Freie Universität Berlin

Bachelorarbeit am Institut für Informatik der Freien Universität Berlin

Arbeitsgruppe Informationssicherheit

Usable Confirmation Method in Online Banking Transactions

Amirhossein Rajabi Pour Kiakolaie
Matrikelnummer: 5090105
amir.rajabi@fu-berlin.de

Betreuer/in: Sandra Kostic
Eingereicht bei: Prof. Dr. Marian Margraf
Zweitgutachter/in: Prof. Dr.-Ing. Maija Poikela

Berlin, August 19, 2024

Selbstständigkeitserklärung

Ich erkläre gegenüber der Freien Universität Berlin, dass ich die vorliegende Bachelorarbeit selbstständig und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel angefertigt habe.

Die vorliegende Arbeit ist frei von Plagiaten. Alle Ausführungen, die wörtlich oder inhaltlich aus anderen Schriften entnommen sind, habe ich als solche kenntlich gemacht.

Diese Arbeit wurde in gleicher oder ähnlicher Form noch bei keiner anderen Universität als Prüfungsleistung eingereicht.

Berlin, den 19. August 2024

Amirhossein Rajabi Pour Kiakolaie

Ami

Abstract

Overlay attacks are a sophisticated threat to online banking platforms, and methods that detect such attacks when technical safeguards fail are absent. This bachelor thesis evaluates a novel method that assists users in detecting overlay attacks during transaction confirmation processes. The research aimed to assess the method's suitability in detecting overlay attacks and its usability through a mixed-methods research approach, leveraging System Usability Scale (SUS) evaluations, overlay detection success rates, and user feedback.

The study involved creating an instructional video to introduce users to overlay attacks and the confirmation method, developing and testing a prototype of this method, and administering two questionnaires—one before and one after the user's interaction with the prototype. The research addressed multiple objectives and research questions, for instance, evaluating the method's suitability in detecting overlay attacks, assessing the system's usability using the SUS, testing the assumptions regarding the design elements, and exploring the correlation between SUS scores and detection success rates.

The results indicate that the method is suitable for detecting overlay attacks for most participants. However, some participants experienced difficulties, and a downtrend in the mean successful attack detection was observed with aging. The mean SUS score reflected good usability but also highlighted areas for improvement. User feedback pointed out several issues, including the cognitive load imposed by the method and the mixed effectiveness of the visual elements.

This study concludes that while the proposed method demonstrates potential for enhancing the security of online banking transactions against overlay attacks, further refinements are needed to improve both its usability and security.

Contents

1 Introduction

In the first chapter of this thesis, we explore the critical need for secure yet usable confirmation methods in online banking. This chapter includes an overview of online banking, highlighting the rapid growth in digital transactions and the concurrent rise in sophisticated cyber threats. We then delve into multi-factor Authentication, discussing various methods and their effectiveness in securing mobile banking. The problem statement identifies key vulnerabilities, particularly focusing on the risk of overlay attacks in transaction confirmation apps. Finally, the chapter outlines the objectives of the study and research questions that will be answered.

1.1 Overview of online banking

With the rapid growth of online money transfers and the digitalisation of the economy, the number of people using online transactions is increasing steadily and quickly [58]. After the COVID-19 outbreak, the number of online banking users increased significantly, with many people opting for digital solutions due to restrictions and safety concerns related to visiting physical bank branches [43]. By 2021, there were approximately 2.5 billion users worldwide, and by the end of 2024, the number is expected to exceed 3.5 billion users worldwide [49], achieving a penetration rate of 42% worldwide, based on today's population [69]. China and far East countries are expected to reach 974 million online banking users, with an approximately 42% penetration rate in 2024 [50, 63]. North America has the highest relative penetration rate by 70% and Europe with 58% of its population using online banking has the second highest relative penetration rate, doubling since 2012 [17, 77, 78].

