UNIVERSITÄT
KOBLENZ · LANDAU
Faculty 4: Computer Science

WeST
People and Knowledge Networks
Institute for Web Science
and Technologies

Multi-model optimization of gaze and touch-based PIN Entry

Master's Thesis

in partial fulfillment of the requirements for
the degree of Master of Science (M.Sc.)
in Web Science

submitted by
Daniyal Akbari

First supervisor:	Prof. Dr. Steffen Staab Institute for Web Science and Technologies
Second supervisor:	Dr. Chandan Kumar Institute for Web Science and Technologies

Koblenz, July 2018

Statement

I hereby certify that this thesis has been composed by me and is based on my own work, that I did not use any further resources than specified – in particular no references unmentioned in the reference section – and that I did not submit this thesis to another examination before. The paper submission is identical to the submitted electronic version.

	Yes	No
I agree to have this thesis published in the library.	□	□
I agree to have this thesis published on the Web.	□	□
The thesis text is available under a Creative Commons License (CC BY-SA 4.0).	□	□
The source code is available under a GNU General Public License (GPLv3).	□	□
The collected data is available under a Creative Commons License (CC BY-SA 4.0).	□	□

....

(Place, Date) (Signature)

Zusammenfassung

Abstract

Knowledge-based authentication methods are vulnerable to Shoulder surfing phenomenon. The widespread usage of these methods and not addressing the limitations it has could result in the user's information to be compromised. User authentication method ought to be effortless to use and efficient, nevertheless secure. The problem that we face concerning the security of PIN (Personal Identification Number) or password entry is shoulder surfing, in which a direct or indirect malicious observer could identify the user sensitive information. To tackle this issue we present TouchGaze which combines gaze signals and touch capabilities, as an input method for entering user's credentials. Gaze signals will be primarily used to enhance targeting and touch for selecting. In this work, we have designed three different PIN entry method which they all have similar interfaces. For the evaluation, these methods were compared based on efficiency, accuracy, and usability. The results uncovered that despite the fact that gaze-based methods require extra time for the user to get familiar with yet it is considered more secure. In regards to efficiency, it has the similar error margin to the traditional PIN entry methods.

Acknowledgement

I take this opportunity to express my profound gratitude and deep regards to my thesis advisors Prof. Dr. Steffen Staab and Dr. Chandan Kumar for their exemplary guidance, monitoring and constant encouragement throughout the course of this project.The door to Dr. Kumar's office was always open to me whenever I ran into a problems or had a question about my research or writing. He consistently allowed this paper to be my own work but steered me in the right direction whenever he thought I needed it. Their input and patience have been invaluable in helping me to learn how to do research, and to navigate some of the of the more emotionally challenging aspects of this project. Their creativity and wisdom are inspirational.

I would also like to acknowledge Mr. Raphael Menges, who I am gratefully indebted to for his effort to help me during the implementation phase and for his very valuable comments on this thesis.

I would also like to thank the participants who were involved in the evaluation process for this project. Without their passionate participation and input, the experiment could not have been possible.

Finally, I must express my very profound gratitude to my mother and dedicate this work to the memory of my father Dr. Mohammad Akbari, for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Contents

1 Introduction

Human-Computer Interaction (HCI) is focused on the joint performance of tasks between humans and computers and how they communicate. There are two components in this exchange of information, Input and Output. Input components identify and sense the user's desired task and information to be communicated to the computer. There are two types of Input methods, Direct and Indirect. The Indirect input refers to the method where the actions of the user will translate into data or commands to be entered into a system. The Indirect input methods have been around since the beginning of Web development and they are still usable and prominent. Keyboard and Mouse are prime examples of indirect interaction methods. The direct input refers to the devices which have no intermediary and the movement of the user's body is equal to the input to the system. The Direct input methods have been introduced to facilitate users to have a natural feeling of interaction, for example, touch interaction.

A touch screen refers to an electronic visual display which is able to detect the presence and location of a touch on the screen. The touchscreen technology can be found in Smartphones, Laptops, Tablets and even ATMs. The touchscreen enables the direct interaction with what is being displayed and requires no intermediary device like a mouse. Touchscreen as a Direct input method enables users to simply point at their target and select it, unlike indirect methods which the user is required to press multiple keys or move the cursor to select an object. Direct and Indirect input methods have been evaluated and the experiment done by D. Schmidt showed that accuracy and time efficiency is much more optimal when using direct conditions [SBG09]. Touch screens are easier to learn and more efficient compared to indirect input methods. It is important to mention that touch screens are known to be practical for selecting large targets and they are not very accurate when used for small targets and they tend to be susceptible to error. During an interaction with a touchscreen interface the user firstly looks at the screen and targets the desired object to be selected and then proceed to actually selecting the object by touching the location of the object on the screen.

Eye tracking had been introduced as a way of direct interaction which could pave the way for new technologies and devices to be introduced for end-users. It is becoming a popular way of interaction. Gaze coordinates could be used to pinpoint the target selected by the user on the screen, then proceed to do a command more efficiently. We hypothesize that the limitations and disadvantages mentioned for touchscreen interfaces can be improved by adding gaze capabilities. It is worth mentioning that gaze is used anyway when people use the touch method and a finger or hand movement follows the user's gaze. Therefore the targeting can be done via gaze and the selection by touch, which will improve the efficiency and also the screen is not obscured by hand during targeting. The multimodal approach will benefit the users with a user-friendly interface and aims to make the PIN-entry secure and to improve the accuracy of target selection and reduce unwanted selections. The focus of this Master thesis is to analyze how the direct method of interaction could be optimized for PIN-entry, i.e, a multimodal approach of enhancing the popular Touch-based interaction using cues and context from gaze signals, so that it would increase accuracy, user-friendliness, efficiency, and security. It should be comparable to the current method of direct interaction and resolve its limitations. One can argue that this is not the most optimal and efficient way of interaction since it can be time-consuming for the user. The touchscreen interfaces are undesirable in some applications since the finger is considered to be a large pointer and cannot accurately point to small areas on the screen. It is also worth mentioning that the finger or hand can block the visibility of the screen during the interactions since most of the time hands have to move in front of the screen to reach the intended target.

User authentication refers to the process in which the identity of the user is proven in a system, to gain access permission. Authentication processes are part of the user's daily interactions, whether it is their email, online banking or online participation to vote. There are three main categories of user authentication. [Joa10]

- **Knowledge-based:** In this category, the user chooses and memorizes a specific password/PIN or combination of password/PIN with a username. The knowledge-based method is the most common method used on computers and smartphones. This method is vulnerable to brute force attacks and improper storage and selection of the password/PIN. This method is susceptible to shoulder surfing. Shoulder surfing is the act of acquiring the individual or private data through direct perception. Whether it is a person who will gather the user's private information by looking over the unaware user or in some case video cameras and vision-enhancing devices are used. Knowledge-based methods might lack the proper security since they are easy to share or even guessed.

- **Object-based:** Refers to a physical object that the user possesses and uses to gain access, like an ID card or token. These objects are usually designed in a way that can be easily carried by the user and they can store user's information on their chips. This method alone is not secure due to the fact that an unauthorized user could steal the object and gain full access. This method can be combined with the knowledge-based method to enhance the security, which is called a two-factor method. For example, the user could set a PIN for their email and also a one-time PIN which will be sent to their private smartphone and it is valid only for that session. The other strategy that almost all banks benefits from is Synchronized OTP (One Time Password) Generator which will generate a code which can be used for a specific transaction and it will expire after. In this case, the user will have to use knowledge-based password first and for a bank transaction, needs an extra code.

- **Inherence-based:** In this method, the user is authenticated based on the user's unique and individual traits. The physical traits such as the fingerprint or retinal scan could be used in this method to grant access to the user. It is worth mentioning that the security of these methods depends on the system since some systems can be easily fooled. In some cases, the user could choose which finger they desire to use for authentication, which will consequently improve the security since only the user could know which finger to use. This method could also be problematic for users who value their privacy and don't prefer to share their biometric information. The other risk that the user's of these methods will face, is that there are attackers who are willing to "steal" the body part or physically force the user to produce biometric authenticator. Which introduce a grave danger for the user's safety and wellbeing.

In this Master thesis, we focus on improving PIN entry and finding new solutions for issues like shoulder surfing and smudge attacks. As the most common authentication method, PIN entry has many security flaws. The user's password or PIN could be observed and memorized by a third party and subsequently, be used to gain access to the user's personal information. Furthermore, even when the third party can't identify what exactly are the entered numbers or characters, it is possible to learn how long the password is, or how many numbers it is consisted of. Therefore it becomes much easier to use brute force to compromise the information. It is crucial to keep in mind that the usability should not be sacrificed for security, which is the case for most of the solutions implemented. Currently, there are solutions that try to combine different entry methods or adopt security tokens, which ultimately increase the security, however, reduces the efficiency and ease of use. We present TouchGaze as a contemporary way of PIN entry which is built on the shoulder of traditional methods. Touchgaze utilizes gaze and touch to eliminate the dangers of the shoulder surfing. Eye tracking data shows the user's gaze and the location of it on the screen. The user interface has a direct influence on accuracy, efficiency, usability, and security. We are going to discuss in details the different designs and choices which were made to implement TouchGaze. In this work, three different authentication methods were implemented and an evaluation has been done to compare them. The results of the experiment are the main contribution along with the proposed and implemented "TouchGaze" method.

2 Background and Related work

2.1 Technical information

The technologies that we have used in this work are explained.

- **Python:** The Python programming language was created by Guido van Rossum, and unlike many popular programming languages like C++ and Java, Python provides a simple but powerful syntax. Python is a high-level programming language, which means that it has to be processed to machine language before it can be run. Which takes time for processing, but on the other hand it is much faster to program in a high-level programming language like Python.

- **PyGaze:** It is an open source software package for Python and offers a platform to create and run complicated trials with an eye tracker in a user-friendly environment. [DMS] Pygaze can be used across most popular operating systems such as Windows, Linux, and Mac OSX. It is also compatible with many different eye tracker devices like SMI device. Pygaze is dependent on two other libraries PyGame and PyschoPy. They are both used to control the mouse, keyboard, and displays. PyscoPy is suitable for complicated stimuli. Pygaze is well documented and there are many experiments done in it. Therefore it makes it accessible to users. In Pygaze it is possible to use the mouse instead of an eye tracker to test the implementation when the eye tracker is not available.

- **IviewRED:** The iViewRED is a dark pupil eye tracking system that uses infrared and computer-based image processing to detect the pupil in real time and translate pupil location as gaze data. [SMI09] This software provides us with the possibility of integration with Python and to receive real-time data from the eye tracker device. The software comes preloaded with high functions like calibration and tracking visualization, which makes it much more convenient to interact with the eye tracker.

2.2 Eye Tracking and Touch

Eye tracking is a promising input medium that facilitates communication between users and computers via the eye movements. Although the measurement of the eye movement has been improved through the years it is still far from perfection. The first ideas to incorporate eye tracking as an input medium belongs to Jacob. he explained his ideas in the paper "What You Look At is What You Get".[Jac91]

Considering the developments and innovations in eye tracking technology, which resulted in increasing precision and reducing cost, but there are still some challenges to use the eye as an input medium. [KMS16] When we consider eye as an input method, there are challenges like calibration errors that we have to face and address. The accuracy problem could be resolved by adapting the layouts and enlarging the targets to reduce the erroneous interactions. The interface should provide appropriate feedback to the user to avoid errors and aid them in the interactions. We can notice these mentioned solutions in GazetheWeb browser which supports gaze-based web browsing and has functionalities like Scrolling, Navigation, Tab Management and Text Input.[KM^{+}17]. The accuracy and usability of gaze-based interaction could be elevated by adding Touch interactions to them. There are some works which are based on the idea of combing Touch and Gaze together.

Pfeuffer and Gellersen [PAG16] have combined gaze and touch in tablets, which enable the user to use the device by touching the device with thumbs and then redirecting the target with the gaze. Therefore it helps the user to hold the device with two hands and use only two fingers for touch and benefit the gaze capabilities. This would help users to do the navigation and browsing task conveniently and accurately. For example Figure 2 shows how the user is browsing a map by targeting the desired position with gaze and then zoom in with the use of touch. Based on the study they have done, most of the users preferred gaze-touch combination compared to only gaze interaction. To address the possible inaccuracy using gaze for targeting specifically for small targets, the idea of CursorShift has been introduced, which is a method that uses gaze to provide users temporal control over cursors during direct-touch interactions. This will enable the user to temporally activate a cursor at the user's gaze position when issuing a tap from the grip position. Then deactivates it after another tap which is also used to perform a 'click' on the cursor's target. Users can comfortably scroll a web page using the thumb with the hand that holds the device. Then, if the user wants to select a hyperlink, the user can utilize the same finger to instance a cursor, drag it precisely, and click the desired link [PAG16].

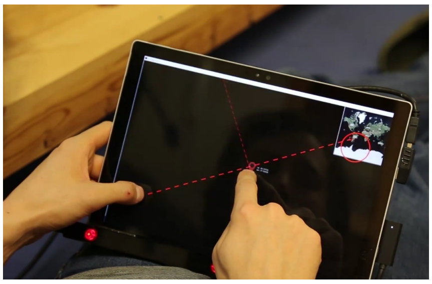


Figure 1: Combination of touch and gaze interaction, browsing a map.

