Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia (MUM '16), December 12–15, 2016, Rovaniemi, Finland

EyeVote in the Wild: Do Users bother Correcting System Errors on Public Displays?

Mohamed Khamis[1], Ludwig Trotter[1], Markus Tessmann[1], Christina Dannhart[1], Andreas Bulling[2], Florian Alt[1]
[1] LMU Munich, Germany
[2] Max Planck Institute for Informatics, Saarbrücken, Germany
{mohamed.khamis, florian.alt}@ifi.lmu.de, bulling@mpi-inf.mpg.de

ABSTRACT
Although recovering from errors is straightforward on most interfaces, public display systems pose very unique design challenges. Namely, public display users interact for very short amounts of times and are believed to abandon the display when interrupted or forced to deviate from the main task. To date, it is not well understood whether public display designers should enable users to correct errors (e.g. by asking users to confirm or giving them a chance correct their input), or aim for faster interaction and rely on other types of feedback to estimate errors. To close this gap, we conducted a field study where we investigated the users willingness to correct their input on public displays. We report on our findings from an in-the-wild deployment of a public gaze-based voting system where we intentionally evoked system errors to see if users correct them. We found that public display users are willing to correct system errors provided that the correction is fast and straightforward. We discuss how our findings influence the choice of interaction methods for public displays; interaction methods that are highly usable but suffer from low accuracy can still be effective if users can "undo" their interactions.

ACM Classification Keywords
H.5.2. Information Interfaces and Presentation: User Interfaces – Input Devices and Strategies

Author Keywords
Public Displays; Smooth Pursuit; Gaze Interaction; Voting

INTRODUCTION AND BACKGROUND
Designing interfaces for interactive public displays is often associated with unique challenges. Previous work has shown that public display users interact for short amounts of time [2, 15], and usually abandon displays if response time is slow [5] or interaction is interrupted [22].

MUM '16, December 12 - 15, 2016, Rovaniemi, Finland

ISBN 978-1-4503-4860-7/16/12…$15.00
DOI: http://dx.doi.org/10.1145/3012709.3012743

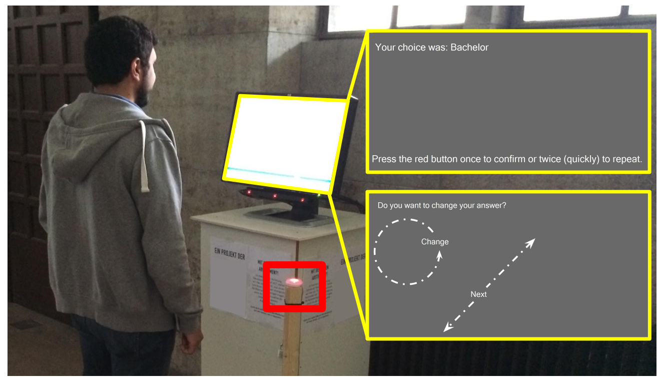

Figure 1. We investigate the willingness of public display users to correct system errors. We deployed a public gaze-based voting system, EyeVote, that occasionally shows intentionally falsified feedback and prompts users to correct their vote. In a two-days field study we experimented with two methods that allow users to correct cases where the system recognizes a selection other than the one the user intended on a situated public display. On the first day we deployed a button-based correction approach (top), while on the second day we deployed a gaze-based correction approach (bottom). Results show that most participants do correct system errors.

At the same time, users of public displays often need to deal with errors. This includes situations where users might want to correct typos when providing feedback or undo a selection.

This motivated us to study the willingness of users to correct their input on public displays. The question of whether to enable public display users to correct their input, or if they are not motivated to do so and hence designers should rely on other means for detecting errors, has not been addressed by prior research before.

Most relevant to our work is the work by Kukka et al. [13] who found that public display users are willing to dismiss error messages and continue interaction if the messages give users an active role (e.g. Press OK to continue). Our work builds over that by understanding if users bother correcting system errors rather than abandoning the display in frustration.