In addition to traditional banks offering various online services, the growth of neobanks, which are digital-only banks without physical branches, has been remarkably fast and they are processing an increasingly large volume of transactions. In 2023, the total value of transactions conducted through neobanks globally reached approximately 5 trillion US dollars, a significant increase from 2.56 trillion US dollars in 2021 [64]. This growth trend is expected to continue, with projections indicating a total of 10 trillion US dollars by 2028. Furthermore, by the end of 2023, there were 250 million neobank users, with an average transaction value of approximately 20 thousand US dollars in this year. This number is expected to increase to 386 million users, with an average transaction value of 27 thousand US dollars, by 2028 [64].

Such growth, while offering undeniable convenience, requires a multi-layered approach to securing transactions. As financial activities move online, so do the sophisticated strategies of cybercriminals. From phishing schemes to malware attacks, user funds are under constant threat. According to a report by SEON [35], the total losses from fraud in the banking sector reached nearly 1.6 billion US dollars in 2022, with mobile banking fraud being a significant contributor. The cost of fraud for financial institutions is substantial, with every 1 US dollar of fraud costing U.S. financial services about 4.23 US dollars when considering legal, processing, and recovery expenses [35]. As another example, the United Kingdom, with 98% online banking penetration and holding first place worldwide, suffered a staggering 7,532 reported cases of fraud only in the banking sector during the second quarter of 2023, resulting in losses exceeding £120 million [45, 44, 63].

Certainly, robust security measures like encryption mitigating these threats and creating a safe infrastructure for online transactions. Strong encryption scrambles data during transmission, making it virtually impossible for attackers to decipher and MFA adds an extra layer of security by requiring users to verify their identity through a secondary factor beyond just a password [59]. Moreover, monitoring online transactions plays a crucial role in preventing cyber attacks by detecting and responding to suspicious activities in real-time [22]. Furthermore, Multi-factor authentication (MFA) has become the basis of security in online banking across the globe [4]. This robust approach requires users to provide multiple forms of verification, significantly preventing unauthorised access to sensitive financial information [56].

Another important aspect of security in online transactions is educating users and society about best practices [23]. By understanding common scams, utilising strong passwords, and practicing smart online habits, individuals become an important line of defense against online financial threats. This collective caution, in addition to robust security measures, promotes trust and safeguards financial well-being in this evolving landscape.

Cyberattacks on online banking transactions employ a variety of sophisticated tactics designed to exploit vulnerabilities in digital systems. World economic forum report [75] indicates that 56% of leaders anticipate generative AI to empower cyberattacks within two years, particularly in phishing, malware development, and misinformation However, accurately assessing the true impact of these attacks is challenging because detailed information about the incidents is often withheld or not disclosed [16].

1.2 Authentication in online banking

MFA can be categorised into three classic types, and each category offers some specific authentication methods [55].

- knowledge-based factors, such as passwords, PINs, or security questions
- Possession-based factors involve physical items like smartphones, security tokens, or smart cards
- Biometric identifiers, such as fingerprint, facial recognition, iris scans, and voice patterns

Specific statistics on the exact breakdown of MFA methods used globally can be challenging to pinpoint due to the privacy of data and varying implementation practices by different banks and regions, consequently, we will observe the first five German banks with the most customers based on a survey, which according to Sparkasse, Volksbank, Commerzbank, ING, Post Bank are the first five banks with the most customers [33]. Reviewing these banks' websites [61, 73, 7, 24, 48] indicated that password and PIN authentication are fundamental for all of these banks. Four out of five banks support biometrics as an additional option for logging into the banking app and all of them offer Push-TAN. Push-TAN is a TAN generated in a second application additional to the online banking app itself, allowing customers to confirm any transaction after receiving a push notification without manually entering the transaction authentication number. Another method offered by all these banks (with some branches of Sparkasse) is Photo-TAN, which also uses a second application for confirming transactions [61, 73, 7, 24, 48].