The other related work is combining direct and indirect touch input which would be consist of two screens. Horizontal touch surface as a direct input device and a Vertical Screen for indirect interaction. By introducing gaze as an additional modality we are able to reduce the time that is needed to reach targets on the vertical screen as well as reduce the effort that is needed to interact with the system. This enables users to comfortably interact with interactive workspaces for a longer time [VMSB15]). In Figure 3 the user is interacting with the horizontal screen through touch input and by gaze input with the vertical screen. There are two proposed interaction techniques Indirect Touch Surface Selection (ITSS) and Indirect Touch Object Selection (ITOS). In ITSS gaze is directed towards the horizontal touch surface, the system maps the touch point to the horizontal screen, and in ITOS user's gaze transfer the initial touch point to the position of the surface where the user is currently looking at. According to the conducted experiments, they come to the conclusion that the performance of these interactions techniques is dependent on the type of task, for example, ITOS outperforms the direct touch and ITSS in terms of tapping speed but when the task is the dragging an object on a screen, direct touch outperformed ITOS and ITSS in speed on the vertical screen but with less accuracy.

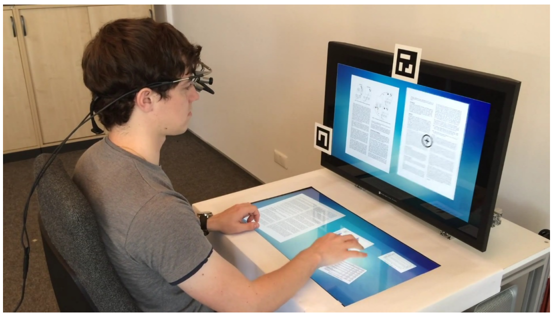
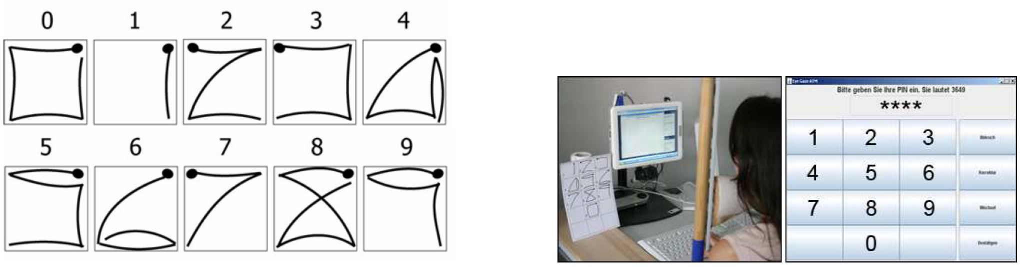


Figure 2: Combining direct and indirect touch input using touch on the horizontal screen and gaze for vertical screen

2.3 Personal Identification Number (PIN) Entry

Kumar has designed "EyePassword" [Kum07], a gaze-based password entry method. They have designed it in two different trigger methods, dwell-based and trigger-based. Trigger-based is done by looking at the keys and selecting them by a trigger key like space bar. The evaluation done on this system showed that Gaze + Trigger is more susceptible to error, but Gaze + dwell performed similar to the keyboard method, in regards to the number of errors. This could be due to the eye and hand coordination. However, the result showed that the keyboard method is more time efficient compared to gaze-based ones but, the increased security is vital and we can overlook the elevated average time for entry.

Gaze gestures could be used as an input method for PIN entry. The work done by A.De Luca proposed to assign specific gestures to each number. [LW07] The concept is that for each number a specific button must be selected and then the eye movement is translated into a gesture. The table of gestures assigned to each number can be seen in Figure 3a and the user study set up in Figure 3b. They have implemented three different prototypes. Dwell-based, Look & Shoot and the gesture-based. The keypad design is based on the German ATM keypad. Based on their evaluation done on these three prototypes, they realized that the gesture-based method took on average 54 seconds per PIN entry. The average time is quite enormous when compared to the traditional ways. On the other hand, the results showed that the gaze-based method performed much better in regards to accuracy and only 9.5 percent of the PINs entered were erroneous. Although the accuracy seems promising for this method, the average time is quite high and reduces the usability of the method. Not to mention that users need to know and understand each gesture or have the gesture table in front of them during PIN entry, which could be a hassle for them.



(a) Gesture table

(b) User study setting

Figure 3: Interfaces

The rotary design [BD12] created by D.Best and A.Duchowski resembles the classic interface of old phones. The numeric keypad design uses gaze dwell time to identify each key entry. To avoid any accidental entry, the user must fixate on "" sign to initiate an entry and then fixate on each number of the 4-digit PIN. In the end, to enter the PIN the user fixates on "#" sign. Thus, for entering a 4-digit PIN, 6 fixations are required. However, in the Rotary design instead of dwelling on each key, the user transitions gaze from the middle of the design to each number and then again in the middle. Their experiment with the two interfaces showed that rotary design performed slightly better than keypad design in regards to accuracy.

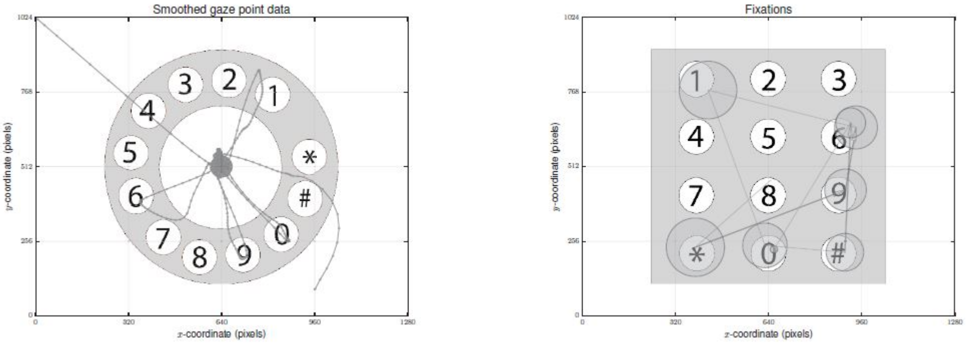


Figure 4: Rotary and keypad interfaces [BD12]

There are some approaches which focus on increasing the noise, therefore the observer could not identify the actual PIN. However, this method increases the noise for the user and decreases the usability and accuracy of the method. It is possible to add graphical identifier on the screen that only the actual user know which one to select. A good example of this approach is "ColorPIN" [LHH10]. In their design, the normal keypad layout is presented with a difference of existence of three different characters which are assigned to each number. The numbers are in different colors of black, white and red. Each time the user enters a key, the order of the letters will be randomized. According to their evaluation, the time for entering a PIN increased considerably compared to the standard PIN entry method. Security increased in "ColorPIN" and only in 2 out of 48 cases the PIN could be identified by a third party. Although the effort and timing to complete each PIN entry is very high.

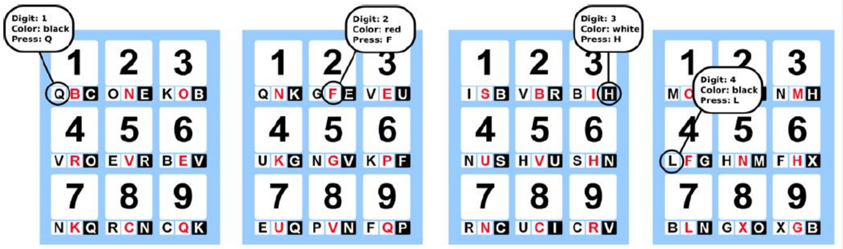


Figure 5: ColorPIN method for entering 4-digit PIN [LHH10]

There are other works related to graphical passwords."PassFace" is a gaze-based graphical password method, which has been designed for ATMs. [DFO08] In the experiment done by them, the user first would go through calibration and then will be shown 5 different faces, which the user has to memorize. Then there is a game like page to familiarize the user with the selection process and to make sure the calibration is done right. Then the user will be shown a grid of 3 x 3 of different faces. The user must choose the familiar face which was shown in the beginning to gain access and successfully be authenticated. The method poses a challenge for the user which they have to remember a face and later will be asked to choose that between 9 different faces. Their study showed that users improved their accuracy in the second session of their experiment and could enter the password successfully more accurately. However, the interface does not possess the usability and familiarity of the traditional keypads which users are used to.

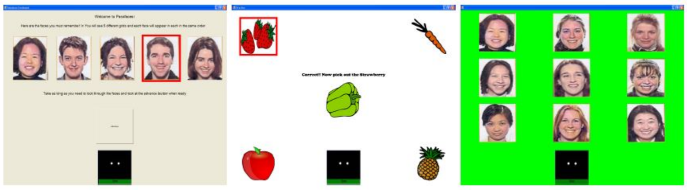
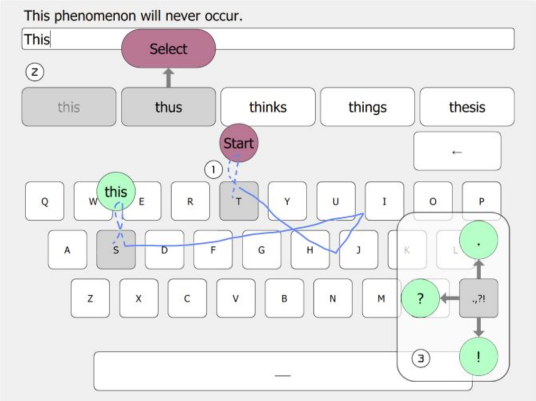


Figure 6: Graphical password method "Passface" [DFO08]

In the related works, we have to mention "EyeSwipe" [Kur16] which are designed for the purpose of dwell free text entry. The design is similar to the touch-based swipe to enter the desired text, but with the use of gaze-path instead of touch. The user starts to indicate the first and last characters by gaze and then move around the vicinity of the other keys. Then the system suggests the most related words based on the gaze-path and the user could select the desired word among them. The spaces between each word are entered automatically to speed up the process. The experiment was done between the dwell-based interface and "EyeSwipe" showed that the typing speed of 9.5 wpm and 11.7 wpm respectively. Based on the opinions of the participants, the user preferred EyeSwipe over dwell-based. The recorded speed is quite exemplary and shows the efficiency of this method.



This phenomenon will never occur.

This

Figure 7: Entry of the word "This" the first and last characters are selected by reverse crossing and the vicinity of the middle characters are looked at[Kur16]

There is another multi-model authentication method, called GazePassTouch. [KAH16] This method is used for authentication on smartphones. The concept of the GazePassTouch is to combine Touch and Gaze for PIN entry. Their assumption is that an attacker only could focus on the hand touching the screen or the movement of the eyes. The method combines the touch-based options of 0 to 9 and gaze-base options of left and right. In GazePassTouch, the user could either touch the first number or look at the right or left side of the screen. An example of an entered PIN would be "1-left-2-right” (3-switches)". The hypothesis is that if the user switches more often between touching the number and gazing right or left, their PIN is more secure. Theoretically, their password could have space of 12^n (10 numbers and 2 gaze options) and "n" here refers to the length of the password itself. [KHBA17] There is another work that is built on the base of GazePassTouch, which is called GTmoPass. GTmoPass is designed for Public displays like ATMs. The user has the same design of GazePassTouch on their smartphone and does the same interactions to be authenticated. The difference is that the smartphone is used as a token to communicate the authentication securely to the public display. Only then the user can gain access to the information requested on the public screen. Based on their evaluations, the accuracy of the GTmoPass decreased compared to GazePassTouch and the average time of GazePassTouch was lower. GTmoPass might not be as practical as it should be for public displays. In our opinion, the GazePassTouch implemented on an ATM could be more efficient. GTmoPass also raise the issue, which users must have their smartphones in hand at all time to gain access to their information in public screens.

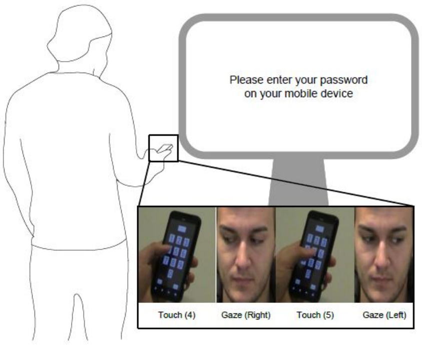


Figure 8: GTmoPass: Two-factor Authentication on Public Displays Using Gaze-Touch passwords and Personal Mobile Devices[KHBA17]

3 Research Problem

As technology advances and becomes ubiquitous, we can see emerge of new devices and terminals in public, Whether it is an Automated Teller Machines (ATM), train ticket terminal or at a fast food chain restaurant. The factor that they have in common is that they are located in public and can be used by anyone. This fact is considered both an advantage and disadvantage of these services. [Luc08] These public screens are observable by direct and indirect third parties, therefore the privacy and security of the users are in jeopardy. But this dilemma is not limited only to the public displays and extends to the private displays like smartphones, which are being used in a public place.

Nowadays almost all devices are password protected and even some users prefer to increase their security by using two-factor authentications. With new authentication methods, we have newer and improved cameras which can provide better resolutions. Therefore, these already built-in devices could be used for the purpose of eye tracking by just adding inexpensive image processing software. Consequently, we can consider eye tracking as a prosperous tool to implement gaze-based password entry methods. These authentication methods could be ideal for the ATM, due to the sensitive nature of the authentication and the fact that all ATM devices are in a public area, the user is risking the exposure of his/her PIN to a third party. The silver lining here is that since the user will stand in front of the ATM it is much easier to use gaze for PIN entry and also the Keypad layout is less likely to be susceptible to errors compared to the QWERTY keyboard. The layout of the PIN entry method contains fewer keys than a QWERTY keyboard which means that button sizes could be larger that results in better accuracy.

The shoulder surfing problem should not be underestimated, whether when it is as apparent as just looking at the screen of someone's smartphone or using hidden cameras to identify the PIN at an ATM. The issue might seem uncomplicated but many have been the victim of it.

The devices which are equipped with touch screen, usually show the fingerprint traces on them in the area which user has touched. These smudges and residues could reveal the points which are touched often and could leak the numbers entered on the keypad. The exploration of these patterns on the screen to identify and extract sensitive information is called "Smudge Attack". [AGM$^+$10] Smudge attack is considered as a threat, since users don't care much for it, and they don't take it into consideration. Even if they do and clean the screen from smudges regularly, it is still possible to detect the smudge point with the use of cameras and computers.