In this work we study how users behave when correcting errors and how the design of feedback mechanisms can assist in error correction. To do so, we deployed a gaze-based voting application on a public situated display (see Figure 1) in which we occasionally showed intentionally falsified confirmation messages to investigate whether users are willing to correct

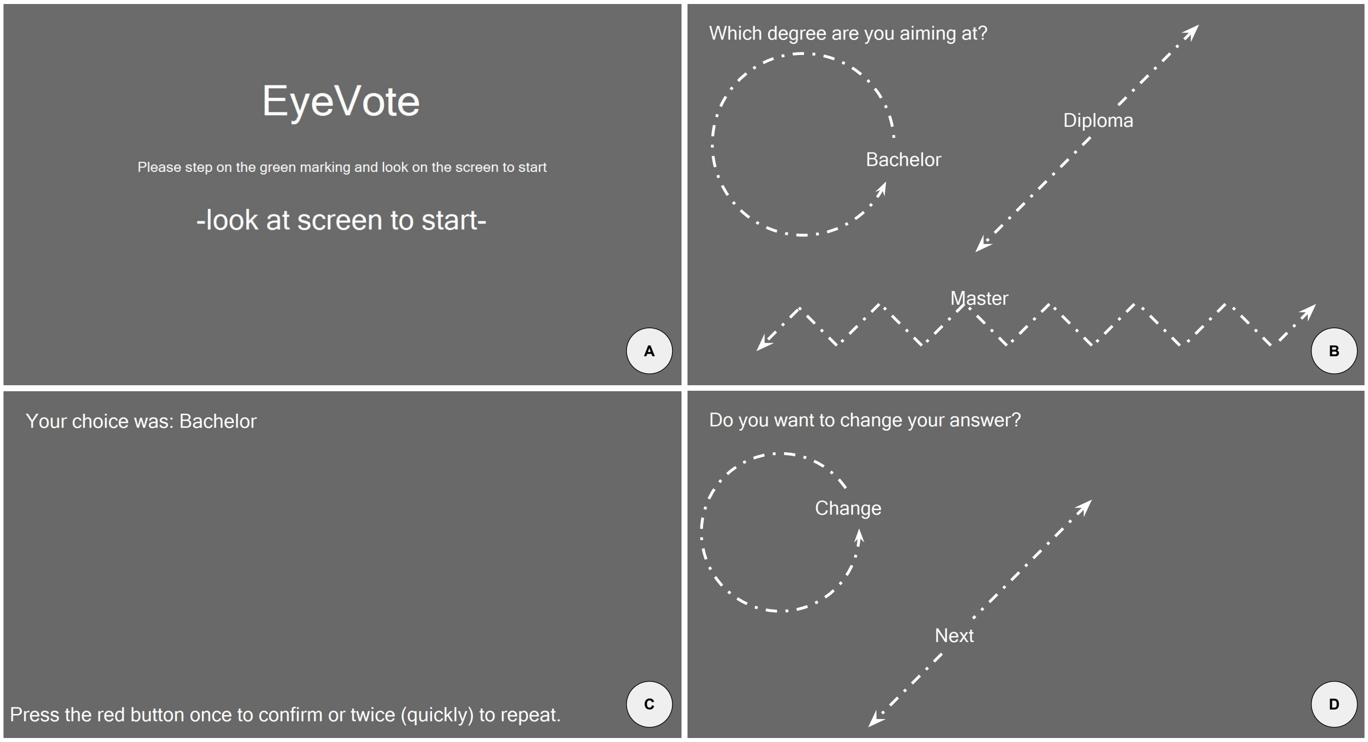

Figure 2. The figure above shows the different views of the EyeVote system. The white dashed arrows are to illustrate the trajectory of the options and were not shown to users. (A) The default view featured a call-to-action label asking participants to "look at the screen to start". (B) After the eyes have been detected, the system shows the first question. The recap view shows the answer detected by the system, in addition to the option to undo the choice. On day 1 users could undo using a button (C). While on day 2 they could undo via gaze (D).

system errors, or if they will abandon the display instead. Using a voting app was motivated by recent research, showing that public displays are a promising medium for collecting votes and opinions [6, 7, 12], encouraging civic discourse [17, 19], and reflecting the population's opinions [16]. Existing systems used touch [6], mid-air gestures [16, 19], feet [17] and recently also gaze [12] for voting in public.

