackson Beatty. 1982. Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological bulletin* 91, 2 (1982), 276.

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

Andrea Chirico, Fabio Lucidi, Michele De Laurentiis, Carla Milanese, Alessandro Napoli, and Antonio Giordano. 2016. Virtual reality in health system: beyond entertainment. a mini-review on the efficacy of VR during cancer treatment. *Journal of cellular physiology* 231, 2 (2016), 275–287.

SangSu Choi, Kiwook Jung, and Sang Do Noh. 2015. Virtual reality applications in manufacturing industries: Past research, present findings, and future directions. *Concurrent Engineering* 23, 1 (2015), 40–63.

Paul M Fitts. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology* 47, 6 (1954), 381.

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

Laura Freina and Michela Ott. 2015. A literature review on immersive virtual reality in education: State of the art and perspectives. *eLearning & Software for Education* 1 (2015), 9.

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 (COGAIN '18)*. ACM, New York. https://doi.org/10.1145/3206343.3206344

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

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, 1477–1480. https://doi.org/10.1145/985921.986094

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.

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.

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.

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

Wee Sim Khor, Benjamin Baker, Kavit Amin, Adrian Chan, Ketan Patel, and Jason Wong. 2016. Augmented and virtual reality in surgery - the digital surgical environment: applications, limitations and legal pitfalls. *Annals of Translational Medicine* 4, 23 (2016).

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

Paul Lubos, Gerd Bruder, and Frank Steinicke. 2014. Analysis of direct selection in head-mounted display environments. In *2014 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, 11–18.

I Scott MacKenzie and William Buxton. 1992. Extending Fitts' law to two-dimensional tasks. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Cystems - CHI '92*. ACM, New York, 219–226.

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

Katsumi Minakata, John Paulin Hansen, I. Scott MacKenzie, Per Bækgaard, and Vijay Rajana. 2019. Pointing by Gaze, Head, and Foot in a Head-Mounted Display. In *Communication by Gaze Interaction (COGAIN @ ETRA'19)*. ACM, New York. https://doi.org/10.1145/3317956.3318150

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.

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

William R Sherman and Alan B Craig. 2018. *Understanding virtual reality: Interface, application, and design*. Morgan Kaufmann.

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. https://doi.org/10.1016/j.ijhcs.2004.09.001 Fitts' law 50 years later: applications and contributions from human-computer interaction.

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

Robert J Teather and Wolfgang Stuerzlinger. 2011. Pointing at 3D targets in a stereo head-tracked virtual environment. In *2011 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, 87–94.

Lucia R Valmaggia, Leila Latif, Matthew J Kempton, and Maria Rus-Calafell. 2016. Virtual reality in the psychological treatment for mental health problems: An systematic review of recent evidence. *Psychiatry Research* 236 (2016), 189–195.