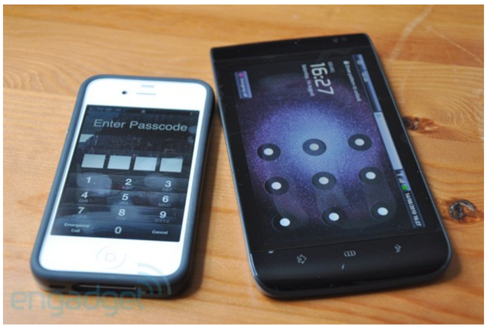


Figure 9: Smudge traces on smartphones[roo16]

Smudge attack is not limited to touchscreen devices and can be related to ATM devices and devices with physical keys. The attacker could come to the position of the device to observe the screen from distance with the use of cameras. For example, in some ATM devices or point of sale (pos) devices, it is possible for the attacker to observe the heat signatures on the keypad to identify the numbers, as well as knowing the order of them. Since the first pressed key tends to show lower heat signatures compared to the last key pressed.

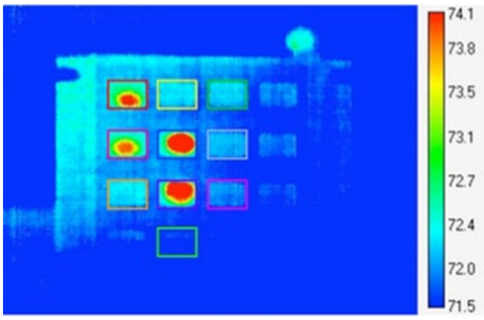


Figure 10: Thermal imaging of ATM keypad [Clu15]

In authentication methods, if we can limit the hands to only carry out trigger rather than targeting and triggering, then the shoulder surfing could be eliminated. The main idea should be to improve the security of authentication methods without increasing the complexity of the user authentication process. The new method should be built on the simple everyday methods which users are familiar with and proved their acceptance among all types of users. In case that the new method presents a completely different and innovative method, it is likely to face the danger of increasing the complexity and could not be applicable to the real-world problem

In this master thesis, we focus on improving the security of the PIN entry method by combining touch and gaze, while touch is used for selection and gaze for targeting. In the meantime, we make an effort to keep the efficiency of the new method desirable, so that it could match the traditional user authentication methods.

4 Methodology

In this chapter, we will explain the set of challenges we face in this work and the approaches we have taken to overcome these obstacles. We will go through the technical aspects of the implementation and the design decisions that we made. Since in this work, the purpose is to combine touch and gaze for the authentication method, it was important to notice the way touch and gaze should work together and not have any interference with each other.

4.1 Challenges

There were several challenges which we had to face to reach the optimal PIN authentication method. There are many factors involved in designing a promising authentication method and if all of the requirements are met then we can consider the method as an alternative for the traditional methods. The final method proposed should be adaptable and could be applied to different platforms. The method needs to have the capability to be used on private and public displays and not be dependent on a specific object that the user possesses. We will go through the factors considered one by one.

4.1.1 Security

The authentication method should reveal the least amount of information about the PIN. Our design should provide users with a higher level of security which is not attainable through other methods. We have discussed shoulder surfing, smudge attack and thermal attack in the last chapter. For the design of the PIN entry method, we considered each of these threats and tried to rectify them in the traditional methods. There were two obstacles to deal with, firstly we needed to eliminate the use of hands for targeting of each key. Secondly, We had to eliminate the number of triggering occurred, to hide the number of digits in the user's PIN. In this manner, the observer could not detect which keys are being entered by the user and additionally has no information regarding the length of the PIN entered. Furthermore, the current position of the gaze should not be displayed on the screen, since it will reveal the PIN. However, it will not be possible for the user to be sure where he/she is looking at, which might be inconvenient.

4.1.2 Usability

The usability of the method should be compared to the traditional way of authentication. It is possible to boost the security by coming up with complicated methods, that each time a user wants to be authenticated they need to put in considerable effort. Consequently, the method should find the balance between usability and security. The design had to be familiar to the user to increase the chance of user acceptance. It is probable that modest changes need to be made to the classic keypad design, although it should not change the integrity of it. Even the slight changes should be considered and tested to see how they affect the results.

4.1.3 Speed

The speed does matter in every aspect of our lives and authentication is not excluded from it. Therefore we need to limit the amount of time needed for the authentication to the minimum. The interaction timing should be acknowledged and it should be compared to the conventional authentication methods. It is only logical that when we increase the security and introducing a new method to the user, the time for each entry would increase. So the objective was to keep the range as close as possible to the common methods like touch PIN entry method.

4.1.4 Error Rate

The other important aspect of the authentication is accuracy since we don't want users to have to repeat PIN entry many times to be successful. Accuracy is crucial in actuality, since in most cases if the PIN entered not correctly for several times the user's account will be blocked. And this problem should be avoided by reducing the inaccuracy of our method. One factor in accuracy is muscle memory, which users are familiar with the interface and have entered their same PIN many times before. But there are other factors involved in the accuracy of our method. The accuracy of the eye tracking device is crucial for our work and the eye drift must be corrected before each PIN entry. Since the eye position is not shown on the screen due to the security reasons, the user is not made sure that the specific key is selected or not. It is possible to implement a feedback method, which will inform the users that a key is entered, but again that would compromise the information about the size of the PIN. We should differentiate between when the user is looking only with eyes and when he/she wants to target the keys with the eye. It is also worth mentioning the position of the touch interaction should not be confused with the gaze position.

4.1.5 Interface

The interface should be designed with all the mentioned attributes in mind. There are many different designs and interactions methods which have been done in recent years. So it was crucial to consider their ideas, and then we could find a better way and not make the same mistakes again. In the interface, we focused on the layout of our design, in regards to how it will look like and how we were going to arrange the keypad. The size of the keys plays an important role as well. For instance, we faced an issue with repeating the same number in one of our designs and we needed to change the design moderately to overcome the matter. As we mentioned in the related work chapter, although the rotary design outperformed the universal keypad design it is worth to acknowledge that most of the youths have never seen a classic rotary phone before. This makes it challenging for them to make a connection to the rotary interface. It is preferable to have an interface which is acceptable for the majority of users.

4.1.6 Interaction

The PIN entry method must contain all the necessary requirements, moreover, it should feel natural for the user and be easy to learn. This means that the user should not face challenging interaction which asks them to put much more effort than they usually do, to be authenticated. This will results in the end user being annoyed with the method and ultimately decide to revert back to traditional ways which they felt complacent with. Eyetracking is not as acknowledged as Touch among people, so the interaction should capture the essence of the gaze and feels natural for them.

4.2 Approach

To address the aforementioned challenges, we have come up with a solution to incorporate touch and gaze. The main idea is to use touch as a trigger and gaze as a targeting mechanism.

In the first step, the keypad interface is chosen and the distance between keys are decided since it will play a role in the accuracy of the method. The method performs drift correction to check if the calibration is still effective for the user. The eye tracker device is connected and recording at all times, but the data is utilized for targeting only when a touch is initiated. The touch is recorded as a boolean object and the position of the touch is not recorded. Therefore it is present either as True or False value. That would activate and deactivate the algorithm to process the gaze position for selection. After the touch is activated, the gaze data will be recorded for selection and the algorithm will search to find the possible key/keys that the user wants to select. Even if the eye position is in close vicinity of a button, that button is selected and entered. If the user's gaze position is outside of all the keys and the touch is activated, the system will give a feedback that the button was not selected. Therefore the user is informed and can try again to enter the same number. The user can submit the PIN by activating the enter function and then the system could compare the entered PIN with the actual PIN. Based on the correctness of the PIN entered the user will get a success or denied page.

4.2.1 Touch PIN entry concept

In the touch method, the procedure is fairly straightforward and the user can trigger each key by simply touching on the key located on the screen and in this method the gaze data is not utilized. In another word, the targeting and selection of each key are done by touch. Like any other simple PIN entry method. After each touch, the position of the touch located on the screen is checked, to determine if it has triggered any keys and if so the number is entered to the PIN value. If the position is outside of the key border nothing will be added to the PIN value. The idea for this design is inspired by the traditional keypads which can be found on ATMs and smartphones. The concept is quite familiar to users and they have adopted this method for many applications.

4.2.2 Touch + Gaze PIN entry concept

In Touch + Gaze method the procedure is different from Touch method and the gaze data play a role in the interaction. The user can target the desired key by gaze and trigger the key by touching any part of the screen. The user's eye position will determine which key is targeted when a touch is initiated. Even if the user touches on a specific key located on the screen, only the location which the user is fixating on with eyes will be selected and the location of the touch on the screen is not considered. To simplify, the targeting is done by gaze and the selection by touch. If we have a 4-digit PIN number, the user will look at the first key and touches to select the digit, and then repeat the process 4 times. so for each entry, the user touches the screen. The idea is that in Touch method the user will look at the desired key and then move the hand to select that specific key, but in Touch + Gaze there is no need for hand movement but only for touching the trigger key, the eye will take on the movement role. The Touch + Gaze concept is based on the work done by Kumar, which entry is done simply looking and pressing the trigger key. [Kum07]

4.2.3 TouchGaze PIN entry concept

The TouchGaze method is similar to the Touch + Gaze method in regard to targeting by gaze and triggering by touch. However, the eye movement is recorded as a gaze path and it will be used to determine the keys selected. The user will start the PIN entry by touching any part of the screen and keeping the touch down until the process is completed. The user will initiate the touch and then can swipe with eyes from the first key until the last key in order and release the touch to confirm the entry. In this method, the need to select each key one by one is eliminated and the user can enter their PIN by use of their gaze path. Unlike the other two designs, TouchGaze does not divulge the number of digits of the PIN entered. The method can identify between key selection and the eye simply moving to get to another key. So let's say the user wants to enter number 1 and 3, in that case, the eyes will pass through 1, 2 and 3. Although the user's gaze is activating only 1 and 3 keys and 2 is not entered. The idea for TouchGaze is from the trace-based touch keyboard. [ZK04] which the user could touch the letters by traveling their touch between these characters on the keyboard. As well as the works done with eye traces, in which the user can enter a word on a virtual keyboard by movement of the eyes on the characters. [HDG13]

4.3 Implementation

We have implemented our interfaces on Windows operating system, using an SMI REDn eye-tracking device. The resolution of the laptop used is 1366*768 pixels. During the interaction, we do not show the current position which the user is looking at since it should be hidden from the screen for security reasons. The interface in Touch and Touch + Gaze in Figure 11 conform to the traditional PIN layout although the interface in TouchGaze Figure 13 has an additional button R. The R button stands for repeat and it is located on the right side of the 0 button. We decided to keep the traditional order of the numbers in our layout so that it corresponds to the visual memory of users. It was possible to randomize the position of each key every time a user wants to enter a PIN number to increase the security even more, but that would diminish the usability of the method. Accordingly, we have decided to keep the traditional layout in all of our designs.

In the implementation phase, we have developed a function that would give feedback to the users after each entry which indicates that a key was entered or not, but for the evaluation phase, this feature was removed.

4.3.1 Touch PIN entry method

The method will have the traditional interface which is presented in Figure 11. There was a delete and enter button in place which we removed for the evaluation purposes but still kept the buttons on the keypad with their functionality deactivated. The user could enter each key by touching the box for each number, to enter the said number. When a touch is set to True the position of the touch is recorded as a tuple of (x,y). The buttons are logically bigger than their appeared boxes, so the algorithm will collect the (x,y) position and calculates if it is located in the vicinity of a specific number. The logical form of the keys is presented in the Figure 12. We have decided on larger logical sizes for the keys compared to what appears on the screen to reduce the possibility of gaze errors, but we have the same layout for the Touch method. Therefore if the touch is outside of the white box but still in the green areas, the closer key to the position is selected.

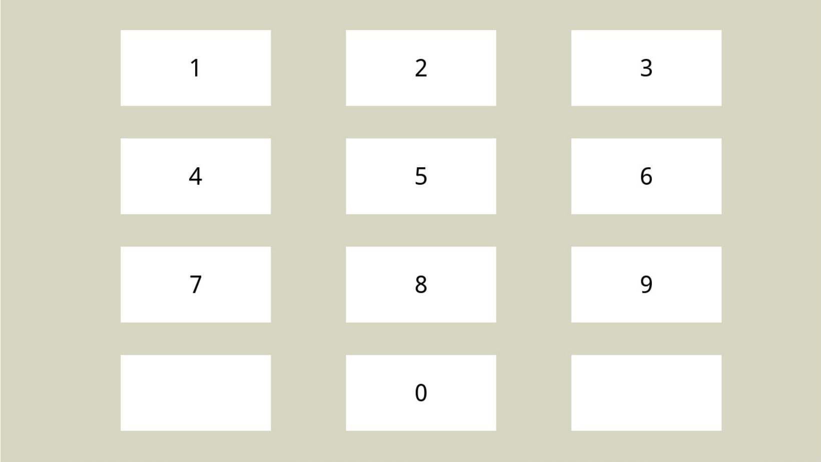


Figure 11: The interface of Touch and Touch + Gaze

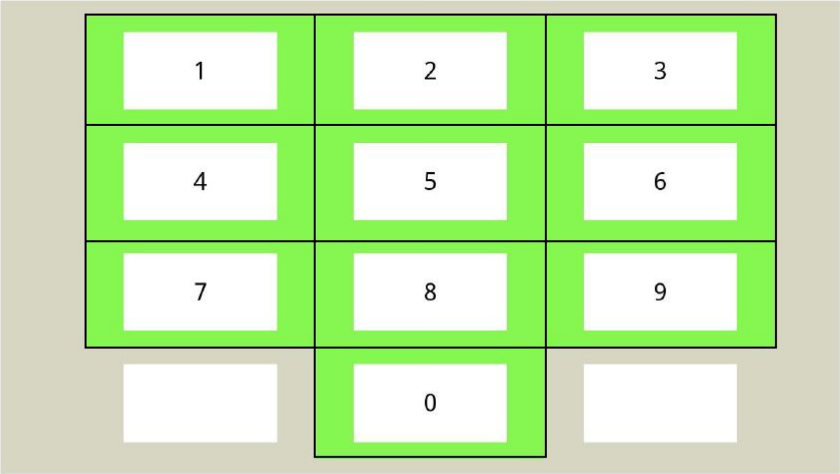


Figure 12: The logical interface of Touch, Touch + Gaze and Touchgaz

The design could include a feedback mechanism so that the users are informed whether a key is selected or not. The background could change color to green if a key selected and change to red when the (x,y) position is outside of all key positions.

