A Fitts' Law Evaluation of Gaze Input on Large Displays Compared to Touch and Mouse Inputs

Vijay Rajanna
Department of Computer Science and Engineering
Texas A&M University, College Station, Texas
vijay.drajanna@gmail.com

Tracy Hammond
Department of Computer Science and Engineering
Texas A&M University, College Station, Texas
thammond@gmail.com

ABSTRACT

Gaze-assisted interaction has commonly been used in a standard desktop setting. When interacting with large displays, as new scenarios like situationally-induced impairments emerge, it is more convenient to use the gaze-based multi-modal input than other inputs. However, it is unknown as to how the gaze-based multi-modal input compares to touch and mouse inputs. We compared gaze+foot multi-modal input to touch and mouse inputs on a large display in a Fitts' Law experiment that conforms to ISO 9241-9. From a study involving 23 participants, we found that the gaze input has the lowest throughput (2.33 bits/s), and the highest movement time (1.176 s) of the three inputs. In addition, though touch input involves maximum physical movements, it achieved the highest throughput (5.49 bits/s), the least movement time (0.623 s), and was the most preferred input.

CCS CONCEPTS

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

KEYWORDS

Fitts' Law; ISO 9241-9; gaze; touch; mouse; foot input; throughput.

ACM Reference Format:
Vijay Rajanna and Tracy Hammond. 2018. A Fitts' Law Evaluation of Gaze Input on Large Displays Compared to Touch and Mouse Inputs. In *COGAIN '18: Workshop on Communication by Gaze Interaction, June 14–17, 2018, Warsaw, Poland*. ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3206343.3206348

1 INTRODUCTION

Human-Computer Interaction involves numerous application contexts and scenarios where hands-free interaction is crucial. Applications demanding rich interactions for efficiency, and users with accessibility needs rely on eye tracking technology as a hands-free input modality [Duchowski 2007]. As an accessible technology, gaze input serves as the primary mode of communication for individuals with severe motor and speech disability [Majaranta and Räihä 2002; Rajanna 2016]. As gaze is increasingly used as an input modality, its applicability is not just limited to accessible technology, but there are emerging use cases such as leveraging gaze-assisted interaction in situationally-induced impairments and disabilities (SIID) [Kane et al. 2008]. In the case of SIIDs, a user's hands are assumed to be engaged in other tasks, and hence unavailable for selecting interface elements or typing by using touch, mouse, or keyboard. For example, the hands of a surgeon performing an operation, a musician playing music, a worker on a factory assembly line, and so on tend to be engaged in a specific task, and hence represent a case of SIID.

Gaze-assisted interaction in a desktop setting as an efficient interaction method or a solution to SIID has been previously explored [Hansen et al. 2008, 2016; Rajanna et al. 2017; Stellmach and Dachselt 2012], and also compared against other inputs [Miniotas 2000; Soukoreff and MacKenzie 2004; Vertegaal 2008; Zhang and MacKenzie 2007]. Gaze-assisted interaction on large displays has various applications as people can interact with public displays, screens in collaborative spaces, operation theaters, etc. While there are various examples of using gaze input on large displays [Hatscher et al. 2017; San Agustin et al. 2010; Vidal et al. 2013], its comparison to other commonly used inputs like touch and mouse are limited. To discuss a few relevant works that explored gaze-assisted interaction on large displays, in an upright stance, Hatscher et al., demonstrated the usability of gaze- and foot-based interaction on a large monitor in operation theaters [Hatscher et al. 2017]. In this setup, a physician performing minimally-invasive interventions can look and interact with medical image data displayed on the large monitor with gaze input. San Agustin et al., developed gaze-enabled public display (55 inches) where users can interact with high-density information like a digital bulletin board with several notes on top of each other [San Agustin et al. 2010].