1.3 Problem statement

The problem that the suggested method for a secure transaction aims to tackle can affect all the aforementioned authentication methods. Many online banking transactions rely on confirmation via a TAN app [61, 73, 7, 24, 48]. Additionally, statistics show that SMS-TAN and biometrics combined with passwords are also commonly used methods [14, 18]. All of these methods add an extra layer of security to transaction confirmation despite their drawbacks. However, an overlay attack on these apps can enable hackers to hide the underlying transaction information, tricking users into unknowingly transferring their money to the hackers. By creating an overlay picture that appears authentic on the user's device, concealing the hacker's bank information with the information the user expects to see, users may unknowingly authorise a transfer without realising the recipient is a hacker. Consequently, there would be no technical evidence to indicate that such an attack occurred on the user's device at the time of the transfer, complicating any legal actions the user might pursue later.

The primary objective of this research is to evaluate a novel transaction confirmation method that combines design elements with technical measures to detect and prevent overlay attacks on online banking platforms. The proposed method aims to help users detect overlay attacks on their online banking apps by using a special generated background for each transaction.

1.4 Objectives and research questions

The objectives of this study are diverse, aiming to thoroughly evaluate different aspects of user interaction with the proposed method and its system design in detecting overlay attacks.

First, the study aims to evaluate the suitability of this method for detecting overlay attacks. This involves examining how effectively users can identify both attack and safe scenarios. Additionally, the evaluation process includes assessing the system's usability using the System Usability Scale (SUS) [5], which provides a quantitative measure of the system's overall ease of use and user satisfaction. This evaluation is crucial for understanding the system's practical usability in real-world conditions and for identifying any potential usability issues that could hinder user performance

Second, the study seeks to gather and analyse user feedback. By capturing detailed feedback from users, the study aims to identify patterns and trends in user experiences and preferences. This analysis will help uncover common challenges and areas for improvement, providing valuable insights that can inform future enhancements to the system.

Third, the study proposes a user interface (UI) design that prioritises reliability and usability. The design of the key system elements was developed with careful consideration of the diverse needs of different user groups. This proposed design can be implemented directly or used as a foundation for further user studies to refine and identify the most suitable UI design.

Finally, the study involves hypothesis testing on attack detection. The study makes assumptions about the difficulty users might face in detecting various scenarios in overlay attacks. These assumptions are then checked through detailed analysis to see if they hold.

This research also aims to address specific research questions to further understand the factors influencing the detection of overlay attacks and the overall usability of the system:

Research Question 1: Is there a correlation between the users' System Usability Scale (SUS) ratings and their performance in detecting overlay attacks and identifying safe transactions?

Research Question 2: How helpful was the information delivered in the introductory video in preparing users to use the confirmation method effectively?

2 Secure online banking background

In this chapter, we will first discuss the various attacks on online banking, focusing on the most common and damaging types. Following this, we will examine overlay attacks in detail. At the end of this chapter, we will review related works and research conducted on this subject.

2.1 Attacks on online banking

One of the most common methods of attacking online banking applications is the use of mobile malware. It includes various malicious software designed specifically to target mobile devices. Trojans, spyware, and ransomware are examples of such malware [13]. Trojans appear to be legitimate applications but execute harmful activities once installed, such as stealing sensitive information [30]. Spyware secretly monitors user activity, capturing data like login credentials and financial information [38]. These malicious apps can significantly compromise the security of mobile banking applications.

Zimperium's Mobile Banking Heists Report claimed that 29 different malware families were detected in 2023, targeting 1,800 banking apps across 61 countries. A staggering 2,178 variants across at least ten distinct banking malware families were identified globally, highlighting the escalating sophistication of cyber threats targeting the financial sector [83, 82].