4.3.2 Touch + Gaze PIN entry method

The keypad in Touch + Gaze is exactly the same as the Touch method. In this method, the gaze data is being collected by function "EyeTracker.sample()" at all time during the interaction. The eye position is being streamed and when the user touches the screen to trigger a key, the last fixation point of the eye is captured by the algorithm to determine which key has been hit and enter the number into the PIN value. If the eye position is outside of all key positions and a selection is initiated, nothing will be entered to the PIN value. After each session of PIN entry, the system would call "eyetracker.drif-correction()" function to make sure the accuracy of the gaze position and to ensure that calibration is still intact. After a certain number of digits have been entered, in our case 4, the system would submit the PIN and compare it to the actual PIN. The PINs are stored with the timing of the entry. For example, let's say the actual PIN number requested is 3555. The user could look at number 3 and tap on the screen to select it and then does the same process for other numbers. To repeat the same button the user could simply look at the same position, in this case, 5 and tap 3 times to enter triple 5s. (gaze@3, tap, gaze@5, tap, tap, tap).

4.3.3 TouchGaze PIN entry method

The gaze data collection is activated when the user touches the screen. The gaze path and the fixation points are collected until the user releases the touch. For identifying the fixation points the current position of the eye is collected and each 200ms the new gaze position is compared to the average of the last data collected until the distance between the current point and the centroid becomes larger than 100 px. Then the time for that fixation and the centroid (x,y) is stored and the timer is set to zero for the next fixation. Pygaze provides fixation functions which can provide us with all the required fixation points, but the algorithm we have developed was better suited for our development Based on the data collected the algorithm calculate the fixation points which have dwell-time of larger than 300 ms. These fixation points then processed to identify the keys entered in their respective order. In Touch and Touch + Gaze for repeated numbers, the user could simply select the same number multiple times, however, in the TouchGaze design, we came up with a new button which repeats the last number entered. This key store the last entered key and each time a new number has entered the value of the R button will be updated to the last key. For example, if a user has a PIN 3555, he/she could start the PIN entry and then focuses on 3, then 5, then R and at the end 5 again. The PIN stored will be "35r55", and we will remove all R characters from our the PIN at the end.("Touch&press", gaze-> 3, 5, R, 5, "Release")

In the initial stages of the development, we used the dummy mode which simulates an eye tracking device by using a mouse, so it wouldn't be necessary to test each component with the eye tracking device.

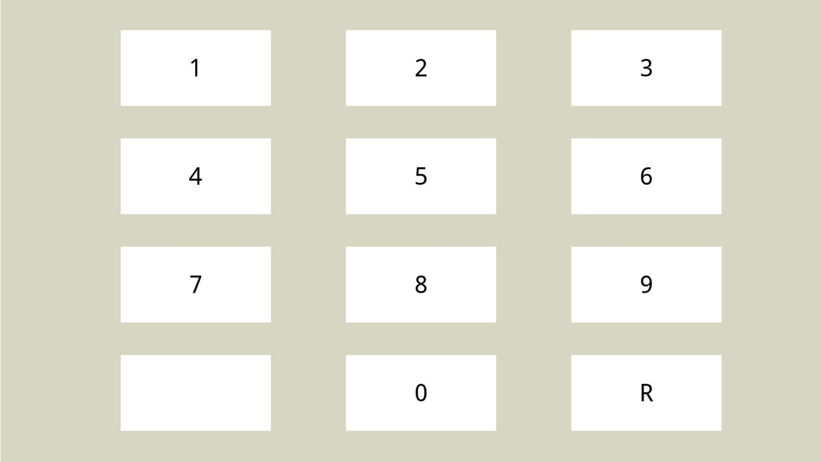


Figure 13: The interface of TouchGaze

5 Evaluation

We have conducted a pilot study before the actual evaluation. There were only two participants in the pilot study and the first participant was given the Touch and Touch+Gaze with an enter key to be selected at the end of each PIN entry. But the second participants were given the interface without the enter key. The pilot study results showed that the enter key increases the time for each PIN entry, therefore we decided to eliminate the Enter key.

5.1 Experimental Set up

The evaluation took place from 26.06.2018 to 29.06.2018 in the research laboratory at the University of Koblenz-Landau, Campus Koblenz. The duration of each experiment was between 35 to 45 minutes. At the beginning of the experiment, the participants were asked to sign an "Informed Consent Form", which was prepared before the experiment. We can see the room setting in Figure 14. The participants were seated in front of the touchscreen laptop which was connected to our eye-tracking device. Base on the instructions of the RED-n Eye Tracking device, the optimal distance of the user and device should be roughly between 50 to 75 cm, which was considered in our set up.

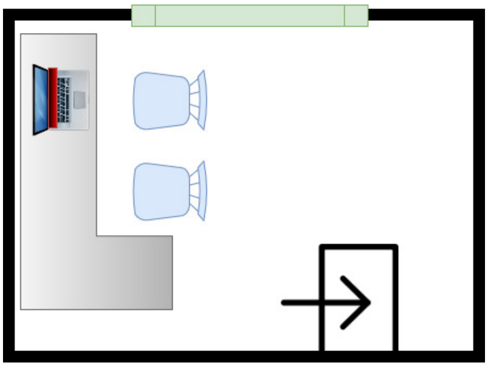


Figure 14: The experiment set up

5.2 Participants

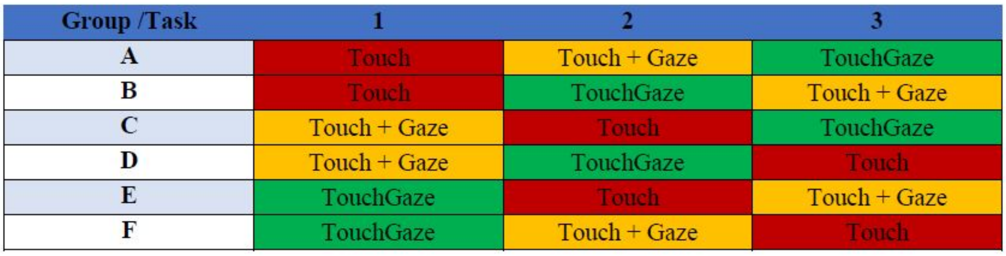
For the evaluation, we have gathered 18 test subjects, between the ages of 22 to 33 years old (Male = 8, Female = 10). All of our participants were students from the University of Koblenz-Landau. The student came from different study backgrounds, such as Computer science, Physics, and Mathematics. Out of our 18 participants, 6 wore glasses and 2 wore contact lenses during the experiment and the remaining had no visual aid. Among all participants, 12 had never used an eye-tracking device before.

5.3 Apparatus

The keypad interfaces were displayed on an 11-inch Lenovo touchscreen monitor with the screen resolution of 1366*768 pixels. The eye tracking device used for the experiment was SMI REDn scientific device which operates at a 60 Hz sampling rate.

5.4 Counter Balancing

The Latin Square is an experimental design which is used to control the variation in an experiment. The order in which treatments are given can actually affect the behavior of the subjects or elicit a false response, due to fatigue or outside factors changing the behavior of the subjects. To counteract this, it was decided to use a counterbalanced design, which reduces the chances of the order of treatment or other factors adversely influencing the results. This design was also used to control the variation in the experiment. To remove the effects of confounding variables on our experiment, the 18 participants were divided into 6 groups. Each group has completed the three phases in a different order, as we can see in Figure 15.



Group /Task	1	2	3
A	Touch	Touch + Gaze	TouchGaze
B	Touch	TouchGaze	Touch + Gaze
C	Touch + Gaze	Touch	TouchGaze
D	Touch + Gaze	TouchGaze	Touch
E	TouchGaze	Touch	Touch + Gaze
F	TouchGaze	Touch + Gaze	Touch

Figure 15: Divided groups

5.5 Procedure

In this study, we focus on the PIN entry accuracy and speed for our three different PIN entry methods. Each participant had completed our three different methods of Touch, Touch + Gaze and TouchGaze based on the order that has been mentioned in their assigned group table. Each participant was given a brief explanation about the experiment upon arrival and a questionnaire regarding demographical information to be filled out. Then the users were seated on a fixed chair so that the calibration could be done more accurately. The test subject was asked to avoid talking and to limit their movements, to keep the calibration accurate. For the calibration, the users had to follow the red dot on the screen which moves in a different part of the screen. If the calibration was not satisfactory, it had to be redone.

Step	Task
1	Signing of the consent form
2	Test subject fills out the pre-test questionnaire
3	Explaining the tasks and what is required from the test subjects
4	Completing the Calibration process if necessary
5	Explaining the first method
6	Start with the first method
7	Completion of the method and checking the state of recorded data
8	Completing the Calibration process if necessary
9	Explaining the second method
10	Start with the Second method
11	Completion of the method and checking the state of recorded data
12	Completing the Calibration process if necessary
13	Explaining the third method
14	Start with the third method
15	Completion of the method and checking the state of recorded data
16	Test subject fills out the post-test questionnaire

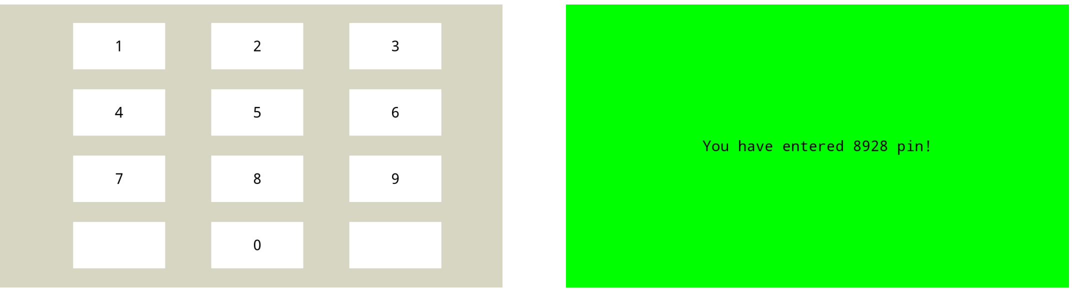
Table 1: Experiment steps

In Table 1, we can see all the steps are taken and tasks which were asked of the test subjects.

5.5.1 Quantitative Evaluation

In the controlled environment, the time and the number of errors are recorded for the user who is entering the PINs. In each experiment, the user will be given a randomly generated 4 letter PIN and right after, they will be shown the keypad layout, so that they can enter the PIN. The similar experiments are done in the same manner and the user has to memorize the PIN then move on to enter them in the PIN interface. All of the subjects had the chance to train and practice with the PIN entry process and get familiar with the layout and the tasks that were expected of them. In our experiment, it was crucial to pay extra attention to the calibration of the eye tracker and observe the participant, so that they wouldn't change their position. Since this could have interfered with the result of the evaluation. The calibration was not applicable to the Touch method therefore, the test subjects had more freedom for movement during Touch evaluation.

- **Touch:** The user will be given a random generated 4-digit PIN and the user needs to remember the PIN and moves on to the next page. In the next page, the keypad is visible and the timer starts then. In this layout, the user only selects each number by touching them on the screen and when the 4-digits are entered they will be directed to the next page. In this page, the user can see either a green success screen or a red error screen based on the correctness of the PIN entered. All the required data is collected and the user can go for the next PIN entry.



You have entered 8928 pin!

(a) Touch interface

(b) Successful PIN entry

Figure 16: Interfaces

- **Touch + Gaze:** Based on the experiment done by M. Kumar, they found out that gaze + trigger is more susceptible to error compare to gaze + dwell.[Kum07] And their experiment was Password entry with a QWERTY keyboard and also with the alphabetical keyboard. We did a similar experiment to them but with a difference of using a Keypad interface.
  The calibration of the eye tracking device is done before this phase and when the calibration is satisfactory the users are asked not to move their head or their chair, Which could interfere with the calibration's settings. After calibration, participants could move on to the PIN entry. Before they can see the layout, a 4-digit number appears on the screen which is a randomly generated number. The user memorizes the PIN and when ready, can go to the next page to enter the PIN. The timer for the entry is started and continues to run until a 4-digit number is entered. In this layout, the user targets the desired number with gaze and then selects them by touching any part of the screen. After they have entered all four numbers, the system will submit the PIN. The screen shows a success green screen for correct entry of PIN and a red error screen fro the wrong PIN entry. All of the data is recorded for that entry and the user can move on to the next entry.