These works demonstrate that in the cases of SIIDs or from a usability perspective it is more relevant to use gaze-assisted interaction on large displays (~ 84 inches). However, the majority of the table mounted trackers (Tobii, EyeTribe, GazePoint, SMI, etc.) are built for 24-inch screens. To achieve the best performance, the optics and IR lights are tuned for the viewing angles that correspond to screens up to 24 inches. This does not mean that we cannot use these trackers on large screens, but they do not track well. Therefore, for someone trying to use these commonly available eye trackers on large displays, it is unknown as to 1) How the accuracy and efficiency of pointing and selecting with gaze input compare against touch and mouse inputs that are used commonly on the large displays? 2) How does the usability of gaze input compare to mouse and touch inputs? 3) What should be size of the targets (this influences the index of difficulty)? 4) While the touch input requires a user to physically move in front of the display to reach different points on the display, and mouse input requires larger movements of

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

ACM ISBN 978-1-4503-5790-6/18/06...$15.00
https://doi.org/10.1145/3206343.3206348

the wrist, do users feel touch and mouse inputs stressful compared to the gaze input? The lack of answers to these queries motivated us to conduct a Fitts' Law evaluation that conforms to ISO 9241-9 standardization[1].

Fitts' Law models the human movement analogous to the way information is transmitted [Soukoreff and MacKenzie 2004]. Different kinds of movement tasks have different indices of difficulties expressed in bits/sec. To perform a movement task, a certain number of bits of information is transmitted by the human motor system. The performance of a movement task can be quantified (throughput) by dividing the number of bits transmitted by the movement time (MT) [Soukoreff and MacKenzie 2004; Zhang and MacKenzie 2007]. Furthermore, Fitts' Law has been used in HCI research in two ways, first, to predict the time it takes (movement time) for a user of a graphical interface to move the cursor to the target and click it. Second, to compare the speed and accuracy of different input methods through a single statistic called throughput [Soukoreff and MacKenzie 2004]. The throughput of an input method is computed as follows:

$$Throughput = \frac{ID_e}{MT} \tag{1}$$

Where, ID_e is the effective Index of Difficulty, and MT is the mean Movement Time. The ID_e is calculated as

$$ID_e = log_2\left(\frac{A_e}{W_e} + 1\right) \tag{2}$$

Where, A_e is the effective distance to the target (amplitude), and W_e is the effective target width which is calculated as

$$W_e = 4.133 \times SD_x \tag{3}$$

Where, SD_x is the standard deviation of the selection coordinates (overshoot or undershoot) in a sequence.

In this study we compared three input methods for pointing and selecting targets on a large display. The three input methods we used were gaze+foot, touch, and mouse. While we initially considered gaze+dwell based activation, we did not include this input due to low gaze tracking accuracy. Though we chose a nominal index of difficulty (max 2.5 bits/s), from the pilot study we found that the participants could not select all the targets without crossing the error threshold. Hence, participants could not complete the experiment when using gaze- and dwell-based selection, and this is a major limitation of using gaze+dwell input on large displays. However, with low indices of difficulties (~ 1.0) gaze+dwell input should be usable, but would constrain the number of targets.

2 PRIOR WORK

Fitts' Law evaluation of gaze input has been primarily conducted in desktop settings, and the common selection methods considered have been gaze+dwell, gaze+click, and mouse [MacKenzie 1992; Miniotas 2000; Miniotas et al. 2006; Soukoreff and MacKenzie 2004; Vertegaal 2008; Zhang and MacKenzie 2007]. Zhang et al., presented the first work on Fitts' law evaluation of gaze input that conforms to ISO 9241-9 [Zhang and MacKenzie 2007]. The authors compared gaze input with short and long dwell times and gaze+Spacebar with mouse input. The Gaze+Spacebar eliminated the waiting time, it was the best selection method among gaze inputs with a throughput of 3.78 bits/s (mouse was 4.68 bits/s). Ware et al., presented a Fitts' law evaluation of gaze input [Ware and Mikaelian 1987]. Three selection methods were used along with gaze: a button press, dwelling, and an onscreen select button. In an experiment where the participants had to select one of the seven menu items arranged vertically, the authors found that irrespective of the selection procedure, the gaze-based selection methods took less than 1 second for target selection.

