Perceptions of Interfaces for Eye Movement Biometrics

Michael Brooks[1], Cecilia R. Aragon[1], and Oleg V. Komogortsev[2]
[1]University of Washington, Seattle, WA,
[2]Texas State University, San Marcos, TX
{mjbrooks, aragon}@uw.edu, ok11@txstate.edu

Abstract

Widespread adoption of biometric authentication technology has been slow because of technical, usability, and acceptability issues. Incorporating feedback from stakeholders throughout the design process is one way to ensure that biometric technologies have the best chance of achieving success. Emerging technology that performs biometric identification through unique eye movement characteristics may offer advantages, but the potential usability and acceptability of interfaces for eye movement biometric authentication is unknown. We explore a series of user interface designs for a simulated eye movement biometric authentication system, presenting an account of our design process, descriptions of the interface designs, and our evaluation. Our findings highlight the importance of feedback design, potential usability issues, users' perceived security benefits, and reasons why users may prefer eye movement biometrics over other systems.

1. Introduction

Passwords, despite enjoying widespread use, have profound and well-documented usability problems that result in security vulnerabilities and high economic costs [1–3]. In the search for alternative authentication solutions for the future, biometric authentication systems have promise because they do not require users to remember a secret or carry a token [4].

Widespread adoption of biometric authentication faces challenges. For authentication, a biometric must be sufficiently universal, distinctive, permanent, quickly and accurately measurable, socially acceptable, and resistant to fraudulent access, a challenging set of criteria [5]. Many of these criteria depend closely on people's ability and willingness to use the technology, and previous work has demonstrated that user interaction with the system has a significant impact on system performance [6–8].

Recent large-scale biometric deployments such as the Unique Identification Authority of India (UIDAI), an ongoing project to issue biometric identification to 1.2 billion Indian citizens [9], demonstrate the pressing need for additional research on usability and user perceptions of biometric technology. Usability and acceptability concerns are sometimes ignored or left until late in security system development, and many have advocated placing a high priority on usability and human-centered design in the early stages of security development [10–12].

To that end, we have designed and studied a set of prototype user interfaces for *eye movement biometric authentication*. Recent advances in biometric technology based on distinctive patterns of eye movement may offer advantages such as resistance to counterfeiting and combination with other biometric modalities [13–15]; these and other dynamic biometrics could soon have a significant impact on people around the world.

Eye movement biometric technology is still too early to be tested in a working system, so our user interfaces are embedded in a *simulated authentication system*, which nevertheless provides useful information about usability issues and user perceptions. This is the first work to directly address usability and acceptability of eye movement biometric technology. Our research has uncovered factors affecting user perceptions towards eye movement biometrics and biometrics in general, and points to several areas that warrant further study.

The main contributions of this paper are 1) insights gained from introducing users to these prototypes 2) a detailed account of our user centered design process, and 3) demonstration of the utility of studying how users interact with security systems early in the research and development process.

2. Background and Related Work

We review research on biometric authentication, gaze-based authentication, eye movement biometrics, and ATM authentication.

2.1. Biometric Authentication

A wide variety of physical and behavioral characteristics have been successfully used for biometric identification [16]. Of these, fingerprint, iris, and face recognition have received the most attention recently [5], and there have been large-scale deployments, such as UIDAI's biometric identification for India's 1.2 billion citizens [9].

However, there are significant barriers to widespread adoption of biometric authentication technology, including usability and accessibility. The ways in which people interact with a biometric system can shape the overall security and performance of the system [6], [7], [17]. For many people, biometric technologies also raise privacy concerns. For example, in some cases biometrics may allow for the possibility of covert identification and data misuse by governments and other organizations [10], [18]. Designers of biometric systems must consider the social and cultural acceptability of their technologies.

Evaluation of biometric systems commonly measures effectiveness in terms of false acceptance and false rejection rates; much more work is needed that attends to usability, user satisfaction, and acceptability [19].

2.2. Eye Movement Biometrics

To our knowledge, the potential of eye movement as a biometric identification technique was first demonstrated by Kasprowski and Ober [20]. This technique uses an eye tracker to record samples of the user's gaze trajectory over a period of time while the user observes a visual stimulus. Research has focused on decreasing the recognition error rate by developing salient features of gaze data and evaluating classification algorithms for distinguishing individuals [13], [14], [21–23]. Although recognition error rates for this technology remain high, recent progress has been promising.