We deemed gaze as interaction modality to be particularly interesting for our research [11], since it supports fast and natural interaction [23] and is hard to observe by others. The latter property is promising since making voters' choices less obvious to surrounding users was shown to be a desired quality in public voting systems [19]. At the same time, gaze interaction is prone to input recognition errors, making it a suitable candidate for investigating users' error correction behavior and the employment of suitable features for error correction.

To obtain a better understanding of how error correction can be implemented and how users react to it, we introduce an undo feature to our voting app that allows users to correct their input by means of different modalities, i.e., gaze and physical touch. We then deployed the system in the real world setting for two days. Findings show that users correct system errors given that the correction method is fast and straightforward.

This paper contributes an investigation of passersbys' willingness to correct errors caused by faulty system detections.

VOTING SYSTEM

EyeVote is a voting system for gaze-enabled public displays. The system displays questions to the public, with multiple floating textual options to choose from (see Figure 2B).

Collecting accurate gaze points requires time-consuming calibration, which is unaffordable on public displays [9]. It was necessary to either use a calibration-free technique (e.g. [21, 24, 18]) or blend the eye tracker calibration into interaction (e.g. [12]). EyeVote uses Pursuits, a calibration-free gaze-interaction technique that relies on correlating the user's eye movements with movements of dynamic stimuli on the screen. The moving stimulus that has the highest correlation to the eye movements, is deemed to be what the user is looking at.

Previous work experimented with different types of stimuli for Pursuits, such as images [21], icons [4] abstract objects [20], game elements [8], letters [14] and digits [1]. Only recently, it was found that users are able to pursue moving text while reading it [12], which means that text strings can also be used as stimuli for the Pursuits approach.

Pursuits Implementation

The system uses the Pearson's product-moment coefficient to calculate the correlation between the eyes movements and the movements of the floating options. Based on pilot experiments and previous work [4, 8, 12, 21], we used the same correlation function with a threshold of 0.7 and a window size of 2000 ms. This means that every 2 seconds, the system computes Pearson's correlation. The floating answer whose movement correlates the most with the eye movement, is deemed to be the object the user is looking at, as long as the correlation is more than 70%. Note that previous systems used more conservative thresholds (e.g. Khamis et al. [12] used a threshold of 90% in their lab study of EyeVote). A larger threshold means that the system will not accept an answer until a high correlation is found, which increases accuracy but might result in spending longer times until the system responses. In our study, we opted for a lower value to increase the responsiveness of the system.

System Walkthrough and Undo Feature

By default, a call-to-action label invites passersby to "Look at the screen to start" (Figure 2A). Once gaze is detected, the interface shows the first question (Figure 2B). After the system correlates the users choice, it proceeds to a recap view, where the detected answer is shown to the user. At this point the user is presented which the choice of changing the detected answer, or proceeding to the next question. On the first day we showed the button-condition (Figure 2C), while the gaze-condition was shown on the second day (Figure 2D). Depending on the user's choice, the system either proceeds to the following question or repeats the last question. After completing 8 questions, the system resets to the default view. If the system loses track of the user's eyes it shows a warning message indicating that eyes are not detected. If eyes are not detected for 8 continuous seconds, the system restarts. The system showed straight forward questions about the favorite band, the user's study program, etc.

Falsified answers

Due to the nature of field studies, it was not feasible to have an experimenter ask every participant whether or not the system detected their vote correctly. Hence, to investigate whether users undo system mistakes, we intentionally introduced *falsified* answers. This was done by showing an answer in the recap view that was not among the options the user had available to choose from. This way we are confident that the system is showing the user a wrong answer, and that the expected behavior is to undo the choice. In every set of 8 questions, fake answers were always shown the first time questions 3 and 6 are answered. For example, the first time a user answers question 3, the system shows "Asics" in the recap view even though it was not among the options. If that user decides to "Change" the answer, the system shows the question again, but this time the system shows the answer that was really detected.