Another prevalent attack vector is phishing where attackers deceive users into revealing personal information by pretending to be a trustworthy entity in electronic communications. This is often executed via emails, SMS, or fraudulent websites that mimic legitimate banking sites. Users are tricked into entering their login credentials or other sensitive information, which attackers then use to gain unauthorised access to their bank accounts [26]. In the year 2023 around 11% of all phishing attacks worldwide targeted banks as shown in figure 1, and another report shows that more than 30% of financial phishing attacks are targeted at Banks [27, 31].

Another type of attack is the Man-in-the-Middle (MitM) attack. It occurs when an attacker secretly intercepts and relays messages between two parties who believe they are directly communicating with each other [15]. In the context of mobile banking, an attacker can intercept the data being transmitted between the user's mobile device and the bank's server. This allows the attacker to capture sensitive information or inject malicious data into the communication stream without the knowledge of either party.

The fourth type of attack that can target mobile banking applications is session hijacking. Session hijacking involves exploiting a valid computer session to gain unauthorised access to information or services in a computer system [54]. Attackers can hijack sessions in mobile banking applications by stealing session cookies or session tokens. Once they obtain these tokens, they can impersonate the user and perform unauthorised transactions[6].

Another type of attack that could facilitate an overlay attack is a zero-day attack.

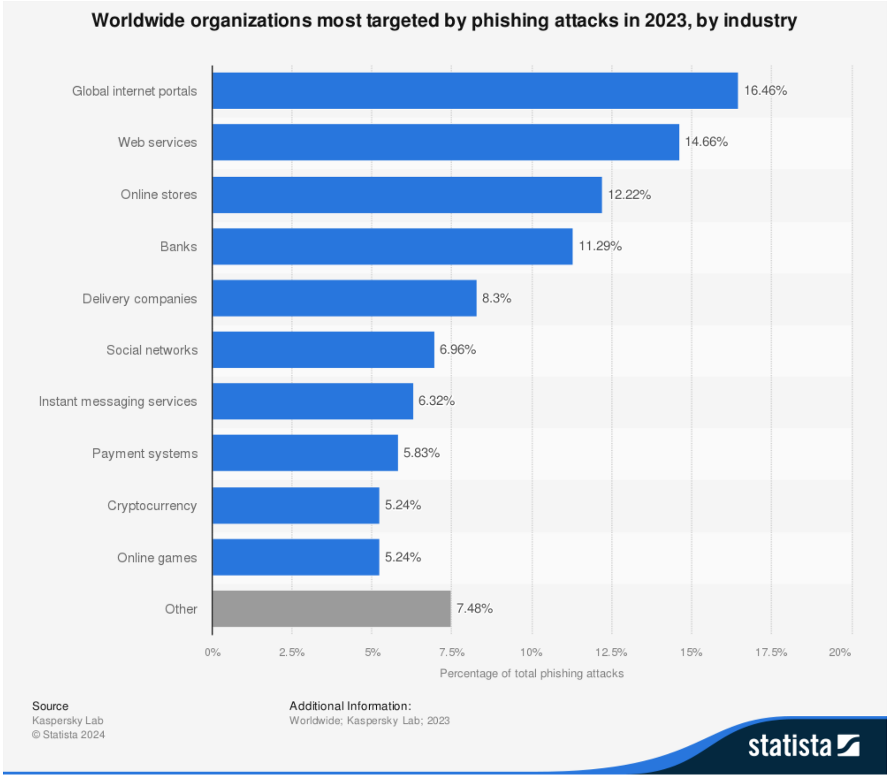


Figure 1: Most targeted organizations worldwide by pishing in 2023 [31]

According to kaspersky [29] a zero-day attack exploits a vulnerability in software or hardware that is unknown to the vendor, leaving no time for the vendor to develop and distribute a patch or fix. They also state that this type of attack is particularly dangerous because it bypass standard security measures and remain undetected for a significant period of time.