Bednarik et al. found that the dynamics of pupil size were the most promising feature of those tested, with a correct identification rate of around 60% in their study population of 12 [23]. Komogortsev et al. explored a different approach to using eye movement for identification that is based on the unique eye globe muscle parameters, or oculomotor plant, of each individual. Their algorithm uses eye movement data to determine characteristics of the oculomotor plant, such as elasticity, length tension, force velocity, and active tension. A recent evaluation of these techniques with a group of 59 individuals achieved a half total error rate of 19% [14]. These studies found that eye movements can be leveraged to identify individuals; we present a complementary user-focused perspective.

2.3. Authentication at ATMs

The usefulness and validity of feedback obtained from usability studies often increases as the context, scenario, and system become more realistic, convincing, and relatable. Therefore, although eye movement biometrics may have wide applicability, in this work we focus on authentication at ATMs (Automated Teller Machines). ATMs provide a use case that study participants can easily relate to: we expected most people to be familiar with using ATMs, and to readily understand the security issues around ATMs. Prior studies on biometric technologies at ATMs emphasized testing with functional prototypes for getting realistic user perceptions [24]. Therefore, our methodology relies on prototype-based studies at various levels of fidelity. Of course, our goal was not to develop an authentication solution meeting the specific requirements inherent to public authentication [25], so we explored interface designs that provided useful insights about eye movement biometrics in general.

3. Prototype Design

Our process for designing user interfaces for an eye movement biometric authentication system involved two stages of prototype development and feedback, converging on a working prototype.

We began by defining the overall flow of the eye movement authentication process. ATM use begins with a Welcome message. After inserting a bank card (effectively claiming an identity), the Calibration step begins, where the eye tracker is calibrated to the user. In the Verification step, the system collects eye movement data for recognition. Finally, the user is either authorized or denied. While all of these steps are important to providing an authentication experience, the Verification step involves the most important design problems for this research; therefore, our efforts focused on the Verification step.

The technical requirements for the verification interface are still in flux as the technology matures, but Holland & Komogortsev have found that the specific visual stimulus used to collect gaze data has only a small effect on biometric recognition accuracy [26]. Nevertheless, the overall interface including the visual stimulus has an impact on usability and acceptability.

Below we discuss the evolution of our designs for the verification interface from low fidelity prototypes to an interactive software prototype.

3.1. Low Fidelity Prototypes

Out of a large collection of initial ideas for verification interface designs, we selected three designs with promising usability. One version included a basic QWERTY onscreen keyboard, with which the user "eye-types" their name, using gaze to activate the keys. The second version showed the letters of the user's first name arranged in a 9-point grid pattern. The user gazes at each of the 9 letters one-by-one. In the third version, the user also enters their name, but here the alphabet was presented on a vertically scrolling bar from which letters are selected using gaze. The designs used a dwell-time mechanism to activate interface elements [27], [28].

Usability studies conducted with prototypes made of paper can uncover design flaws before an expensive working prototype is built [29]. We created prototypes from paper and cardboard for each of the three verification interface designs and conducted an exploratory usability

study with 9 participants. Participants were mostly students and staff recruited from technology- and engineering-focused departments at our university. Participants were asked to authenticate and complete cash withdrawals using the paper ATM, with each of the three different verification interface designs. We used a concurrent think-aloud protocol and semi-structured interview to discover usability issues and understand user preferences [29].

From this study we learned that the vertically scrolling alphabet selector was too unfamiliar and difficult to control for most participants, so we removed it from future iterations. Participants stressed the importance of speed, so we reduced the number of distinct screens in the authentication process by simplifying on-screen instructions and integrating them into the interfaces where they were required. Participants were concerned about their ability to comfortably control the interface using their gaze, so we put effort into providing clear visual feedback, which is known to be important in interfaces using gaze interaction [30].

3.2. High Fidelity Prototypes

We developed a high fidelity software prototype in C# to run on a Windows PC connected to a Tobii X120 eye tracker, which records gaze data at 120 Hz (Figure 1).

As we refined the working prototype, we added carefully tuned visual feedback such as a red dot that smoothly follows the user's gaze and animations that make gaze-based button activation easy and predictable. We also created several new alternative designs. Once the software was testable, we brought 8 more participants into the lab to try out 5 different verification stimulus designs. Participants were again asked to withdraw money from the ATM while thinking aloud. After going through all of the prototypes, we asked follow-up questions to understand problems encountered.