FIELD STUDY

We deployed a 27 inch display (1920×1200 pixels) in large university hall that is expected to be reached by many students, academics, and university staff members. The display was equipped with a Tobii EyeX Controller. A squared marker was placed on the floor with a distance of one meter to the display to guide passersby to standing in the eye tracker's range.

Design

The study ran for two days and covered two conditions: (1) button-based undo, and (2) gaze-based undo.

On the first day, we deployed the button-based undo approach. A red button was placed next to the display, participants were asked to press the button once to proceed to the next question, or twice to repeat the question (see Figures 1 and 2C).

On the second day, we deployed the gaze-based undo approach. After each question, participants were asked if they want to change their answer (Figure 2D), and were shown two floating textual options that can be selected via Pursuits. Following the circular trajectory of the word "Change" would repeat the last question again. While following the linear trajectory of the word "Next" would show the following question.

Results

We logged the raw gaze data, all presented and selected answers, as well as button and gaze based undos. The system launched the first question 187 times in total during two days. This means that there were 187 instances where the system detected a user standing within the marked area and facing the screen. We refer to these instances as "interactions".

On Day 1 there were 106 interactions, that is, gaze was detected at 106 different instances. Out of which, there were 49 instances where at least one question was answered. In total, 220 questions were answered on the first day. On Day 2, at least one question was answered in 30 out of the 81 times in which a user interaction was detected. In total, 243 questions were answered on the second day.

Undos

We distinguish two cases where the "undo" feature was used by users: (1) cases where users corrected falsified answers, and (2) cases where users corrected unaltered answers. The former are cases where we are confident that the system showed a wrong answer that is was not among the choices, while the latter are cases where the system might or might not have shown a wrong answer (see Figure 3).

Out of the 220 questions answered on Day 1, 37 were falsified answers. Out of those, 22 answers (59.5%) were changed using the undo feature. While on Day 2, there were 31 falsified answers out of 243 answered questions. Users corrected 27 out of 31 (87%) falsified answers.

Note that users did not have any motivator to correct their input apart from their intrinsic motivation. Therefore, we interpret the correction rates of 59.5% and 87% as indicators that the users are willing to correct input mistakes.

The undo rate for the gaze method is higher compared to the button method. We attribute this to the use of the same interaction method for selection and correction which is seemingly better accepted by the users.

User dropouts

As this field study was not supervised, users could join and leave at any time. We relied on the number of dropouts as a measure of satisfaction and how users cope with the system. We define a dropout as a situation where the system lost eye contact for more than 8 continuous seconds.

Figure 4 illustrates the number of users who answered each question, as well as the number of users who dropped out after answering each question. It is noticeable that 20 out of 29 users (41%) dropped out after answering the first question in

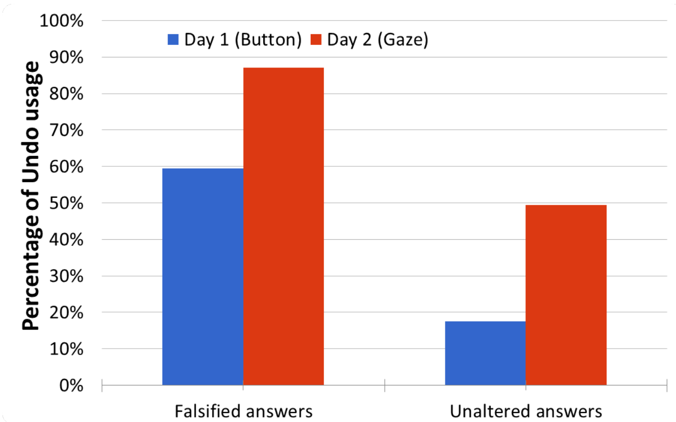

Figure 3. The figure illustrates the relative undos for falsified and unaltered questions for each day. The figure shows that users corrected the majority of the falsified answers on both days (59.5% and 87% respectively). Gaze-based undo was used more frequently than the button-based approach, we believe that this is due to the use of gaze for both selection and correction, which was better perceived by users.