2.2 What is an overlay attack

Among the various possible cyber-attacks on mobile devices, overlay attacks are particularly sophisticated. These attacks allow hackers to partially cover the device's screen and mimic the legitimate interface, making it difficult for users to detect the attack [81]. The first step in an overlay attack involves infecting the mobile device through the methods previously mentioned. The malicious action can remain undetected on the device until the opportune moment. Only then does the second part of the attack, which is the creation of the overlay, take place. A significant danger of overlay attacks is that even vigilant users may not detect them if the recreation of the user interface is flawless. The third part of the attack occurs when the user continues to interact with the malicious interface. At this point, the likelihood of a successful attack is very high, and after this step, the stolen information can be sent to attackers or a fraudulent transaction can be completed [70, 71].

For instance, an attacker could target an online banking application by capturing the transfer amount, IBAN, name, and other essential details while replacing the destination IBAN with their own. The user would unknowingly confirm the transaction based on the displayed information, unaware of the underlying manipulation. Similar attacks can be perpetrated on cryptocurrency exchanges and mobile wallets, where a long wallet address is the only requirement for transferring funds. If malware successfully activates and the user initiates a cryptocurrency transfer, attackers can create an overlay mimicking the expected interface. The legal repercussions of such cryptocurrency attacks are potentially more complex than those involving traditional banking, as recovering stolen funds might be more challenging. Nevertheless, these attackers likely possess efficient methods for cashing out the illegal proceeds.

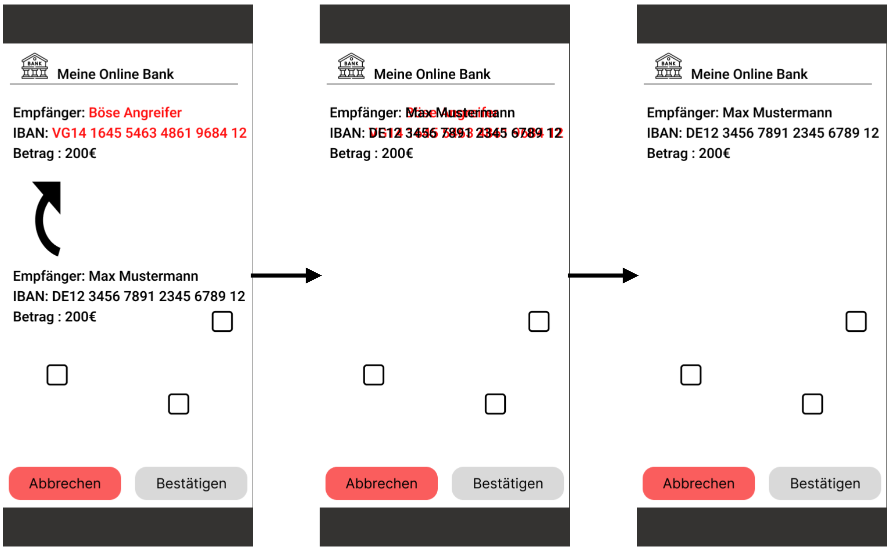

Meine Online Bank
Empfänger: Böse Angreifer
IBAN: VG14 1645 5463 4861 9684 12
Betrag : 200€
Empfänger: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Betrag : 200€
Abbrechen
Bestätigen
Meine Online Bank
Empfänger: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Betrag : 200€
Abbrechen
Bestätigen
Meine Online Bank
Empfänger: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Betrag : 200€
Abbrechen
Bestätigen

Figure 2: Overlaying the IBAN

2.3 Related works

As of the first quarter of 2024 [62], android held a dominant market share in the global mobile operating system landscape, accounting for approximately 70% of devices and Apple's iOS followed with around 29% as a strong contender. Given the worldwide adoption of these two platforms, the continuous enhancement of their security is imperative.