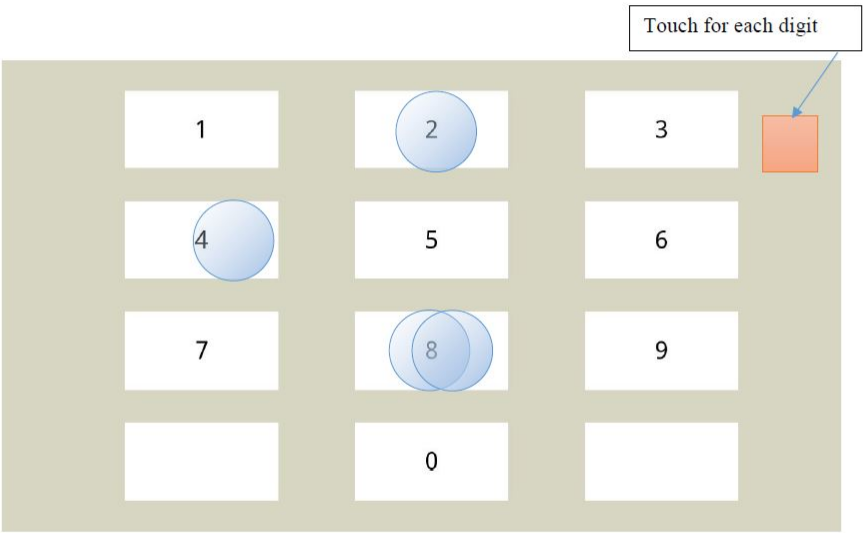


Figure 17: Touch + Gaze interface: The user looks at 4 and touch once, then 2 and 8 which will register 4288 PIN

- **TouchGaze:** D. Best and A. Duchowski designed a rotary PIN entry layout which looks like the old telephones. They evaluated and compared their design against normal PIN layout and the final results shown that rotary design did slightly better in regards to accuracy. They used * symbol to start the PIN entry and a pound symbol to finish the entry. [BD12]
  Considering the speed of dwell against gaze path A. Kurauchi have experimented with Eyeswipe, which shown to be much more faster than the dwell-based method to type in a QWERTY keyboard. [Kur16]
  Like the other two methods, the 4-digit number is shown to the user and they move on to the layout. For our experiment, the users were shown the layout which has a small change to the previous layouts. We have the "R" button which stands for repeat. The user could start the entry process by touching any part of the screen and keeping the touch down. In the meantime, they can look for their desired number in the order they want and move from each button on to the next by their gaze. At the end when they are done they release the touch and the PIN is entered. In case that a number is repeated twice or more in a row the user could look at the number and then move to the R button to repeat the previous number. The amount of time that took for each user to enter each PIN was recorded and in case the password which was entered was different from what we expect from them, an error was recorded.

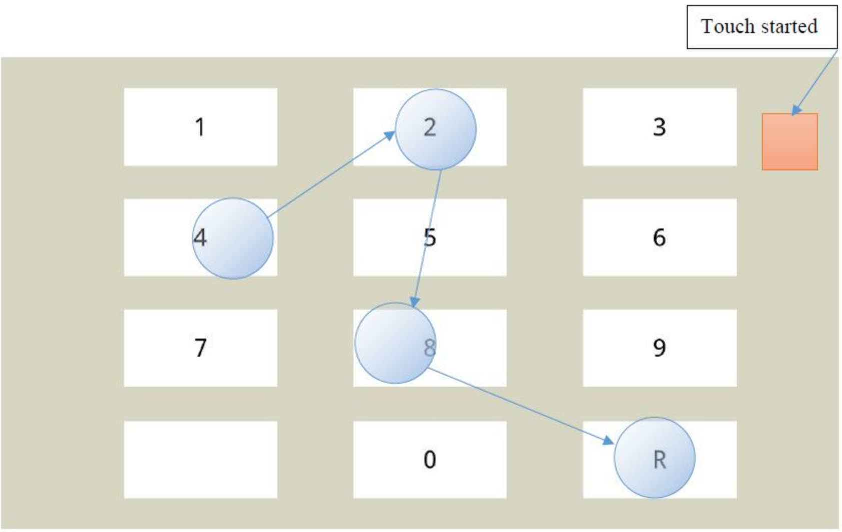
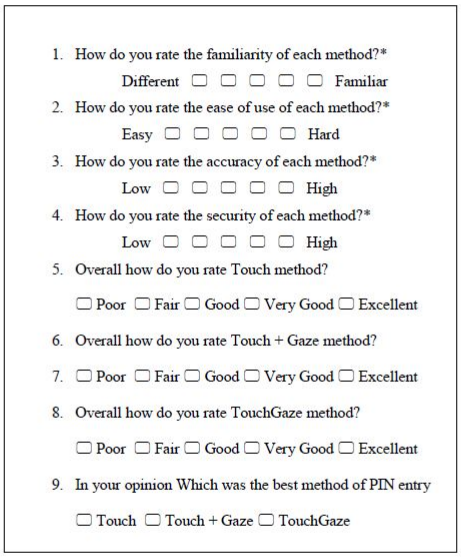


Figure 18: TouchGaze interface: The user touches any part of the screen and holds the touch, then looks at 4, 2, 8 and R to enter 4288

5.5.2 Qualitative Evaluation

At the end of the experiment, the users were given a questionnaire in which they could provide their feedback about the experiment and to rate each entry method based on ease of use, accuracy, speed, and security. For the last questions, we asked each participant about which was the best PIN entry method for them and if they have any feedback about them. In the Figure 19, the questions asked from the test subjects are shown. They can give their feedback and rate the three methods based on their experience.



1. How do you rate the familiarity of each method?*
   Different □ □ □ □ □ Familiar
2. How do you rate the ease of use of each method?*
   Easy □ □ □ □ □ Hard
3. How do you rate the accuracy of each method?*
   Low □ □ □ □ □ High
4. How do you rate the security of each method?*
   Low □ □ □ □ □ High
5. Overall how do you rate Touch method?
   □ Poor □ Fair □ Good □ Very Good □ Excellent
6. Overall how do you rate Touch + Gaze method?
7. □ Poor □ Fair □ Good □ Very Good □ Excellent
8. Overall how do you rate TouchGaze method?
   □ Poor □ Fair □ Good □ Very Good □ Excellent
9. In your opinion Which was the best method of PIN entry
   □ Touch □ Touch + Gaze □ TouchGaze

Figure 19: The questionnaire filled out by the participants after completing the evaluation. (The questions with * are asked for each method individually)

5.6 Results

In this master thesis, we hypothesized that TouchGaze can perform and compete with Touch and Touch + Gaze methods in regard to Speed and accuracy. The (ANOVA) method presented us with a significant result (F(2,51)=21.38785652, p<.05) among these methods. After analyzing the collected data during our evaluation with all three methods, this is the result for accuracy and meantime presented in the Figures 20 and 21.

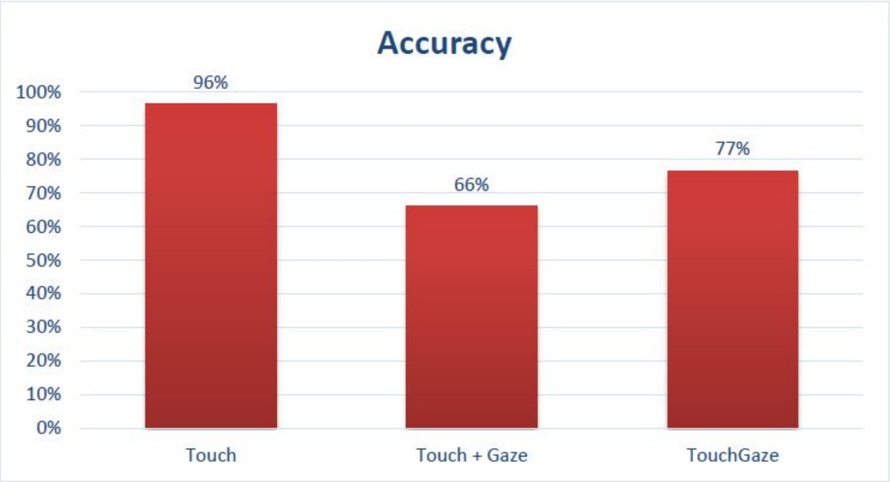


Figure 20: Accuracy across all three methods

As we can see in the chart the accuracy was the lowest in Touch + Gaze and as we expected it was very High for Touch method. This is due to participant's prior knowledge and experience of the Touch method. TouchGaze accuracy was recorded to be better compared to Touch+Gaze but still has a gap from the Touch method.

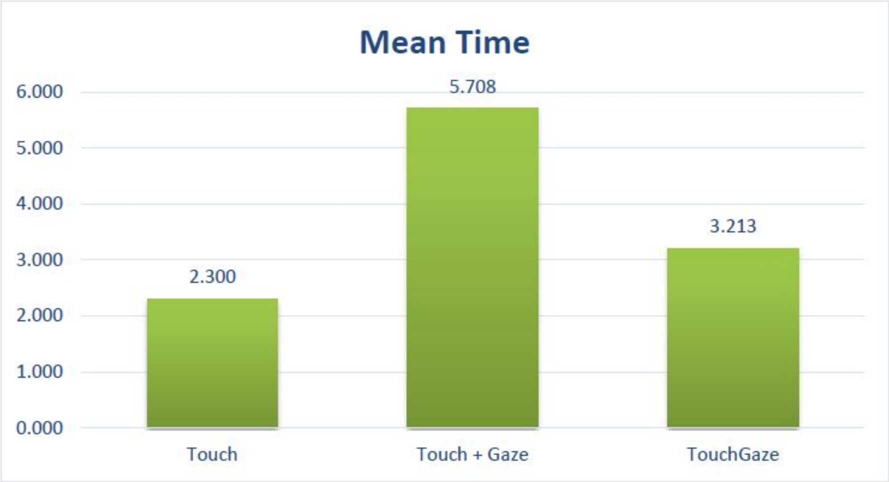


Figure 21: Mean Time for PIN entry across all three methods

As we can see in the Figure 21 the Touch + Gaze has the highest mean time among the three methods. However, it is significant to observe the gap between Touch and TouchGaze which is minimal. We anticipated the high mean time for Touch + Gaze, although the result for TouchGaze was not expected to be this low.

Each participant has 40 entered different PINs in each method. We divide the first 10 entry as the Training data and the remaining 30 as the Actual evaluation data. Now we can look deeper into our collected data to see how each participant performed in the Training stage compare to the evaluation stage. In the training stage the user might not be very familiar with the interaction and the entry method, but as he/she advances and gets experienced with the method, we expected the speed and accuracy would be improved.

In this Figure 22, we can observe the average time for each participant, to enter a PIN via the Touch method. As we expected almost for all test cases the timing has improved from training to the actual evaluation. For most of the users, the time gap is not considerable, because Touch is very familiar for most users and they have previous experience with it.

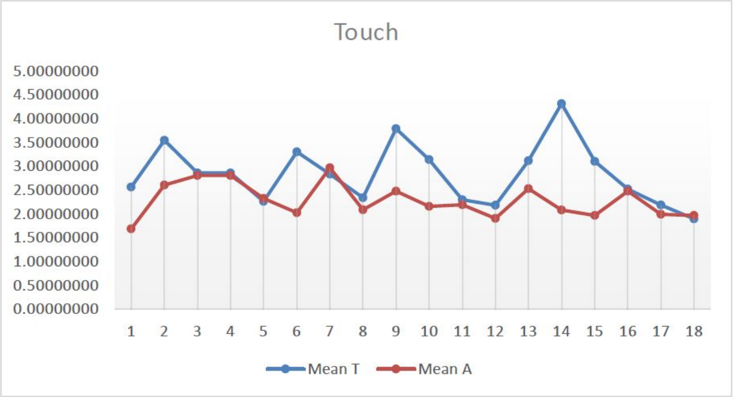


Figure 22: The mean time of Touch method during Training and Actual evaluation for all participants

As we can see in the next figure, all participants have a better mean time during actual evaluation. In the training stage, users needed more time to adapt to the new method. For 60 percent of our them, this was the first time using an eye tracking device. The maximum average time recorded was for the second participant, who didn't feel comfortable with the method and changed position often which affected the calibration of the eye tracker.

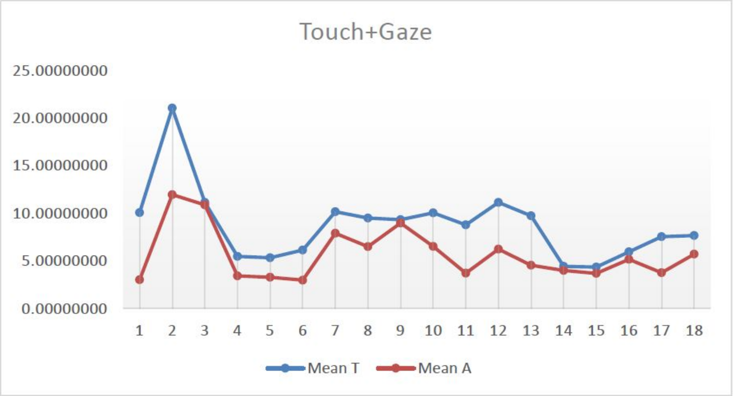


Figure 23: The mean time of Touch+Gaze method during Training and Actual evaluation for all participants

The TouchGaze does not follow the same pattern as the other two methods as we can observe in Figure 24. Only half of the participants fall into the same category in which their actual time is better than their training. However, in the rest of the cases, both training and actual average time are very close together. We could consider this as a result of the nature of the interaction itself. Although Touch + Gaze might be a new way of PIN entry it follows the same logic as the Touch. However, TouchGaze is a different story and needs more training to be comfortable for the user. Considering all this, the maximum mean time recorded belongs to the 9th participants and it is close to 6 seconds. Which is a quite a gap if we compare it to the 20 recorded in Touch + Gaze.

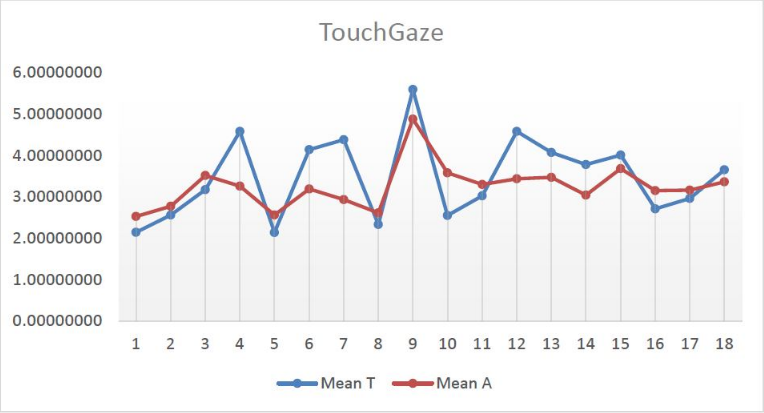


Figure 24: The mean time of TouchGaze method during Training and Actual evaluation for all participants

The second attribute we are going to discuss is the accuracy of each method.