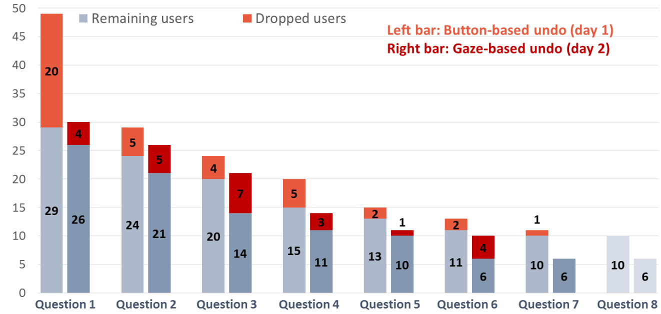

Figure 4. The figure reflects the number of dropped-out users; for example, on the first day 49 users answered question 1, 20 of which dropped out afterwards. We suspect the lower number of dropouts at the first question in day 2 is due to using the same modality for selection and confirmation. While the relatively higher dropout rates in the second day at the falsified questions (questions 3 and 6), is likely due to eye fatigue.

the first day in which the button-based undo was deployed. While on the second day, relatively few users dropped out after the first question, but relatively more dropped out after questions 3 and 6, in which the falsified answers were shown.

Interestingly, the number of users who answered all eight questions is very similar on both days: 20.4% on Day 1, and 20% on Day 2.

DISCUSSION

To our surprise, many users (59.5% and 87% on both days respectively) corrected their input when we displayed falsified answers. Although prior work reports that people are likely to abandon the display if they find it unresponsive or faulty [2]. Our results indicate that passersby are willing to fix system mistakes as long as the option is available and feasible.

By examining Figure 4, we expect that the high dropout rate for the first question on the first day is due to the use of two separate modalities for selection and confirmation; we used gaze for voting, and the button for confirming the answer. On the other hand, although there are less dropouts at the first question on the second day, there are more dropouts at the falsified answers. By the time users reached the falsified questions (questions 3 and 6) they had already answered and confirmed many questions via gaze. Hence we expect that the temporal demand led to fatigue, which discouraged some users from completing the 8 questions. Previous work had reported that performing multiple consecutive gaze-based selections is tiring to the users [12].

The results suggest that users are indeed willing to correct system errors. Figure 4 shows that although relatively more users dropout at system errors, the majority of users corrects them and continues interacting. With this conclusion we encourage the use of highly usable interaction techniques even if this leads to sacrificing some system accuracy, and relying on users to correct the occasional system errors. Taking EyeVote as an example, future systems can offer a *dynamic undo function*, which can be realized by introducing an additional threshold; if the highest correlation between the eye movements and one of the trajectories is higher than the selection threshold (set to 70% in our implementation), a more conservative threshold is checked (e.g. 90%) and if the correlation is lower than that (i.e. between 70% and 90%) the user is presented with a confirmation message, otherwise the user proceeds to the following question.

Although the use of a button to confirm or revoke an interaction is intuitive, the results suggest that users were more likely to dropout in the case of the button-based undo compared to the gaze-based approach. We believe that the use of the same modality for interaction and confirmation maintains a straightforward flow and is less confusing to the passersby. Hence we recommend future systems to use the same modality for both interaction and confirmation.

While we investigated the willingness to correct input on a gaze-enabled display, the results are also applicable to other modalities. For example, systems can reduce dwell times when selecting via mid-air hand gestures; this increases the responsiveness of the system but makes it more error prone, hence the system should also allow users to undo their actions.