In the case of overlay attacks, there are some cautions about Android accessibility features that create an opportunity for attackers to abuse these functions. In the current Android OS, there is a significant tension between the high usability and the severe security threats posed by overlays. Without effective countermeasures, attackers can exploit overlays to fully compromise and control the UI feedback loop of Android devices [81] Google introduced methods starting from Android 12, that allow developers to hide overlay windows for specific activities. This method has been highlighted as a powerful and effective approach to mitigate overlay-based attacks on Android devices [21, 60]. However, while this feature adds a robust layer of protection, it is essential to complement it with other security practices.

In 2017, Fratantonio et al. [20] published the instructions for implementing an overlay attack on Android named Cloak and Dagger. The research uncovered several design flaws that enable an Android app with two allowed permissions to carry out severe and covert attacks. The first permission [1] allows an app to create windows that appear on top of all other apps, often used for overlays that remain visible regardless of the foreground app [12]. The second permission [2] allows an app to act as an accessibility service, observing and interacting with other applications on behalf of users [11]. The experiments of Fratantonio et al. [20] demonstrated that it is straightforward to publish an app on the Play Store and that context-aware clickjacking, which is a transparent form of overlay, silent installation of a God-mode app, and keystroke inference attacks are both feasible and discreet. Furthermore, the study suggests improvements for Google to resolve these issues.

Yan et al. [79] study addresses this issue by exploring the possibility of detecting overlay-based malicious apps at the app market level. It includes a comparative analysis of the overlay behavior between benign and malicious apps. Based on the insights gained from this study, the "OverlayChecker" system was designed and deployed to quickly and automatically detect overlay-based malicious apps.

Another recent research in 2024 has identified several vulnerabilities and proposed various countermeasures against overlay attacks. A study by Zhou et al. [81] conducted a systematic examination of unprotected windows in Android system apps, leading to the development of another tool also named "OverlayChecker" that is designed to identify and address these vulnerabilities. Their findings revealed 49 vulnerable system app windows across multiple Android versions, prompting significant security updates from major mobile vendors like Google and Samsung.

Creating overlays on Apple iOS is not feasible, as the only overlays possible are notifications, which appear only at the top part of the screen. However, in 2023, Kaspersky detected a zero-day vulnerability in the Apple native messaging app "iMessage" that could potentially give complete control of an iPhone to a hacker. Such vulnerabilities could serve as entry points for overlay attacks or other types of cyberattacks [28].

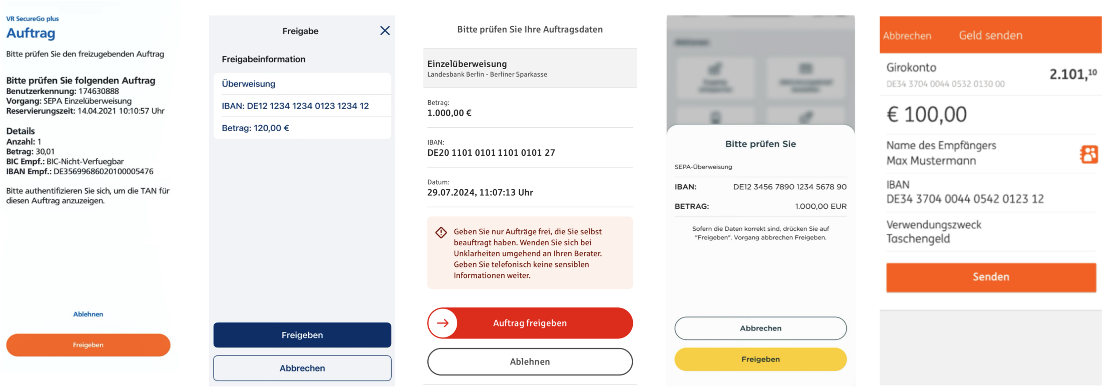
During the research on counter methods and solutions for overlay attacks, it became evident that there is a lack of approaches based on human-centered design and leveraging user cognitive abilities. While technical solutions are crucial, educating users and designing interfaces that enhance security awareness is equally important. Training users to recognise suspicious overlays and employing design elements that highlight legitimate interface components can significantly improve the detection rates of overlay attacks. Ensuring that users are well-informed about the permissions requested by applications and guiding them to enable necessary security settings can also reduce the risk of falling victim to overlay attacks. Therefore, the subject of this thesis and the proposed method represent a new idea that needs to be discovered and tested on users for its usability and suitability to reveal an overlay attack, which is the focus of this research.