Miniotas et al., tested the validity of the findings from Ware et al. [Ware and Mikaelian 1987], by comparing the performance of an eye tracker and a mouse in a simple pointing task [Miniotas 2000]. The authors found that the selection time is longer for the eye tracker than for the mouse by a factor of 2.7. Zhai et al., proposed MAGIC: Manual and Gaze Input Cascaded Pointing to improve the usability of gaze input [Zhai et al. 1999]. A Fitts' law evaluation was conducted by using 3 input methods which included an isometric pointing stick and two versions of MAGIC pointing. The authors found that the completion time and target distance did not completely follow Fitts' law when using MAGIC pointing, but when considering both target size and target distance the data fit the Fitts' law but relatively poorly. Pointing with two version of MAGIC achieved a higher performance (4.55 and 4.76 bits/s) than manual input (3.2 bits/s).

Vertegaal et al., presented a Fitts' Law evaluation of gaze+click, gaze+dwell, mouse, and stylus [Vertegaal 2008]. The performance of these selection methods were compared in selection of large visual targets. The index of difficulty varied from 1.28 bits/s to 3.6 bits/s. The authors found that gaze+click outperformed the mouse by 16%. However, eye tracking inputs suffered a high error rate of 11.7% for manual click and 43% for dwell time click. Eye tracking with manual click with an index of performance of 10.9 bits/s appeared to be the best trade off between speed and accuracy. We observe that all the Fitts' Law evaluations we discussed were conducted in a desktop setting, and the participant was always seated while using the various input methods. Contrary to this typical setup, our work performs Fitts' Law evaluation of the gaze input on a large display, while comparing it against mouse and touch inputs. Also, the participants used the three inputs in an upright stance.

3 FITTS' LAW EXPERIMENT DESIGN

For the Fitts' Law experiment we used the software [2] developed by Soukoreff and MacKenzie [MacKenzie 1992; Soukoreff and MacKenzie 2004]. Specifically, we used Fitts' Task Two which is a multi-directional point-and-select task. For each trial the target to be selected is highlighted in red color, and once the highlighted target is selected, the target that is opposite to the current target gets highlighted. In accordance with the previous Fitts' law studies on gaze pointing, we used a nominal index of difficulty that ranged from 2.0 to 2.5 [Zhang and MacKenzie 2007]. Hence, the amplitude, i.e., the distance to the target we chose were 1650 px and 1250 px, and the target widths were set to 350 px and 450 px. Figure 1 shows the Fitts' Law experiment setup on a large display.

3.1 Selection Methods

We chose three selection methods: 1) gaze+foot, 2) touch, and 3) mouse, and all the three methods were used in an upright stance.

[1]https://www.iso.org/standard/30030.html [last accessed Jan 23rd 2018]

[2]http://www.yorku.ca/mack/FittsLawSoftware/ [last accessed Jan 23rd 2018]

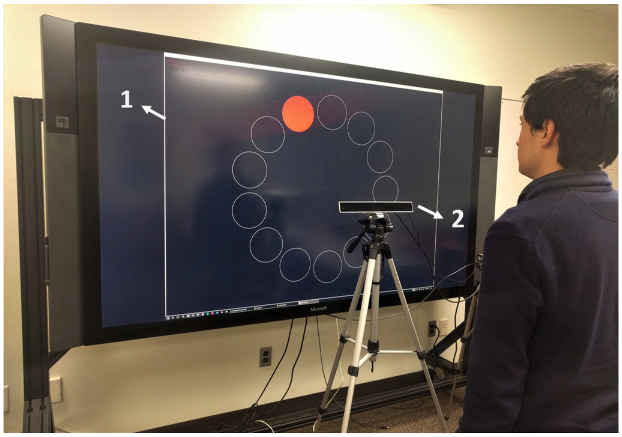

Figure 1: Fitts' Law Experiment: a participant, in an upright stance, performing a multi-directional point-and-select Fitts' Law task (1) shown on a large display. Also, an eye tracker is mounted on a tripod (2).

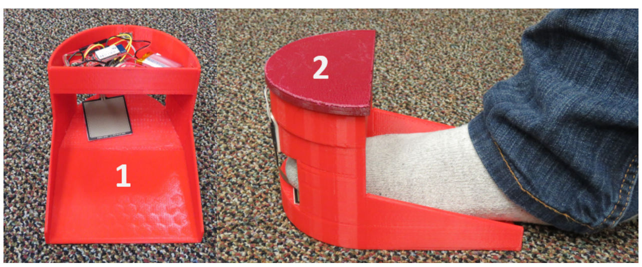

Figure 2: The foot controller used in the gaze+foot selection method. 1 - a force sensitive resistor, microcontroller, and bluetooth module in a 3D printed case, 2 - foot interaction.