It is not unexpected that all of the participants have an average accuracy of more than 90 percent both in training and actual stages in Touch method as we can see in Figure 25. It can be observed from the chart that many of test subjects scored 100 percent and didn't have any error. The classical PIN method entry should behave in this manner and not be pruned to errors. Some participants reported that they made a mistake during entry or forgot the PIN given to them to be entered altogether. Which is not the case in the real world since the user often knows and remembers the code.

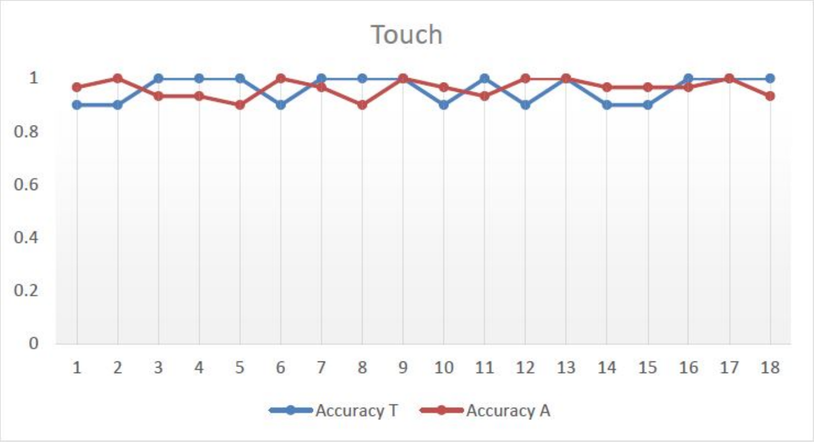


Figure 25: The accuracy of Touch method during Training and Actual evaluation for all participants

The average accuracy results in Touch + Gaze does stay with little difference. In Figure 26, there are some users like 12 and 17 who improved their accuracy compared to the training stage. It is important to know that most users were concerned about improving their time as we saw in the chart, therefore this could affect their accuracy. Once again we have user number 2 which had the highest average time, and here number 2 has the least accuracy among all participants.

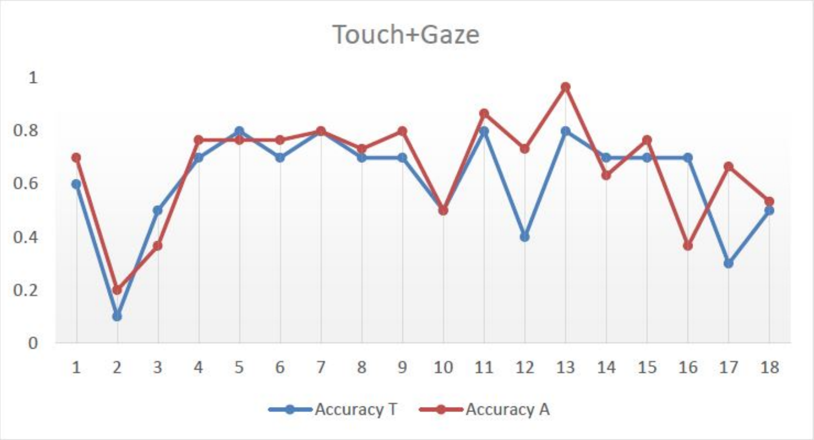


Figure 26: The accuracy of Touch+Gaze method during Training and Actual evaluation for all participants

The TouchGaze average accuracy seems promising in Figure 27 since almost all except 2 participants improved their accuracy by practicing and entering more PINs. This suggests that the method is practical and could gain user acceptance if they train with it. On the other hand, it could be argued that there is room for improvement of the system. The average training phase accuracy was 58 percent and it increases to 77 percent in the actual evaluation phase, which is quite a significant leap.

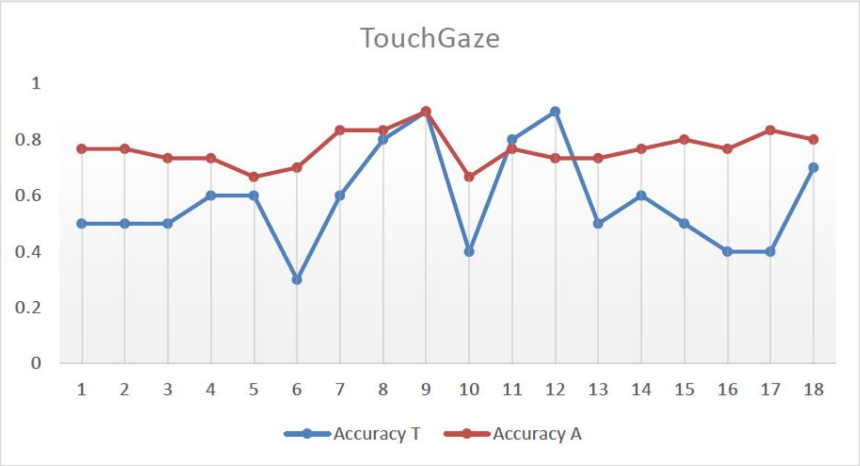
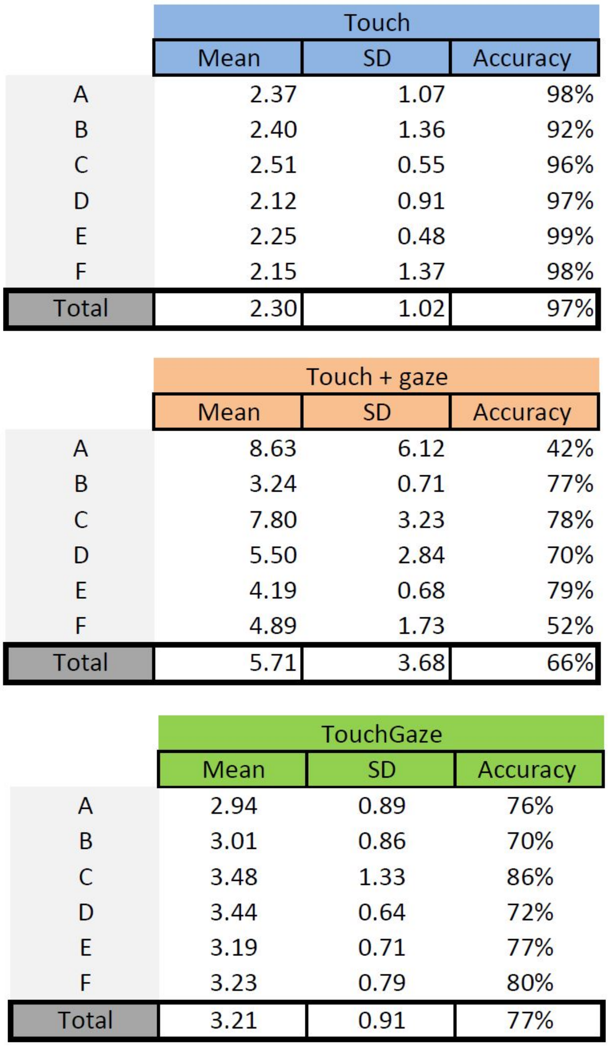


Figure 27: The accuracy of TouchGaze method during Training and Actual evaluation for all participants

In the Figure 28 we have calculated the meantime and Standard Deviation (SD) for each group, to complete the trial with our three designs. The tables present a higher deviation in Touch + Gaze when the standard deviation in Touch and TouchGaze are smaller. This shows that in Touch and TouchGaze we have much fewer outliers compare to Touch + Gaze. The information of the six different groups and the order which they have completed the methods shows that the ordering had little to no effect on the average time and accuracy.



	Touch		
	Mean	SD	Accuracy
A	2.37	1.07	98%
B	2.40	1.36	92%
C	2.51	0.55	96%
D	2.12	0.91	97%
E	2.25	0.48	99%
F	2.15	1.37	98%
Total	2.30	1.02	97%

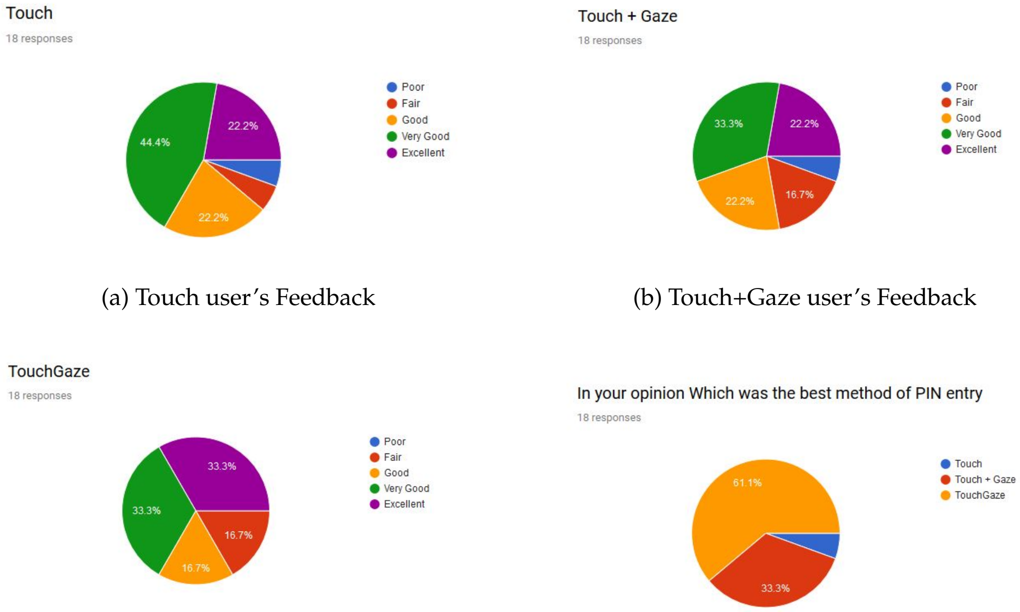
	Touch + gaze		
	Mean	SD	Accuracy
A	8.63	6.12	42%
B	3.24	0.71	77%
C	7.80	3.23	78%
D	5.50	2.84	70%
E	4.19	0.68	79%
F	4.89	1.73	52%
Total	5.71	3.68	66%

	TouchGaze		
	Mean	SD	Accuracy
A	2.94	0.89	76%
B	3.01	0.86	70%
C	3.48	1.33	86%
D	3.44	0.64	72%
E	3.19	0.71	77%
F	3.23	0.79	80%
Total	3.21	0.91	77%

Figure 28: The average time, standard deviation and accuracy of each group in our three methods.

At the end of the evaluation, we have asked each participant to fill out another questionnaire to give their feedback and opinions about each method. It was evident for us that the majority of the user is familiar with the Touch interaction. That is why they have mentioned that Touch was the easiest method for them to use. The Touch method is well established among all types of users and they practice this method many times every day, but based on our questionnaire more than 70 percent of our participants were new to the TouchGaze method. It is understandable that the participant did not have a prior experience with a method similar to TouchGaze, and it takes time for them to adapt to this method 72 percent of the users were satisfied with the accuracy of TouchGaze and 100 percent were satisfied with Touch.

However, when we asked the users opinion about the security of each method, the majority of them preferred Touch + Gaze and TouchGaze. Since they could see the merits of the two methods over Touch during the experiment. At the end of the questionnaire, we asked the participants to rate each method based on their opinion, from poor to excellent. All three got a good rating from our participants but the interesting fact is that nobody voted "Poor" for TouchGaze, when there were some who thought Touch and Touch+Gaze were poor. To compare all the three methods, we asked each participant to choose the best method in their point of view. TouchGaze was chosen by 11 persons, Touch + Gaze comes in second by 6 persons and Touch by only 1 person. The results are compelling for this experiment and displayed the potential of TouchGaze and how it could be part of the future designs.



(a) Touch user's Feedback

(b) Touch+Gaze user's Feedback

(c) TouchGaze user's Feedback

(d) User's overall preferred method

Figure 29: User's feedback

5.7 Discussion

In this section, we will discuss and interpret the results of the evaluation. Based on all entries we have collected, we have decided not to include the first 10 entries in each method, since we considered that as training data. However, the accuracy and the mean time of the Touch method did not improve considerably. It is safe to consider that participants didn't need any training for this method.

The case was different for Touch + Gaze and TouchGaze since they were fairly new methods to most users. The recorded data from the Touch + Gaze method suggests that the meantime and accuracy improve with time. But there were many reports from our participants that this method was tiring for them and their eyes. This could be a factor for the underwhelming results of Touch + Gaze.

The TouchGaze method structure is different from the other two methods and in the beginning, the error rate and timing were high for users. But after training, they improved rapidly in their time and accuracy. We can conclude that the users needed more training with the system and then the final results of the evaluation would have been much closer to the Touch method.

To test our idea about Touchgaze and Touch + Gaze, we can look at Figure 30 to see if practicing more is actually a factor in meantime and accuracy. In Figure 30 we are observing the error rate for two of our methods and the chart is demonstrating the extensive decline of errors in the last 10 entered PINs compared to the first 10 training entries. In the last 10 entries of TouchGaze, we have the average accuracy of 82 percent which is much higher than our total average of 77 percent mentioned before.