Out of the five designs that we tested, one interface in particular emerged as a good balance of both usability and perceived security for most participants (this was the basis for the *Targeting* design, described later and shown in Figure 2). We included this design in the final prototype and added one new design, where the user reads a selection of text to provide eye movement data, as described in [31].

3.3. Final Prototype Interface Design

The prototype resembles a typical ATM interface featuring a central display screen surrounded by a plastic housing, a dark blue screen background, and white text and interface elements for reading ease. Below the screen is a bank card slot. Other physical components such as a cash dispenser, deposit receptacle, receipt printer, or physical keypad, were omitted.

Users first see the welcome screen with instructions to insert a bank card. The system verifies that the user's eyes are visible by recording a small number of gaze samples. If the eyes are outside a preset acceptable region or untracked, then a positioning screen appears (similar to the built-in Tobii eye tracker head positioning interface). Instructions may appear to move closer or away.

Figure 1: The high fidelity prototype, showing a welcome screen. The eye tracker is below the monitor.

Next, the system runs a 5-dot calibration procedure. The results of the calibration are collected *via* the Tobii SDK: if less than 80% of the calibration samples were usable, or if the samples were too far from the expected calibration targets, an error screen appears showing the head positioning widget. Following calibration, the verification screen appears with a visual stimulus, such as the onscreen keyboard or typing pad, while gaze data is recorded for authentication.

We developed two different variants of the stimulus, *Targeting* and *Reading*, described below. If the percentage of valid samples collected is below 80%, another error screen appears. The user may attempt the authentication again, or choose to recalibrate. Following collection of gaze data, a "Verifying…" message and progress bar appear for 1-2 seconds, giving the appearance of processing the authentication (although in fact the decision is preset by the study moderator). If the user is not recognized, then an error screen appears that allows another attempt.

3.4. Interface Variations

In the *Targeting* variant of the verification stimulus, the system displays the message "Gaze at the highlighted circle" and displays a grid of nine circles (Figure 2). Each circle contains a small "plus" icon to help center the user's gaze. All of the circles are shaded except for one, which is highlighted. A small red dot on the screen follows the user's gaze. When the red dot enters the highlighted circle, the circle begins to fade to green, and the white border of the circle thickens slightly. After 0.5 seconds, the button flashes white, and fades, having been activated. Another circle becomes highlighted, and the process repeats. Each

circle must be activated once, in a random sequence, to complete verification.

In the *Reading* variant, the system displays "Read the following text" and then shows a selection of text. To the left of the text is a vertical "progress bar" that begins close to the top of the screen. As the user reads, the system detects where the user's gaze is located in the text and smoothly advances the bar to a matching position. Once the bar approaches the bottom of the text, it becomes solid white and a "Finished Reading" button becomes available in the lower right corner of the screen. Once the button has appeared, the user may click the button to finish.

The text displayed was selected from Lewis Carroll's "The Hunting of the Snark." We selected this text because it is difficult and nonsensical, requiring careful reading [31]. The poem is also long enough to show different selections each time a participant uses the prototype.

In order to provide a familiar point of reference for comparing the eye tracking authentication processes, we also created a traditional PIN-based authentication variant that *does not use eye tracking*. The participant uses the mouse to type their previously provided PIN on a 10-digit on-screen keypad. Although a direct comparison in performance between this variant and the biometric variants is not our intention – it is clear that PIN authentication is currently faster and simpler than eye movement biometric authentication – this baseline may help contextualize the results.

4. Evaluation

We conducted a lab study to investigate the efficiency, ease of use, and perceived security of our user interfaces for eye movement biometric security systems. Participants were asked to authenticate at the ATM and withdraw money. Our measures focused on how long people took to complete the task, how often system events like recalibration or repositioning occurred, and people's perceptions about the system's usability and security.

4.1. Experiment Design and Procedure

We compared the Tracking, Reading, and PIN interfaces in a within-subjects experiment. Participants completed 8 authentication attempts (trials) in a row on each of the three prototypes. We counter-balanced the order of the interfaces to mitigate learning effects.