For gaze+foot input, the eye tracker was mounted on a tripod, and placed in between the user and the display as shown in Figure 1. The eye tracker was removed when using the mouse and touch inputs. For the mouse input, the participant used a wireless mouse placed on a tall, height adjustable table placed next to the user, and for the touch input, the participant directly touched the screen to select targets.

To achieve the gaze+foot interaction, we enhanced a gaze+foot input system developed by Rajanna et al. [Rajanna and Hammond 2016], which consists of an eye tracking module and a foot controller (Figure 2), and the on-screen cursor follows the user's gaze. To select a target the user first places the cursor on the target by focusing on it, and then selects it by pressing a pressure sensor, attached to the foot controller, with the foot. The foot controller connects to the eye tracking system over Bluetooth, and the entire circuitry is placed inside a portable 3D printed case.

3.2 Display and Gaze Tracking

The experiment was conducted on a Microsoft Surface Hub [3], a large (84-inch) touch enabled display. We used a Gazepoint GP3 HD tracker for eye tracking. Since the tracking was not accurate enough around the left and right edges of the 84-inch screen, the interaction space was set to 69 inches. The tracker had a manufacturer reported accuracy of 0.5° to 1.0° of visual angle, and had a sampling rate of 150 Hz. However, to test the accuracy of the tracker for our setup with a large display, we recruited 7 (6 M, 1 F) participants, and repeatedly recorded the tracking accuracy values (following the standard calibration) on a 9-points grid interface we developed. A total of 39 accuracy values were recorded, and the average tracking accuracy was 4.6° of visual angle (min 2.6°, max 9°).

[3]https://www.microsoft.com/en-us/surface/devices/surface-hub/tech-specs

3.3 Participants and Procedure

For the Fitts' Law experiment we recruited 23 participants (19 M, 4 F) with their ages ranging from 19 to 32 (μ_{age} = 23). Data from 4 participants were excluded since they could not complete the gaze input due to low tracking accuracy. Also, 3 participants who were wearing glasses removed their glasses (for better gaze tracking accuracy) during the experiment. At the beginning of the study, each participant was briefed about the Fitts' Law task and the kind of inputs they would be using for target selection. For each input method (e.g., mouse) the participant completed one sequence of trials to familiarize themselves with the system before the actual data collection began. The participants used three input methods–gaze+foot, mouse, and touch–for target selection, and the order of input methods used by the participants was counterbalanced according to the Latin square design.

For each input method the participant completed 4 blocks of target selection task, and each block had four sequences of trials as we used two amplitudes (1650 px and 1250 px) and two target widths (350 px and 450 px). In each sequence, there were 13 trials, hence, a total of 3,952 trials (13 trials × 4 seq × 4 blocks × 19 participants) were completed for each input. Also, a total of 11,856 trials (3,952 × 3 inputs) were completed from all the three inputs. The participants were allowed to rest for a minute between each block, and in the case of gaze input, the participants were re-calibrated if the calibrated stance was disturbed between the blocks.

4 RESULTS AND DISCUSSION

We conducted a one-way ANOVA with replication on the four dependent variables (DVs): 1) movement time, 2) throughput, 3) error rate, and 4) effective target width. The independent factor was the 'selection method' which had three levels: 1) mouse, 2) touch, and 3) gaze. Table 1 shows the result of ANOVA on the DVs, and also the mean and standard deviation of the selection methods for each DV.

We observe that the factor 'selection method' is significant ($p < 0.05$) for all the four DVs, i.e., the value of a DV differs among the selection methods. Out of all the selection methods, 'touch' achieves the highest throughput (5.49 bits/sec), consequently it has the least movement time, error, and effective target width. Similarly, 'gaze' input has the lowest throughput (2.33 bits/sec), consequently it has the highest movement time, error, and effective target width. Post-hoc tests with Bonferroni correction showed that for DVs movement time, throughput, and effective target width the difference between each pair of the selection methods, (mouse, touch) (mouse, gaze) (touch, gaze), was significant (p <

Table 1: Fitts' Evaluation: ANOVA and post-hoc analysis (p values highlighted in gray indicate significance at $\alpha = 0.05$).