CONCLUSION AND FUTURE WORK

Our results indicate that users will tolerate input errors and correct them if the system allows that. Therefore, we encourage the use of highly usable metrics even if they reduce accuracy provided that correction mechanisms are implemented. In our study, users corrected most of the falsified inputs (59.5% for the button-based undo and 87% for the gaze-based undo). Although the dropout rate was higher when falsified answers were shown during the gaze-based undo condition, the majority of the users corrected them before proceeding to the following question.

In this work we evaluated an undo feature for public displays using Pursuits. One direction for future work is to experiment with other gaze-input methods such as gaze gestures [3, 10]. Additionally, more modalities can be experimented with and compared such as mid-air hand gestures, touch and head gestures (e.g. nodding to confirm interactions).

REFERENCES

1. Dietlind Helene Cymek, Antje Christine Venjakob, Stefan Ruff, Otto Hans-Martin Lutzand Simon Hofmann, and Matthias Roetting. 2014. Entering PIN codes by smooth pursuit eye movements. *Journal of Eye Movement Research* 7(4):1 (2014), 1–11.

2. Nigel Davies, Sarah Clinch, and Florian Alt. 2014. *Pervasive Displays: Understanding the Future of Digital Signage* (1st ed.). Morgan & Claypool Publishers.

3. Heiko Drewes and Albrecht Schmidt. 2007. *Interacting with the Computer Using Gaze Gestures*. Springer Berlin Heidelberg, Berlin, Heidelberg, 475–488. DOI: http://dx.doi.org/10.1007/978-3-540-74800-7_43

4. Augusto Esteves, Eduardo Velloso, Andreas Bulling, and Hans Gellersen. 2015. Orbits: Gaze Interaction for Smart Watches Using Smooth Pursuit Eye Movements. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (UIST '15)*. ACM, New York, NY, USA, 457–466. DOI: http://dx.doi.org/10.1145/2807442.2807499

5. Miriam Greis, Florian Alt, Niels Henze, and Nemanja Memarovic. 2014. I Can Wait a Minute: Uncovering the Optimal Delay Time for Pre-moderated User-generated Content on Public Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1435–1438. DOI: http://dx.doi.org/10.1145/2556288.2557186

6. Luke Hespanhol, Martin Tomitsch, Ian McArthur, Joel Fredericks, Ronald Schroeter, and Marcus Foth. 2015a. Situated Interfaces for Engaging Citizens on the Go. *interactions* 23, 1 (Dec. 2015), 40–45.

7. Luke Hespanhol, Martin Tomitsch, Ian McArthur, Joel Fredericks, Ronald Schroeter, and Marcus Foth. 2015b. Vote as you go: Blending interfaces for community engagement into the urban space. In *7th International Conference on Communities and Technologies (C&T)*. ACM, University of Limerick, Limerick, Ireland, 29–37.

8. Mohamed Khamis, Florian Alt, and Andreas Bulling. 2015. A Field Study on Spontaneous Gaze-based Interaction with a Public Display Using Pursuits. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers (UbiComp/ISWC'15 Adjunct)*. ACM, New York, NY, USA, 863–872. DOI: http://dx.doi.org/10.1145/2800835.2804335

9. Mohamed Khamis, Florian Alt, and Andreas Bulling. 2016a. Challenges and Design Space of Gaze-enabled Public Displays. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct (UbiComp '16)*. ACM, New York, NY, USA, 1736–1745. DOI: http://dx.doi.org/10.1145/2968219.2968342

10. Mohamed Khamis, Florian Alt, Mariam Hassib, Emanuel von Zezschwitz, Regina Hasholzner, and Andreas Bulling. 2016b. GazeTouchPass: Multimodal Authentication Using Gaze and Touch on Mobile Devices. In *Proceedings of the 34th Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. ACM, New York, NY, USA, 6. DOI: http://dx.doi.org/10.1145/2851581.2892314

11. Mohamed Khamis, Andreas Bulling, and Florian Alt. 2015. Tackling Challenges of Interactive Public Displays Using Gaze. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers (UbiComp/ISWC'15 Adjunct)*. ACM, New York, NY, USA, 763–766. DOI: http://dx.doi.org/10.1145/2800835.2807951