Before each new interface, the test facilitator explained the authentication procedure, asking participants to imagine that the ATM they were using was operated by their bank. Participants were allowed to believe that the variants using eye tracking would genuinely be authenticating them, although they were actually configured to automatically grant access on every trial except for the 5th trial, when the system would deny access. Our hope was that this deception would result in more genuine experience and realistic reporting of frustration.

Figure 2: Targeting verification stimulus. A red dot follows the user's gaze as they activate each circle.

After the first trial for each interface, participants filled out a questionnaire designed to measure first impressions, with four items asking about difficulty, confusion, frustration, and enjoyment. Next, the participants completed 7 more trials. After the last trial, there was a second questionnaire that included the same questions as the first one, but added questions about perceived security, invasiveness, and privacy concerns.

Once all three authentication variants had been tested, participants completed a worksheet where they could rank the three authentication designs from 1 (best) to 3 (worst) with respect to Ease of Use, Security, and Overall Preference. After completing the survey, participants were interviewed about the reasons behind their rankings.

Concluding the session, the facilitator debriefed participants, explaining that the authentication systems that had appeared to identify them on the basis of their eye movements had actually been rigged to always grant access except on the 5th trial.

In addition to the two questionnaires and the ranking worksheet that participants completed, the prototype was instrumented to record system events, including overall task duration, the duration of each calibration and verification, and occurrences of error screens.

4.2. Participants

We recruited 22 participants through emails to university department mailing lists and word-of-mouth; 17 were students (13 undergraduates). Most said that they had some professional experience or interest in technology. Participant ages ranged from 17 to 36, with a mean of 26. Gender was unevenly distributed (16 male, 6 female). Although most types of corrective lenses did not appear to significantly interfere with eye tracking, 7 participants

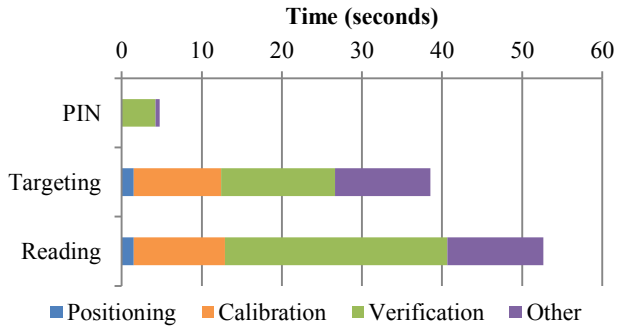

Figure 3: A breakdown of mean time spent during authentication on the three ATMs.

wore glasses and 2 wore contact lenses. Eight of the participants had previously seen or used an eye tracker.

Three of the 22 participants had been involved with earlier prototyping exercises. Prior experience with the research may have influenced these participants' reactions to the system, but their data was still useful – outside of the lab, users bring a variety of unpredictable backgrounds, skills, and interest levels to their interaction with biometrics. Additionally, the interface had changed significantly since earlier versions of the system; these 3 participants experienced the same kinds of usability issues as other participants.

5. Results

Below we provide quantitative and qualitative results from the usability evaluation, including task completion and errors, authentication time, ease of use, and perceived security and acceptability.

5.1. Authentication Time

The time taken to use the prototype showed a pronounced difference between eye tracking authentication and the traditional PIN system. In creating the prototype, we made many design decisions that affected the authentication time, based on reasonable assumptions about the data and processing requirements of eye movement biometric technology. Of course, these assumptions may be inaccurate. For example, the decision to use 3 stanzas of poetry in the Reading prototype and 9 targets in the Targeting prototype both have an impact on the time required to authenticate.

The timing results are not intended to be a demonstration that the system is, or is not, fast enough to be usable, but rather to provide a ballpark figure for how long the prototype takes to use given the design decisions that we made. It is also important to consider the effects that overhead processes such as head positioning, eye tracker calibration, and errors have on the total time, and how these might be reduced, as in [32]. The comparison with a PIN variant serves to provide a commonly-understood baseline for interpreting the timing of the other variants.

We analyzed timing *via* a mixed-effects model analysis of variance, taking *Interface* and *Trial* as repeated measures, while *Participant* was modeled as a random effect because the levels of this factor were drawn randomly from a population. While the denominator degrees of freedom are higher for this type of analysis, statistical significance is no easier to achieve because wider confidence intervals are used [33]. There was a significant difference in the total times ($F(2,105.6)=2608.5$, $p < 0.001$). Users took an average of 4.7 ± 2.4 seconds to authenticate with a PIN, 39 ± 13 seconds with the Targeting prototype, and 53 ± 21 seconds with the Reading prototype.