The discussed method in this thesis suggests a new design and functionality for TAN

[1] SYSTEM_ALERT_WINDOW

[2] BIND_ACCESSIBILITY_SERVICE

confirmation and as another example of related works in this field, we can observe examples of five German banks Push TAN confirmation method shown in figure 3. These banks are Sparkasse[3], ING[4], Comerzbank[5], Volksbank[6] and Postbank[7]. A white background screen with a simple design is the common background for all of these applications, along with the amount of money and the receiver's IBAN, and three of them also mention the name of the receiver and Date and time of the requested transfer. In a successful overlay attack on these applications the hackers will have to replace the destination IBAN with their own and reproducer the IBAN entered Originally from the user on the white background which would be impossible for the user to detect in such an environment.



VR SecureGo plus

Auftrag

Bitte prüfen Sie den freizugebenden Auftrag

Bitte prüfen Sie folgenden Auftrag
Benutzerkennung: 174630888
Vorgang: SEPA Einzelüberweisung
Reservierungszeit: 14.04.2021 10:10:57 Uhr

Details
Anzahl: 1
Betrag: 30,01
BIC Empf.: BIC-Nicht-Verfuegbar
IBAN Empf.: DE35699686020100005476

Bitte authentifizieren Sie sich, um die TAN für diesen Auftrag anzuzeigen.

Freigabe

Freigabeinformation

Überweisung
IBAN: DE12 1234 1234 0123 1234 12
Betrag: 120,00 €

Bitte prüfen Sie Ihre Auftragsdaten

Einzelüberweisung
Landesbank Berlin - Berliner Sparkasse

Betrag:
1.000,00 €

IBAN:
DE20 1101 0101 1101 0101 27

Datum:
29.07.2024, 11:07:13 Uhr

Geben Sie nur Aufträge frei, die Sie selbst beauftragt haben. Wenden Sie sich bei Unklarheiten umgehend an Ihren Berater. Geben Sie telefonisch keine sensiblen Informationen weiter.

Bitte prüfen Sie

SEPA-Überweisung

IBAN: DE12 3456 7890 1234 5678 90
BETRAG: 1.000,00 EUR

Sofern die Daten korrekt sind, drücken Sie auf "Freigeben". Vorgang abbrechen Freigeben.

Girokonto 2.101,10
DE34 3704 0044 0532 0130 00

€ 100,00

Name des Empfängers
Max Mustermann

IBAN
DE34 3704 0044 0542 0123 12

Verwendungszweck
Taschengeld

Figure 3: Five German banks Push TAN confirmation method screen

VR Secure Go Plus [1], Postbank BestSign [10], S-pushTAN [61], Commerzbank photoTAN [8], ING Banking to go [25]

[3] https://www.sparkasse.de/pk/produkte/konten-und-karten/finanzen-apps/s-pushtan.html