12. Mohamed Khamis, Ozan Saltuk, Alina Hang, Katharina Stolz, Andreas Bulling, and Florian Alt. 2016. TextPursuits: Using Text for Pursuits-Based Interaction and Calibration on Public Displays. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. ACM, New York, NY, USA, 12. DOI: http://dx.doi.org/10.1145/2971648.2971679

13. Hannu Kukka, Jorge Goncalves, Tommi Heikkinen, Olli-Pekka Suua, Yifei Zuo, Hannu Raappana, Mohamed Abdellatif, Olli Okkonen, Raúl Jiménez, and Timo Ojala. 2015. Touch OK to Continue: Error Messages and Affective Response on Interactive Public Displays. In *Proceedings of the 4th International Symposium on Pervasive Displays (PerDis '15)*. ACM, New York, NY, USA, 99–105. DOI: http://dx.doi.org/10.1145/2757710.2757723

14. Otto Hans-Martin Lutz, Antje Christine Venjakob, and Stefan Ruff. 2015. SMOOVS: Towards calibration-free text entry by gaze using smooth pursuit movements. *Journal of Eye Movement Research* 8(1):2 (2015), 1–11.

15. Jörg Müller, Robert Walter, Gilles Bailly, Michael Nischt, and Florian Alt. 2012. Looking Glass: A Field Study on Noticing Interactivity of a Shop Window. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 297–306. DOI: http://dx.doi.org/10.1145/2207676.2207718

16. Erica C. Ostermann, Long Ma, Daniel Sussman, and Susan R. Fussell. 2015. CommunityConnect: An Interactive Display for Educational Residential Settings. In *Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing (CSCW'15 Companion)*. ACM, New York, NY, USA, 175–178. DOI: http://dx.doi.org/10.1145/2685553.2699000

17. Fabius Steinberger, Marcus Foth, and Florian Alt. 2014. Vote With Your Feet: Local Community Polling on Urban Screens. In *Proceedings of The International Symposium on Pervasive Displays (PerDis '14)*. ACM, New York, NY, USA, Article 44, 6 pages. DOI: http://dx.doi.org/10.1145/2611009.2611015

18. Yusuke Sugano, Xucong Zhang, and Andreas Bulling. 2016. AggreGaze: Collective Estimation of Audience Attention on Public Displays. In *Proc. ACM Symposium on User Interface Software and Technology (UIST)*.

19. Nina Valkanova, Robert Walter, Andrew Vande Moere, and Jörg Müller. 2014. MyPosition: Sparking Civic Discourse by a Public Interactive Poll Visualization. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW ’14)*. ACM, New York, NY, USA, 1323–1332. DOI:http://dx.doi.org/10.1145/2531602.2531639

20. Eduardo Velloso, Markus Wirth, Christian Weichel, Augusto Esteves, and Hans Gellersen. 2016. AmbiGaze: Direct Control of Ambient Devices by Gaze. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS ’16)*. ACM, New York, NY, USA, 812–817. DOI: http://dx.doi.org/10.1145/2901790.2901867

21. 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 (UbiComp ’13)*. ACM, New York, NY, USA, 439–448. DOI:http://dx.doi.org/10.1145/2493432.2493477

22. Robert Walter, Gilles Bailly, and Jörg Müller. 2013. StrikeAPose: Revealing Mid-air Gestures on Public Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ’13)*. ACM, New York, NY, USA, 841–850. DOI: http://dx.doi.org/10.1145/2470654.2470774

23. 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. DOI: http://dx.doi.org/10.1145/302979.303053

24. Yanxia Zhang, Andreas Bulling, and Hans Gellersen. 2013. SideWays: A Gaze Interface for Spontaneous Interaction with Situated Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ’13)*. ACM, New York, NY, USA, 851–860. DOI:http://dx.doi.org/10.1145/2470654.2470775