Figure 3 shows the breakdown of how this time was spent in different parts of the process. The PIN design contained no steps aside from verification, but the other two ATMs included time for positioning and calibration. There was no significant difference in the time spent on *positioning* or on *calibration* between the two eye tracking variants, as expected. Calibration took an average of 11 ± 2.8 seconds while positioning took 1.5 ± 4.8 seconds. Verification on the Targeting prototype took an average of 14 ± 9 seconds, while on the Reading prototype, verification took 28 ± 15 seconds. This difference was significant ($F(1, 92.3) = 882.109$, $p < 0.001$). The remaining time (*Other*, in Figure 3) was spent on error screens and transitions, and was not significantly different between the two eye tracking ATMs.

5.2. Perceived Security and Acceptability

Participants answered a set of scale questions about each verification stimulus: confidence in the ATM's security, how personally invasive using the ATM felt, concern that their data could be misused, and risk of covert identification. Perhaps because the questions asked about abstract, unfamiliar concepts, or because each interface was rated in isolation, there were no significant differences detected between the three designs.

Table 1: Number of users who assigned each rank to the three designs for Ease of use, Security, and Preference. 1st place is best, 3rd is worst.

		PIN	Targeting	Reading
Ease of use	1st	21	2	2
	2nd	1	16	4
	3rd	0	4	16
Security	1st	6	11	10
	2nd	6	10	8
	3rd	10	1	4
Overall Preference	1st	6	12	6
	2nd	12	6	3
	3rd	4	4	13

After having seen all of the interfaces, participants comparatively ranked the Security of the three designs. These rankings were significantly different according to a Friedman test ($\chi^2(2) = 6.026$, $p = 0.049$). Participants ranked the two eye tracking variants higher for security than the PIN variant, but differences between each pair were not significant in a post hoc analysis using Wilcoxon Signed Ranks tests. The breakdown of rankings for Security is shown in Table 1.

When asked to explain their rankings for security, several participants said that they were unsure how to answer the question because they didn't really understand how the eye tracking biometric system worked. Many said that they ranked the security of the Targeting and Reading designs highly because they were confident that a biometric technology could be stronger than a PIN. Not everyone felt this way – one participant said that because of the recognition error, he wasn't confident in the reliability of the mapping between his biometric data and his identity: "my gaze print is not the only gaze print I have." Participants granted biometrics an assumed security advantage over PINs, but for some this confidence quickly eroded when the biometric failed to identify them reliably.

Most of the participants did not find PINs to be trustworthy and gave reasons such as the risk of shoulder surfing. Of those that did prefer the PIN for Security, some cited its familiarity. One person preferred the PIN because, unlike a biometric, it could be unique for each service that a user is registered with, reflecting a commonly cited privacy concern with biometric technologies. Several users who gave the PIN interface a low security rank also commented that because the ATM screen displayed the key pad on the screen, they felt the PIN could be more easily stolen by observers. Perhaps if the interface had used a hardware keypad for PIN entry, the PIN interface would have received higher scores on Security.

5.3. Ease of Use

We assessed ease of use through a questionnaire with 5-point scales measuring difficulty, confusion, frustration, and enjoyment. Among the three designs, Friedman tests found the ratings to be significantly different in all four areas ($p \leq 0.018$). We tested for pairwise differences using post hoc Wilcoxon Signed Ranks tests with Bonferroni corrections (using $\alpha = 0.0166$), finding that the Targeting variant was significantly more Enjoyable than the PIN variant ($Z = -3.2$, $p = 0.001$). Unsurprisingly, PINs were rated significantly less Difficult than either Targeting ($Z = -3.3$, $p = 0.001$) or Reading ($Z = -2.6$, $p = 0.009$). PINs were also rated significantly less Frustrating than both Targeting ($Z = -2.9$, $p = 0.004$) and Reading ($Z = -3.2$, $p = 0.001$). Pairwise differences for Confusion were not significant.

When asked at the end of three sessions to rank the three designs based on ease of use, most participants found the PIN easiest to use, followed by Targeting and Reading (Table 1). A Friedman test found a significant difference for Ease of Use ($\chi^2(2) = 31.0$, $p < 0.001$). Pairwise Wilcoxon Signed Ranks tests with a Bonferroni correction found that the PIN design was rated significantly easier to use than both the Targeting design ($Z = -4.2$, $p < 0.001$) and the Reading design ($Z = -4.1$, $p < 0.001$). The difference between Targeting and Reading was not significant.