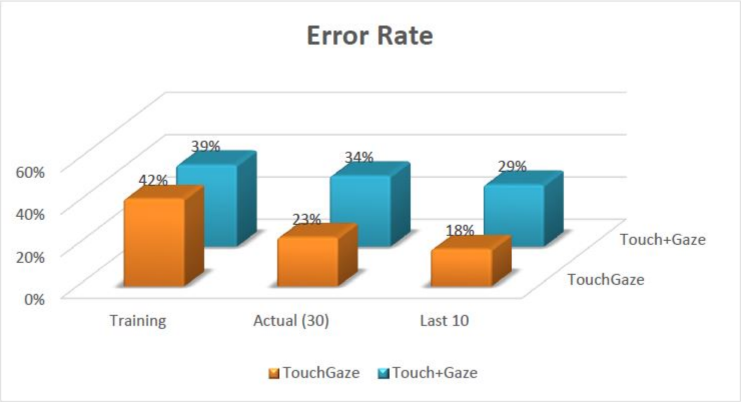


Figure 30: The error rate of TouchGaze and Touch + Gaze during training and in the last 10 entries

To study this case, even more, we can review the average time for the training and the last 10 entries. The Figure 31 demonstrate the reduction of average time from the training compared to the last 10 entries. For Touch + Gaze, we can see the decrease of 3.04 seconds in average time.

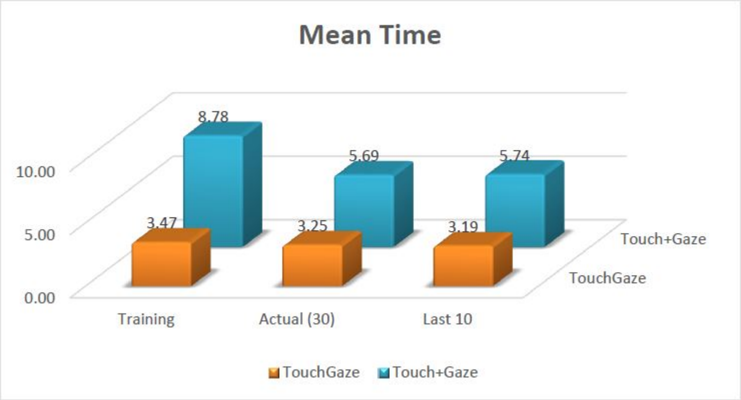


Figure 31: The mean time of TouchGaze and Touch + Gaze during training and in last 10 entries

We believe that TouchGaze's performance can be improved if the users have more training and experience like they have in Touch method. Even if we consider the current TouchGaze's accuracy which is 77 percent, we can compare it to the similar "Rotary" design. [BD12] According to their experiment and the data collected by them, the rotary design performed the average rate of the correctness of 71 percent which is below TouchGaze. And their keypad interface recorded the average accuracy of 64 percent which is below 66 percent accuracy of Touch + Gaze.

Overall we believe that TouchGaze exceeded our expectations and performed admirably.

6 Conclusion

In this Master thesis, we focused on shoulder surfing issue and other similar attacks that could compromise the security of PIN entry methods. We have combined Touch and Gaze to address these problems and implemented an innovative, yet secure authentication method. We analyzed many different types of research that have been done similar to our work and identified weaknesses and strengthens in their methods. There were many different approaches to shoulder surfing dilemma, which some of them were practical, but the majority of them tried to come up with complex interaction to improve the security. These methods would add extra steps to the interaction to either confuse the observer or make sure only the actual user would know how to enter the correct answer. We found few ideas which tried to use gaze to hide the targeting from the observer and in the meantime keep the method efficient, such as the Rotary design which was inspired by the classical rotary phones.

It is safe to assume that PIN entry methods will be part of the foreseeable future, however, the attacks on them are grave and could lead to a leak of information. It is true that different authentication methods like fingerprint are widespread. But the problem with biometric and token-based methods is that they can be stolen and used by a malicious individual to gain access to the user's sensitive information. In the password and PIN method, the concept is different and the users usually don't write them down and memorize them. Although they are still vulnerable to different types of attacks. The alternative authentication method proposed, challenges these wide ranges of attacks and makes the old eavesdropping methods obsolete. In the Touch + Gaze the user will enter each character one by one and for each of them touches the trigger once. Therefore, the observer can't know which digits have been entered but can know that the user touched for example 4 times and this will reveal that the length of the PIN is 4 digits. Therefore, the user's information might be compromised by a brute force attack. TouchGaze, on the other hand, provides users with an innovative PIN entry method. For each interaction, the user starts the entering process by initiating press touch and takes the path between the desired digits by gaze and finally submitting the PIN by simply releasing the touch. This method will not expose the numbers which the users have entered and it has an edge over the Touch + Gaze since the length of the PIN is not revealed and reduces the user's effort to touch the screen many times.

The experiment which we have undertaken had 3 different stages. Which each group of participants completed them in a different order

- **Touch** This stage proved to be the most accurate and efficient way to enter a PIN. Since all users have worked with this method for several years and it is widely accepted. However, when it comes to security this method is lacking the required features to protect the user's PIN. That is why most of our participants rate the security of this method very low and only 1 participant chose this method as the best overall method.

- **Touch + Gaze** This method demonstrated to be at the bottom of our list, due to low accuracy and high entry time. The method performed sub-par among our participants and compared to Touch method there was a large distance between their overall average time. But the users rated the security of this method higher than the Touch method.

- **TouchGaze** The TouchGaze method showed that it can be a contender and demonstrated that the TouchGaze method requires slightly more time for each entry compares to classical methods like the Touch. The accuracy of Touchgaze is acceptable considering the contemporary state of its design. And the results proved that by training, user's can perform better in regards to accuracy. If we consider the security of this method, most users rated the method as very secure and overall they chose this method as their best method.

TouchGaze and Touch + Gaze methods are designed on the shoulder of the Touch method and we tried to improve the already proved method. The comparison of the TouchGaze and Touch method is about the requirements of the user and their consideration of the strength and weaknesses of each method. The Touchgaze method could have many real-world applications and could become an additional authentication method which provides higher security and could pave the way for new methods to emerge.

7 Future Work

For further studies, we propose the testing of the Touchgaze in a larger group of participants, which is done in a different world applicable setting. The further evaluation should be done in different phases and needs time for the users to find familiarity with the method so that the user acceptance of TouchGaze could be evaluated. The method should be applied and tested on smartphones with the usage of the already built-in camera, instead of an eye-tracking data. This could be challenging due to the small screens in smartphones and also the accuracy of the gaze position. We have a few other ideas for prospective work, which are as follows.

- **Security analysis** TouchGaze and Touch + Gaze are resilient towards thermal and smudge attacks by design, however, the methods were not tested for observer attacks. In future works, we want to do a security analysis, to confirm our hypothesis that a third party could not identify the PIN. The threat model should be done in two different ways, active and passive attacker. In the active method, an observer should be physically present when the user is entering his/her PIN. And the attacker can observe the number of identified PINs. In the passive method, the interactions of the user are recorded by a camera and then used to identify the PIN. Many different observers could see the video and try to detect the correct PIN. By doing this experiments, we can discover the security rate of the methods.

- **Improving usability** The data gathered from our evaluation demonstrated that the usability of TouchGaze has room for improvement. Further research should be done in the future to enhance the method and improve the ease of use for users. It is possible to create another experiment with the same participants and do them in different stages. So that the users are familiar with the interaction method and have experience. Then we can identify the ways we can improve the usability of the TouchGaze. It is possible to change the user interface slightly to examine if it would improve the timing and accuracy. The other point we can focus on is regarding the R button since not all users can comprehend the concept and might create confusion for them. So it is possible to work on that front as well.

- **User Profile** It is feasible to collect the data for each user during our interaction and use machine learning techniques to create a profile of each user and the way they proceed to enter their PIN. Therefore, this would add an extra security measure for the method. So whenever a malicious individual wants to hack into someone's account, they need more than the PIN and need to have the similar behavior to the user. For instance, collecting the data regarding the Touchpoint of the users. We can record this data and the system would analyze the similarity of the interaction regarding timing, touch point, and eye fixations. This extra security measure should be tested again in regards to efficiency, accuracy, and security.

Overall we can do much more experiment with TouchGaze and add additional features to improve the usability and the security of our method. The important point to be acknowledged is that security and usability should be in a harmonic balance. Neither one should be compromised to benefit the other. The method and the improved version of the method must follow the mentioned requirements for a successful authentication method.

8 Bibliography

References

[AGM$^+$10] Adam J Aviv, Katherine L Gibson, Evan Mossop, Matt Blaze, and Jonathan M Smith. Smudge attacks on smartphone touch screens. *Woot*, 10:1–7, 2010.

[BD12] Darrell S Best and Andrew T Duchowski. A Rotary Dial for Gaze-based PIN Entry. pages 69–76, 2012.

[Clu15] Graham Cluley. This simple iphone case can be used to steal atm pins, Jan 2015.

[DFO08] Paul Dunphy, Andrew Fitch, and Patrick Olivier. Gaze-contingent passwords at the ATM. (2):2–5, 2008.

[DMS] Edwin S Dalmaijer, Sebastiaan Mathôt, and Stefan Van Der Stigchel. PyGaze : an open-source , cross-platform toolbox for minimal-effort programming of eye-tracking experiments. pages 1–16.

[HDG13] Sabrina Hoppe, Florian Daiber, and Dfki Gmbh. Eype - Using Eye-Traces for. pages 1–4, 2013.

[Jac91] J K Jacob. The Use of Eye Movements in Interaction Techniques : What You Look At is What You Get. 9(3):152–169, 1991.

[Joa10] Hans Joachim. Authentication methods. pages 1–48, 2010.

[KAH16] Mohamed Khamis, Florian Alt, and Mariam Hassib. GazeTouchPass : Multimodal Authentication Using Gaze and Touch on Mobile Devices. pages 2156–2164, 2016.

[KHBA17] Mohamed Khamis, Regina Hasholzner, Andreas Bulling, and Florian Alt. GTmoPass : Two-factor Authentication on Public Displays Using Gaze-Touch passwords and Personal Mobile Devices. 2017.

[KM$^+$17] Chandan Kumar, Raphael Menges, , Daniel Müller, and Korok Sengupta. Gazetheweb: A gaze-controlled web browser. In *Proceedings of the 14th Web for All Conference on The Future of Accessible Work*, page 25. ACM, 2017.

[KMS16] Chandan Kumar, Raphael Menges, and Steffen Staab. Eye-Controlled Interfaces for Multimedia Interaction. pages 6–13, 2016.

[Kum07] Manu Kumar. USER INTERFACE DESIGN. (May), 2007.

[Kur16] Andrew Kurauchi. EyeSwipe : Dwell-free Text Entry Using Gaze Paths. pages 1952–1956, 2016.

[LHH10] Alexander De Luca, Katja Hertzschuch, and Heinrich Hussmann. ColorPIN – Securing PIN Entry through Indirect Input. pages 1–4, 2010.

[Luc08] Alexander De Luca. EyePass - Eye-Stroke Authentication for Public Terminals. pages 3003–3008, 2008.

[LW07] Alexander De Luca and Roman Weiss. Evaluation of Eye-Gaze Interaction Methods for Security Enhanced PIN-Entry. (November):28–30, 2007.

[PAG16] Ken Pfeuffer, Jason Alexander, and Hans Gellersen. Partially-indirect bimanual input with gaze, pen, and touch for pan, zoom, and ink interaction. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 2845–2856. ACM, 2016.

[roo16] Smudge attacks - upgrade your phone's passcode to at least six digits, Oct 2016.

[SBG09] Dominik Schmidt, Florian Block, and Hans Gellersen. A comparison of direct and indirect multi-touch input for large surfaces. In *IFIP Conference on Human-Computer Interaction*, pages 582–594. Springer, 2009.

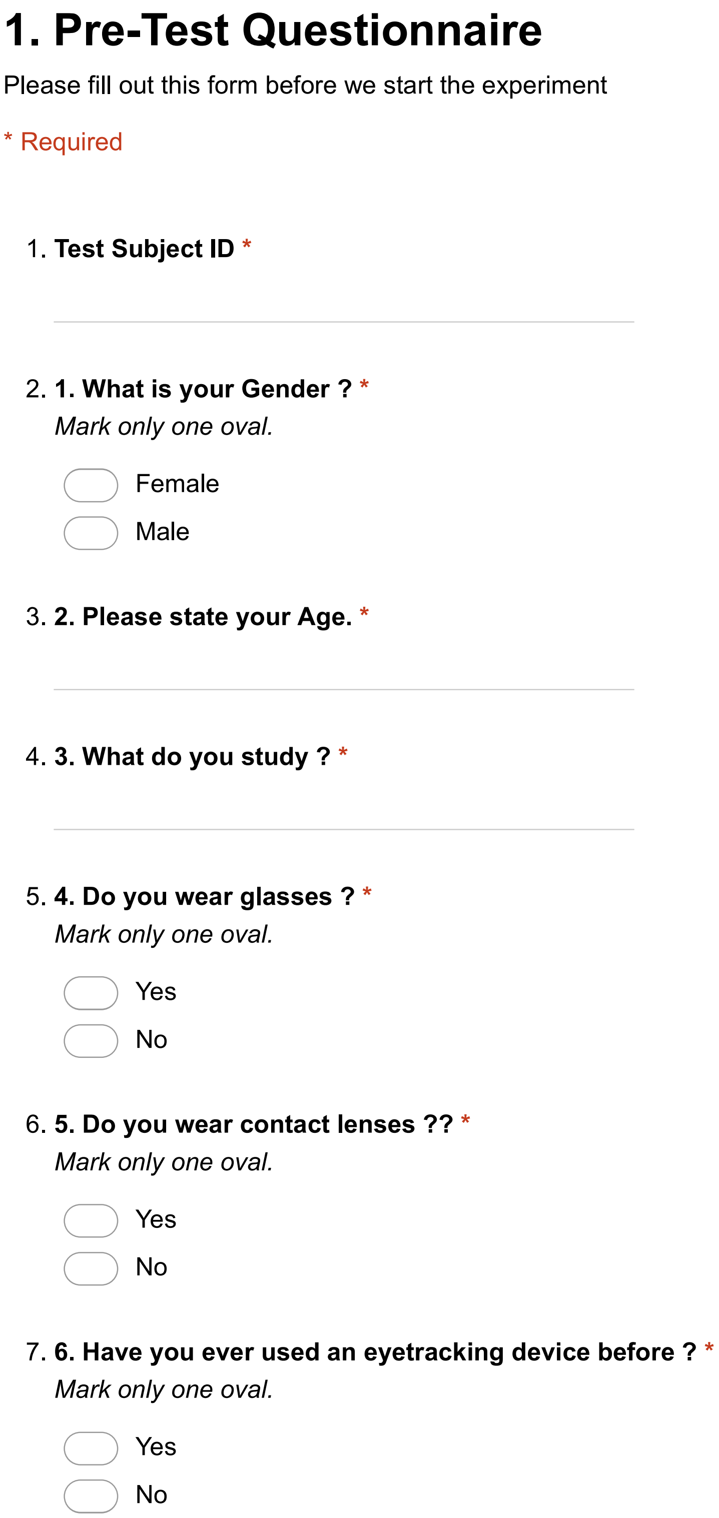
[SMI09] iView X System Manual iView X Manual. (August), 2009.