Selection Method [Ms, Th, Gz]	Mean	Std. Dev	ANOVA
Movement Time (ms)	Ms = 777.884	152.23	F(2,606) = 242.196
	Th = 623.253	139.50	$p = 0.000$
	Gz = 1176.54	561.14	
Throughput (bits/s)	Ms = 3.449	0.885	F(2,606) = 755.789
	Th = 5.498	1.487	$p = 0.000$
	Gz = 2.331	1.001	
Error Rate (%)	Ms = 1.0374	3.3501	F(2,606) = 161.763
	Th = 0.6073	3.0006	$p = 0.000$
	Gz = 8.7298	9.9549	
Effective Target Width (pixels)	Ms = 301.671	207.508	F(2,606) = 51.659
	Th = 198.462	272.184	$p = 0.000$
	Gz = 407.084	285.128	

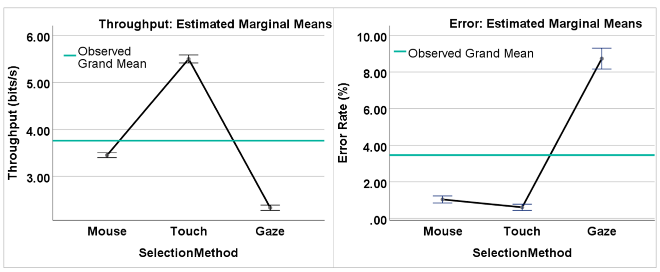

Figure 3: Comparison of estimated marginal means for DVs 'Throughput' and 'Error Rate' for the three selection methods. The error bars represent standard error of the mean.

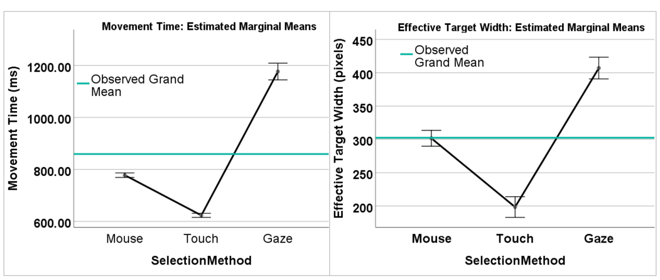

Figure 4: Comparison of estimated marginal means for DVs 'Movement Time' and 'Effective Target Width' for the three selection methods. The error bars represent standard error of the mean.

0.05). However, for the DV error, the difference between each pair of selection methods was significant except for the pair (mouse, touch) where p = 0.32 > 0.05. Figure 3 and Figure 4 compare the means of the three selection methods for each DV.

Though theoretically it appears that gaze input should achieve a higher throughput than the mouse and touch inputs, since an eye movement between two distant targets is quicker [Vertegaal 2008] than the mouse or touch, the results contradict our assumption. This is due to the fact that though the user may move the cursor quickly from target A to the vicinity of target B, placing the cursor exactly on target B and selecting it consumes more time due to lower tracking accuracy on the large display. Therefore, the results suggests that there are two ways to improve the throughput of gaze-based selection on large displays. First, by reducing the index of difficulty of the task, i.e., primarily by increasing the target width, and also by reducing the distance between the targets. Second, by developing eye trackers exclusively for the large displays (larger than 24 inches). Also, in the interviews the participants shared that with gaze+foot interaction it is essential to achieve the synchronization between pointing with gaze and selecting with foot.

Furthermore, we wanted to understand if the users' performance, specifically throughput and error, improve as they progress from block 1 to block 4 for a given selection method (e.g., touch). Hence, we conducted a one-way ANOVA with replication on the dependent variables 'throughput' and 'error' for each of the selection methods, and the independent factor was 'block.' Table 2 shows the block mean and standard deviation for various DVs, and corresponding ANOVA results.

Table 2: Fitts' Evaluation: ANOVA of block performance (p values highlighted in gray indicate significance at $\alpha = 0.05$).