When prompted to explain their rankings for Ease of Use, many participants said that they ranked the PIN design highly because it was familiar and fast. One participant said "anyone can do it" and that it did not require much focus. This is a desirable characteristic for authentication interfaces where distractions and other tasks typically demand the user's attention. The Reading variant was described as slow and difficult; some participants complained about the difficulty of the reading selection, saying that they became "detached" and were "just looking at the words" after a while. For several participants who were not native English speakers, the reading was particularly arduous. Clearly the choice of which passage to read has a significant impact on ease of use and universality of the interface.

On the Targeting variant, several participants explained that they had difficulty completing the procedure because of inaccuracy in the gaze position that made it hard to target the circles on the screen. On the other hand, participants said that the clear and intuitive feedback provided by the red gaze indicator made it easier to understand what was happening. The clear task and the way the user's progress through the task was evident made one participant feel like they had more control over the process, although another user felt that this was an "annoying hoop to jump through." In contrast, the visual feedback on the Reading variant was less immediate and some participants were confused about whether the reading "progress bar" was an indicator of their progress (it was) or was meant to be followed. Stimuli that require precise gaze control of the interface, as in the Targeting variant, become less usable if eye tracking accuracy degrades, but may also be more enjoyable to use.

5.4. Overall Preference

Most participants selected the Targeting design as the most preferred, followed by the PIN and the Reading designs (Table 1), but differences were not significant.

Most participants explained their rankings for Overall Preference as a balance of Ease of Use and Security, citing some of the reasons already discussed. Many participants said that they preferred the Targeting design because it was fun, interesting, and felt like a game, while PINs were boring. However, one user gave a low rating to the Targeting design and felt that game-like interaction was inappropriate in a banking context. Those who preferred the Reading variant found the reading itself interesting.

One participant was excited by the prospect of personalizing the reading selection. Security interfaces which incorporate interactive game-like elements and personalization may be more engaging and enjoyable, but this may not always be appropriate.

6. Discussion and Future Work

6.1. Perceptions of Eye Movement Biometrics

The prototype was set to deny access on the 5th trial, so that all participants experienced exactly one authentication failure. Some participants attempted to explain this error in the interview: one participant said that the authentication had probably failed because he had stopped to reread something in the poem. On the Reading ATM, because of the distracting red "Recording" light and the progress bar, one participant said "At times when I would get distracted to look at these things, and I would fail the authentication." Comments about errors involved generalizations instead of focusing on specific failures, suggesting that users were not precisely aware of when failures occurred or how often they occurred. These explanations were volunteered, not directly prompted by interview questions, and usually assumed that the user was responsible for errors. Biometric technologies must provide excellent feedback in order to be usable [6]. In this case, insufficient feedback around errors led to poor understanding of when and why recognition failed. Knowing why errors occur may lead users to be more lenient of occasional false rejections.

Participants' were asked to explain their ratings of the Security of each design. Several participants were simply more confident that a biometric technology could provide stronger security than a PIN – often the reason given was that PINs can more easily be stolen by observers. On the other hand, some participants trusted biometric security systems because they perceived biometrics as newer, more complex, or more sophisticated.

Aside from comments about biometric technologies in general, participants also discussed the security of the specific designs used in the study. Some participants believed that the Reading interface would be more secure because it might collect data that was more personal and nuanced than the Targeting interface. Others preferred the Targeting interface because the eye movement associated with reading could more easily be covertly captured during daily activities. Regardless of whether the user interface for the biometric system actually has a significant effect on security, interface design choices do affect *perceptions* of security in ways that may be difficult to anticipate.

6.2. Barriers for Eye Movement Biometrics

There are many barriers to be overcome before biometric identification through eye movement becomes a viable option. Some of these are technical: for most large-scale applications, identification error rates must be very low. Our participants were immediately sensitive to these issues: several asked whether an individual's eye movements are stable enough over time to be consistently recognizable (for example, when the user is tired vs. after drinking coffee).

In addition to these challenges, for many people the prospect of any type of biometric identification raises serious privacy and safety issues, which can depend on culture or religion [10]. In our study, few people mentioned such concerns, but we believe that this is due to the novelty of eye tracking and sampling bias in our participant pool.