[VMSB15] Simon Voelker, Andrii Matviienko, Johannes Schöning, and Jan Borchers. Combining direct and indirect touch input for interactive workspaces using gaze input. In *Proceedings of the 3rd ACM Symposium on Spatial User Interaction*, pages 79–88. ACM, 2015.

[ZK04] By Shumin Zhai and Per Ola Kristensson. The Word-Gesture Keyboard : Reimagining Keyboard Interaction. pages 43–52, 2004.

List of Figures

Appendices



1. Pre-Test Questionnaire

Please fill out this form before we start the experiment

* Required

1. **Test Subject ID** *

2. **1. What is your Gender ?** *
   Mark only one oval.
   ○ Female
   ○ Male

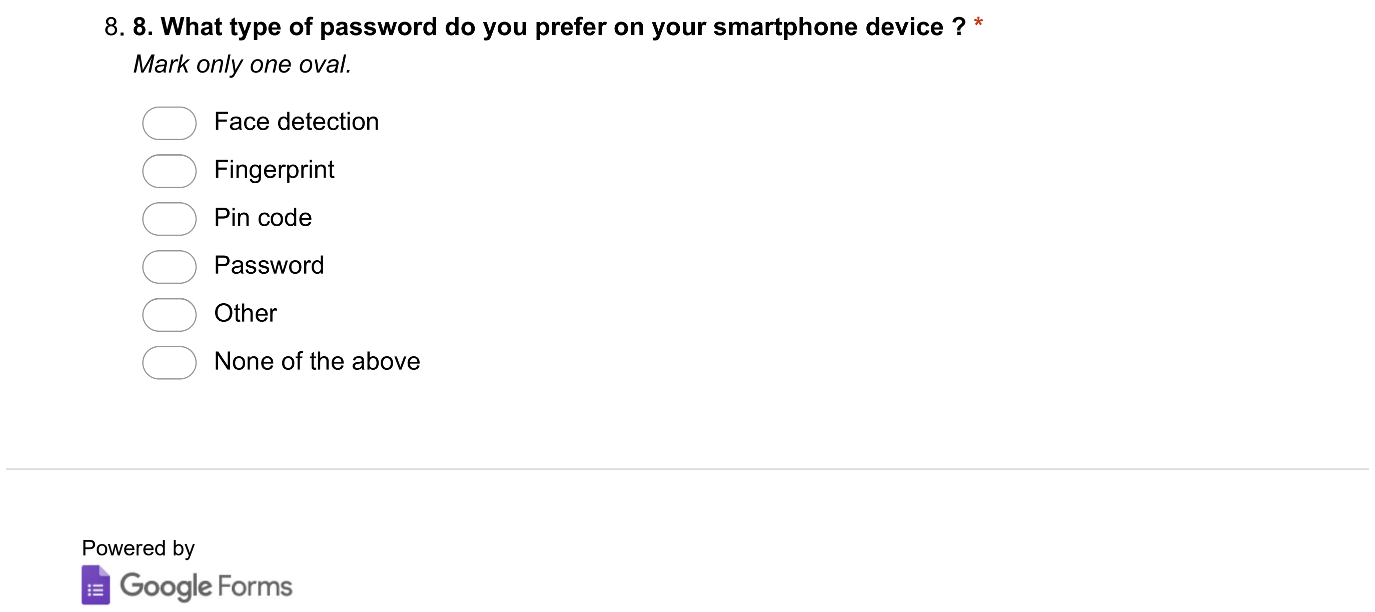
3. **2. Please state your Age.** *

4. **3. What do you study ?** *

5. **4. Do you wear glasses ?** *
   Mark only one oval.
   ○ Yes
   ○ No

6. **5. Do you wear contact lenses ??** *
   Mark only one oval.
   ○ Yes
   ○ No

7. **6. Have you ever used an eyetracking device before ?** *
   Mark only one oval.
   ○ Yes
   ○ No



8. **8. What type of password do you prefer on your smartphone device ?** *

Mark only one oval.

○ Face detection

○ Fingerprint

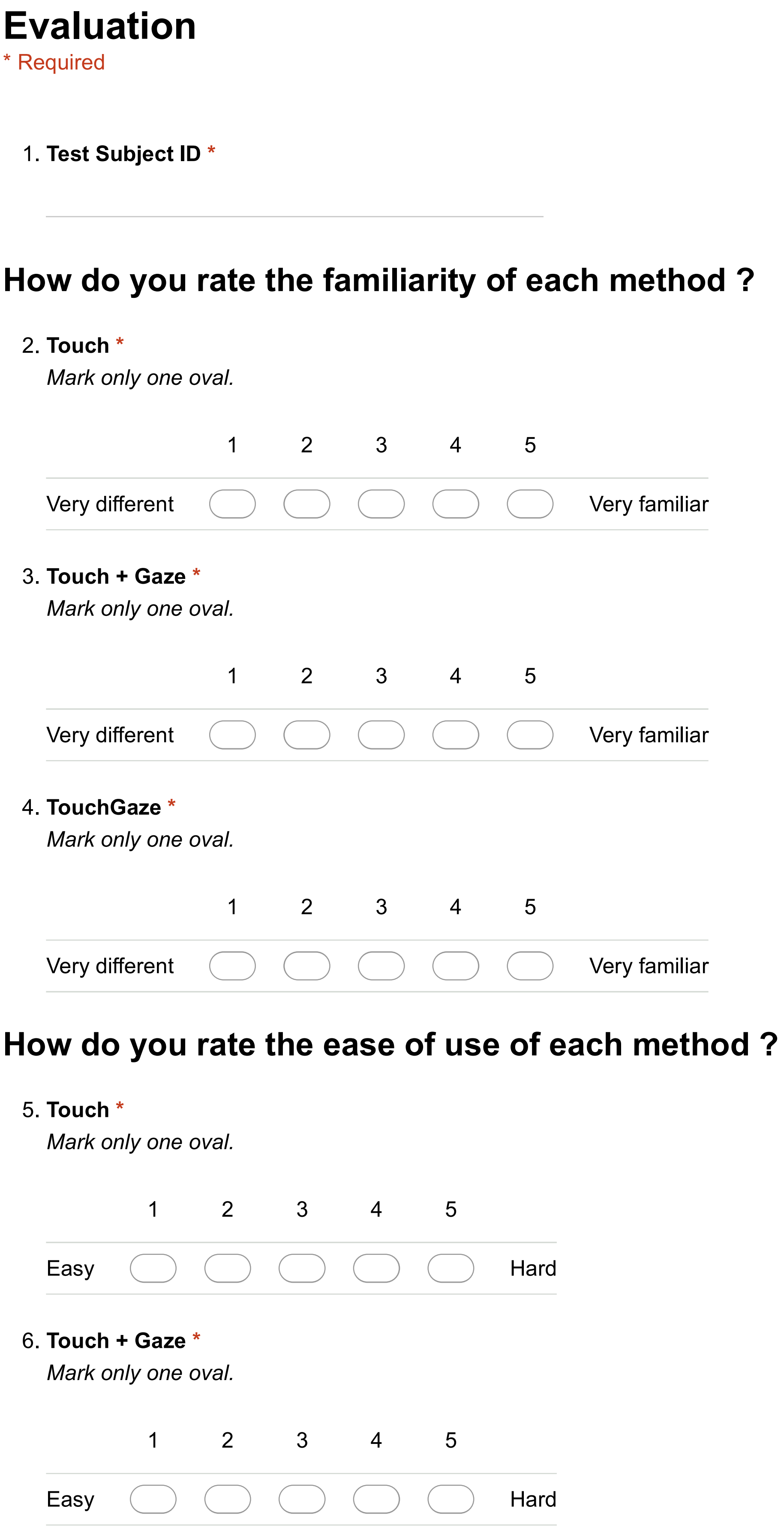
○ Pin code

○ Password

○ Other

○ None of the above

Powered by Google Forms



Evaluation

* Required

1. **Test Subject ID** *

How do you rate the familiarity of each method ?

2. **Touch** *

Mark only one oval.

	1	2	3	4	5	
Very different	○	○	○	○	○	Very familiar

3. **Touch + Gaze** *

Mark only one oval.

	1	2	3	4	5	
Very different	○	○	○	○	○	Very familiar

4. **TouchGaze** *

Mark only one oval.

	1	2	3	4	5	
Very different	○	○	○	○	○	Very familiar

How do you rate the ease of use of each method ?

5. **Touch** *

Mark only one oval.

	1	2	3	4	5	
Easy	○	○	○	○	○	Hard

6. **Touch + Gaze** *

Mark only one oval.

	1	2	3	4	5	
Easy	○	○	○	○	○	Hard

7. **TouchGaze** *

Mark only one oval.

	1	2	3	4	5	
Easy	○	○	○	○	○	Hard

How do you rate the accuracy of each method ?

8. **Touch** *

Mark only one oval.

	1	2	3	4	5	
Low	○	○	○	○	○	High

9. **Touch + Gaze** *

Mark only one oval.

	1	2	3	4	5	
Low	○	○	○	○	○	High

10. **TouchGaze** *

Mark only one oval.

	1	2	3	4	5	
Low	○	○	○	○	○	High

How do you rate the security of each method ?

11. **Touch** *

Mark only one oval.

	1	2	3	4	5	
Low	○	○	○	○	○	High

12. **Touch + Gaze** *

Mark only one oval.

	1	2	3	4	5	
Low	○	○	○	○	○	High

13. **TouchGaze** *

Mark only one oval.

	1	2	3	4	5	
Low	○	○	○	○	○	High

Overall how do you rate this method ?

14. **Touch** *

Mark only one oval.

○ Poor
○ Fair
○ Good
○ Very Good
○ Excellent

15. **Touch + Gaze** *

Mark only one oval.

○ Poor
○ Fair
○ Good
○ Very Good
○ Excellent

16. **TouchGaze** *

Mark only one oval.

○ Poor
○ Fair
○ Good
○ Very Good
○ Excellent

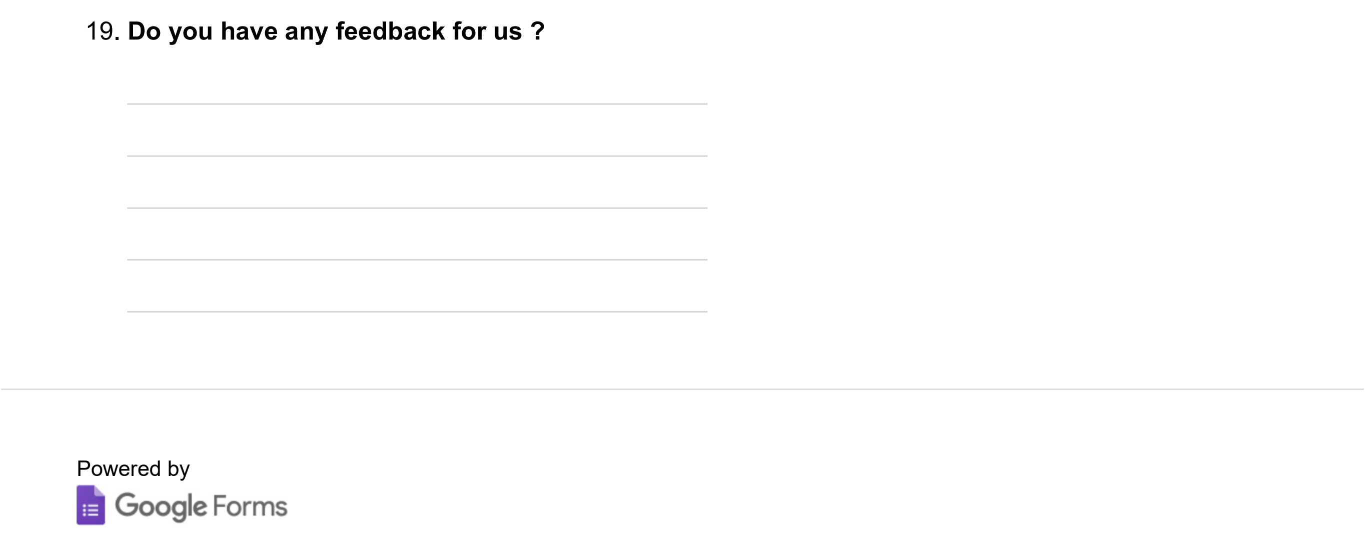
Feedback

17. **In your opinion Which was the best method of PIN entry** *

Mark only one oval.

○ Touch
○ Touch + Gaze
○ TouchGaze

18. **Do you have any suggestion to improve the PIN entry ?**



19. **Do you have any feedback for us ?**

Powered by

Google Forms