Block [B1 to B4]	Mean [Std. Dev]	ANOVA
Mouse Throug -hput (bits/s)	B1 = 3.19 [0.91], B2 = 3.48 [0.80]	F(3,225) = 7.75
	B3 = 3.65 [0.91], B4 = 3.46 [0.86]	$p = 0.000$
Touch Throug -hput (bits/s)	B1 = 4.78 [1.39], B2 = 5.58 [1.22]	F(3,225) = 13.23
	B3 = 5.68 [1.56], B4 = 5.93 [1.51]	$p = 0.000$
Gaze Throug -hput (bits/s)	B1 = 2.09 [1.01], B2 = 2.49 [1.03]	F(3,225) = 5.61
	B3 = 2.29 [0.97], B4 = 2.44 [0.95]	$p = 0.001$
Mouse Error Rate (%)	B1 = 0.91 [3.76], B2 = 0.91 [2.50]	F(3,225) = 1.68
	B3 = 0.70 [2.23], B4 = 1.61 [4.40]	$p = 0.172$
Touch Error Rate (%)	B1 = 0.80 [3.45], B2 = 0.10 [0.88]	F(3,225) = 1.38
	B3 = 1.01 [3.83], B4 = 0.50 [2.90]	$p = 0.250$
Gaze Error Rate (%)	B1 = 10.3 [10.7], B2 = 8.90 [9.23]	F(3,225) = 1.38
	B3 = 7.89 [9.31], B4 = 7.79 [10.4]	$p = 0.250$

We observe that the factor 'block' is significant for DVs gaze, mouse, and touch throughputs. The throughput generally increases as the user progresses from block 1 to block 4, which is an indication that the users' performance does improve with more exposure to the selection method. However, we also see that the difference in 'error' between the blocks is not significant for all the DVs. This suggests that the users get quicker in selecting targets with subsequent blocks, however, the accuracy of selection remains the unchanged.

4.1 Subjective Feedback

Each participant rated their experience of using the three selection methods on a Likert scale (1-very low to 7-very high) for various physiological measures. Figure 5 summarizes the mean value of

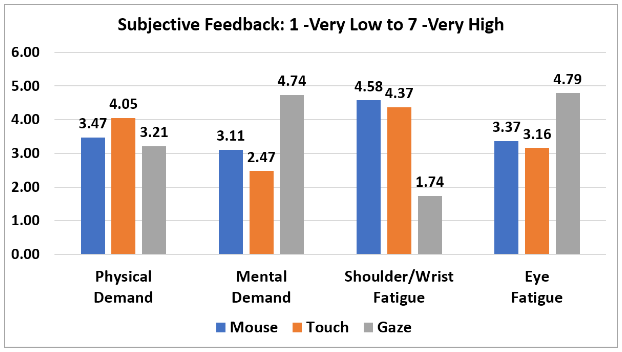

Figure 5: Subjective Feedback - lower score is better.

each measure. As we may expect, touch and mouse results in increased shoulder/wrist fatigue and physical demand on the large displays compared to the gaze input. On the contrary, the gaze input results in increased mental demand and eye fatigue compared to touch and mouse inputs. However, gaze input has an added advantage of enabling hands-free interactions that are crucial in the cases of SIIDs. We further analyzed the ratings using a one-way ANOVA with replication and considering 'selection method' as the independent factor. We found that 'selection method' is not a significant factor for DV 'Physical Demand' [$F(2,36) = 1.506$, $p > 0.05$]. However, 'selection method' is a significant factor for DVs 'Mental Demand' [$F(2,36) = 19.09$, $p < 0.05$], 'Eye Fatigue' [$F(2,36) = 16.128$, $p < 0.05$], and 'Shoulder/Wrist Fatigue' [$F(2,36) = 34.283$, $p < 0.05$]. Also, for DVs 'Mental Demand', 'Eye Fatigue', and 'Shoulder/Wrist Fatigue' where the ANOVA results are significant, post-hoc tests with Bonferroni correction showed that the effect was due to the difference between (gaze, touch) and (gaze, mouse), but the difference between (touch, mouse) was not significant.

5 CONCLUSION

As new scenarios emerge making the gaze input more relevant on large displays, we wanted to compare the usability and efficiency of gaze input to other commonly used inputs. Hence, we compared the gaze input against mouse and touch inputs on a large display in a Fitts' Law evaluation that conforms to ISO 9241-9 standardization. Though gaze enables faster cursor movements between the targets theoretically, we found that gaze input had the lowest throughput and highest error rate. On the contrary, although touch results in increased shoulder/neck fatigue, it achieves the highest throughput and lowest error rate, and was the most preferred input.