6.3. Designing Eye Movement Biometrics

Our designs for eye movement biometric authentication interfaces relied on prototype-based studies; this was particularly helpful was in refining feedback provided by the system. It would have been difficult to design effective visual feedback without engaging in iterative evaluation with users external to the design process. Details like the precise dwell times, animation rates used by gaze-activated buttons, and the amount of smoothing used for the red gaze indicator dot play a major role in the usability of the Targeting interface, in particular. Additionally, the screens that provide feedback in error conditions turned out to require more careful design than we had anticipated. Participant comments reinforced the powerful effect that these messages and indicators had on their usage and interpretation of how the security system worked.

Our research explored ideas for what kind of visual stimulus to present to users during the verification step, where the eye movement data is recorded for recognition. After numerous iterations and redesigns, the Targeting interface seemed to be the most successful because it provides a balance of ease of use and perceived security. The highly interactive but simple game-like design was preferred by many participants who said it was fun and effortless to use. Of course, as the novelty of using an eye tracker wears off, we might expect enjoyment to decrease. There are certainly many interface designs we did not consider or investigate. In particular, we believe that the space of interactive, game-like designs for authentication systems merits future study.

7. Conclusion

Recent work on eye movement biometrics has focused on overcoming technical challenges such as recognition error rates. We have presented our initial investigation of user interfaces for eye movement biometric systems, including an account of our design process, the designs we created, and the findings from a lab usability study with 22 people. This work was the first that we are aware of to address how users perceive eye movement biometrics and what significant usability problems and design challenges

exist for this technology.

There is growing interest in eye movement biometric identification. The IEEE Fifth International Conference on Biometrics: Theory, Applications, and Systems (BTAS 2012) featured a competition on eye movement verification and identification [34]. With programs such as UIDAI increasingly applying biometric technology at a massive scale, the importance of understanding human interaction with these systems is clear.

8. Acknowledgments

This research was supported by the National Institute of Standards and Technology (NIST) under grant 60NANB10D213. We thank the Laboratory for Usability Testing and Evaluation at the University of Washington for equipment, and Professors Jacob Wobbrock and Jan Spyridakis for their assistance with statistical tests.

References

[1] J. "Jofish" Kaye, "Self-reported password sharing strategies," in *Proc. CHI 2011*, 2011, p. 2619.

[2] P. G. Inglesant and M. A. Sasse, "The true cost of unusable password policies: password use in the wild," in *Proc. CHI 2010*, 2010, pp. 383–392.

[3] A. Adams, M. Sasse, and P. Lunt, "Making passwords secure and usable," in *Proc. HCI 1997*, 1997.

[4] A. K. Jain, A. Ross, and S. Prabhakar, "An Introduction to Biometric Recognition," *IEEE TCSVT*, vol. 14, no. 1, pp. 4–20, Jan. 2004.

[5] A. K. Jain, "Biometric recognition: overview and recent advances," in *Proc. CIARP 2007*, 2007, vol. 4677, no. 2.

[6] Y. Choong, M. F. Theofanos, and G. Haiying, "Fingerprint Self-Captures: Usability of a fingerprint system with real-time feedback," in *Proc. BTAS 2012*, 2012, pp. 1–7.

[7] M. F. Theofanos, R. J. Micheals, S. Orandi, B. Stanton, and N. F. Zhang, "How the subject can improve fingerprint image quality," *Journal of Electronic Imaging*, vol. 17, no. 1, p. 011007, 2008.

[8] T. Ko and R. Krishnan, "Monitoring and Reporting of Fingerprint Image Quality and Match Accuracy for a Large User Application," *Proc. AIPR Workshop 2004*, pp. 159–164, 2004.

[9] UIDAI, "Unique Identification Authority of India." [Online]. Available: http://uidai.gov.in/.

[10] M. A. Sasse, "Usability and trust in information systems," in *Trust and Crime in Information Societies*, Robin Mansell and B. S. Collins, Eds. Edward Elgar, 2005, pp. 319–348.

[11] D. G. tom Markotten, "User-centered security engineering," in *Proc. EurOpen/USENIX NordU2002*, 2002.

[12] M. E. Zurko and R. T. Simon, "User-centered security," in *Proc. NSPW 1996*, 1996, pp. 27–33.