ACKNOWLEDGMENTS

We would like to thank "Microsoft Corporation" for sponsoring a "Microsoft Surface Hub." We are also thankful to Dr. John Paulin Hansen (Technical University of Denmark), Dr. Scott Mackenzie (York University), Mohammad Khan, Josh Cherian, and the members of Sketch Recognition Lab for their assistance and feedback.

REFERENCES

Andrew T Duchowski. 2007. Eye tracking methodology. *Theory and practice* 328 (2007).

Dan Witzner Hansen, Henrik H. T. Skovsgaard, John Paulin Hansen, and Emilie Møllenbach. 2008. Noise Tolerant Selection by Gaze-controlled Pan and Zoom in 3D. In *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications (ETRA '08)*. ACM, New York, NY, USA, 205–212. https://doi.org/10.1145/1344471.1344521

John Paulin Hansen, Haakon Lund, Florian Biermann, Emillie Møllenbach, Sebastian Sztuk, and Javier San Agustin. 2016. Wrist-worn pervasive gaze interaction. In *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications*. ACM, 57–64.

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

Shaun K. Kane, Jacob O. Wobbrock, and Ian E. Smith. 2008. Getting off the Treadmill: Evaluating Walking User Interfaces for Mobile Devices in Public Spaces. In *Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '08)*. ACM, New York, NY, USA, 109–118. https://doi.org/10.1145/1409240.1409253

I Scott MacKenzie. 1992. Fitts' law as a research and design tool in human-computer interaction. *Human-computer interaction* 7, 1 (1992), 91–139.

Päivi Majaranta and Kari-Jouko Räihä. 2002. Twenty years of eye typing: systems and design issues. In *Proceedings of the 2002 symposium on Eye tracking research & applications*. ACM, 15–22.

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

Darius Miniotas, Oleg Špakov, Ivan Tugoy, and I. Scott MacKenzie. 2006. Speech-augmented eye gaze interaction with small closely spaced targets. In *Proceedings of the 2006 symposium on Eye tracking research & applications - ETRA '06*. ACM Press, New York, New York, USA, 67. https://doi.org/10.1145/1117309.1117345

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

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

Vijay Rajanna, Seth Polsley, Paul Taele, and Tracy Hammond. 2017. A Gaze Gesture-Based User Authentication System to Counter Shoulder-Surfing Attacks. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. ACM, New York, NY, USA, 1978–1986. https://doi.org/10.1145/3027063.3053070

Javier San Agustin, John Paulin Hansen, and Martin Tall. 2010. Gaze-based Interaction with Public Displays Using Off-the-shelf Components. In *Proceedings of the 12th ACM International Conference Adjunct Papers on Ubiquitous Computing - Adjunct (UbiComp '10 Adjunct)*. ACM, New York, NY, USA, 377–378. https://doi.org/10.1145/1864431.1864444

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.

Sophie Stellmach and Raimund Dachselt. 2012. Look & Touch: Gaze-supported Target Acquisition. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 2981–2990. https://doi.org/10.1145/2207676.2208709

Roel Vertegaal. 2008. A Fitts Law Comparison of Eye Tracking and Manual Input in the Selection of Visual Targets. In *Proceedings of the 10th International Conference on Multimodal Interfaces (ICMI '08)*. ACM, New York, NY, USA, 241–248. https://doi.org/10.1145/1452392.1452443

Mélodie Vidal, Andreas Bulling, and Hans Gellersen. 2013. Pursuits: spontaneous interaction with displays based on smooth pursuit eye movement and moving targets. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. ACM, 439–448.

Colin Ware and Harutune H. Mikaelian. 1987. An Evaluation of an Eye Tracker As a Device for Computer Input2. 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

Shumin Zhai, Carlos Morimoto, and Steven Ihde. 1999. Manual and Gaze Input Cascaded (MAGIC) Pointing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99)*. ACM, New York, NY, USA, 246–253. https://doi.org/10.1145/302979.303053

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