[13] P. Kasprowski, "Human identification using eye movements," *Praca doktorska, Politechnika Œląska*, 2004.

[14] O. V. Komogortsev, A. Karpov, L. R. Price, and C. R. Aragon, "Biometric Authentication via Oculomotor Plant Characteristics," in *Proc. ICB 2012*, 2012, pp. 1–8.

[15] O. V Komogortsev, A. Karpov, C. D. Holland, and H. P. Proença, "Multimodal Ocular Biometrics Approach: A Feasibility Study," in *Proc. BTAS 2012*, 2012.

[16] A. Albrecht, M. Behrens, T. Mansfield, W. Mcmeechan, M. Rejman-Greene, M. Savastano, P. Statham, C. Schmidt, B. Schouten, and M. Walsh, "BioVision: Roadmap for Biometrics In Europe to 2010," *Biometrics*, 2003.

[17] L. Coventry, "Usable biometrics," in *Security and usability: designing secure systems that people can use*, L. F. Cranor and S. Garfinkel, Eds. 2005.

[18] S. Prabhakar, S. Pankanti, and A. K. Jain, "Biometric recognition: Security and privacy concerns," *Security & Privacy, IEEE*, vol. 1, no. 2, pp. 33–42, Mar. 2003.

[19] D. Toledano, R. F. Pozo, A. H. Trapote, and L. H. Gomez, "Usability evaluation of multi-modal biometric verification systems," *Interacting with Computers*, vol. 18, no. 5, pp. 1101–1122, Sep. 2006.

[20] P. Kasprowski and J. Ober, "Eye movements in biometrics," *Biometric Authentication*, pp. 248–258, 2004.

[21] F. Deravi and S. P. Guness, "Gaze Trajectory as a Biometric Modality," in *Proc. Biosignals 2011*, 2011.

[22] O. V. Komogortsev, S. Jayarathna, C. R. Aragon, and M. Mahmoud, "Biometric identification via an oculomotor plant mathematical model," in *Proc. ETRA 2010*, 2010.

[23] R. Bednarik, T. Kinnunen, A. Mihaila, and P. Fränti, "Eye-movements as a biometric," in *Image Analysis*, vol. 3540, Springer, 2005, pp. 780–789.

[24] L. Coventry, A. De Angeli, and G. Johnson, "Usability and biometric verification at the ATM interface," in *Proc. CHI 2003*, 2003.

[25] A. De Luca, M. Langheinrich, and H. Hussmann, "Towards understanding ATM security: a field study of real world ATM use," in *Proc. SOUPS 2010*, 2010, pp. 1–10.

[26] C. D. Holland and O. V. Komogortsev, "Biometric Verification via Complex Eye Movements: The Effects of Environment and Stimulus," in *Proc. BTAS 2012*, 2012, no. Btas.

[27] M. Kumar, T. Garfinkel, D. Boneh, and T. Winograd, "Reducing shoulder-surfing by using gaze-based password entry," in *Proc. SOUPS 2007*, 2007, p. 13.

[28] A. De Luca, R. Weiss, and H. Drewes, "Evaluation of eye-gaze interaction methods for security enhanced PIN-entry," in *Proc. OZCHI 2007*, 2007, no. November.

[29] J. S. Dumas and J. C. Redish, *A practical guide to usability testing*. Intellect Ltd, 1999.

[30] D. H. Koh, S. M. Gowda, and O. V. Komogortsev, "Input evaluation of an eye-gaze-guided interface: Kalman Filter vs. Velocity Threshold Eye Movement Identification," in *Proc. EICS 2009*, 2009, p. 197.

[31] C. Holland and O. V Komogortsev, "Biometric identification via eye movement scanpaths in reading," in *Proc. IJCB 2011*, 2011.

[32] S.-W. Shih, Y.-T. Wu, and J. Liu, "A calibration-free gaze tracking technique," in *Proc. ICPR 2000*, 2000, vol. 4.

[33] K. Z. K. Gajos, J. O. J. Wobbrock, and D. S. Weld, "Improving the performance of motor-impaired users with automatically-generated, ability-based interfaces," in *Proc. CHI 2008*, 2008.

[34] P. Kasprowski, O. V. Komogortsev, and A. Karpov, "First Eye Movement Verification and Identification Competition at BTAS 2012," in *Proc. BTAS 2012*, 2012.