[4] https://www.ing.de/hilfe/banking-to-go/girokonto-features/

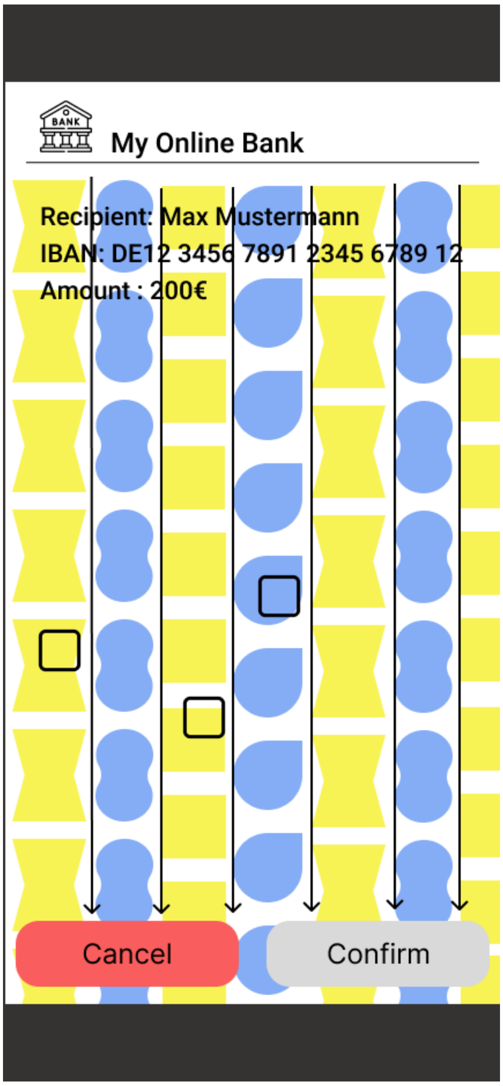
[5] https://www.commerzbank.de/konten-zahlungsverkehr/service/tan-verfahren/phototan/

[6] https://www.vr.de/privatkunden/unsere-produkte/was-ist-ein-girokonto/vr-securego-plus-app.html

[7] https://www.postbank.de/privatkunden/services/online-banking/bestsign.html

2.4 Description of the proposed method

The new method proposes a background, shown in figure 4. It consists of four columns of different shapes that repeat from top to bottom of the screen. Each of the four shapes has a different size and repeats at regular intervals. After the fourth column, the shapes continue to repeat in the same pattern until the entire smartphone screen is filled. This design uses only two colors for the shapes and a background color. The color of each shape remains consistent. For each transaction, the app will generate a random background, making it difficult for any attacker to recreate it in a matter of seconds. If a user notices any irregularity in the coloring, spacing between shapes, or changes in the shapes used, it would be an obvious sign of an overlay on the screen. Additionally, there are three checkboxes randomly placed in different parts of the screen. Users can cancel the transaction with the "cancel" button, and they must check all three boxes to confirm the transaction.



Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€

Figure 4: New mobile transaction confirmation

3 Research methodology

3.1 Research design

This study employs a mixed-methods research design to investigate participants' online banking behaviors, evaluate an interactive prototype for detecting overlay attacks, and assess user satisfaction and usability through the System Usability Scale (SUS). In order to gather insight about the usability and suitability of this confirmation method in practice, and to comprehensively explore the objectives and research questions, the study utilizes both qualitative and quantitative methods. The quantitative component involves collecting and analyzing survey responses and success rates in identifying transactions, while the qualitative component involves gathering open-ended feedback on the functionality and usability. This mixed-methods approach integrates both quantitative and qualitative data collection and analysis to provide a comprehensive understanding of the research problem. This approach is supported by Creswell and Plano Clark's [9] framework for designing and conducting mixed-methods research.

This research assumes varying levels of difficulty for participants in recognizing different overlays. Slides with a partially different background are presumed to be the easiest to identify, while partial discoloration of a shape is considered the most difficult. The other two overlay types are regarded as having a normal level of difficulty. One of the objectives of this study, as stated in the Introduction, is to test these assumptions by analyzing the results. These assumptions are based on the extent of the abnormality they create in each transaction screen.

3.2 Data collection methods

According to A. Tashakkori and C. Teddlie [66], there are six primary methods of data collection: questionnaires, interviews, focus groups, tests, observations, and secondary data (e.g., personal and official documents, physical data, archived research data).

In this study, we employ three of these methods: questionnaires, tests, and observations. The questionnaires gather both quantitative and qualitative data on participants' online banking habits and their experiences with the prototype. The test method involves participants interacting with a prototype to assess their ability to identify overlay attacks, thereby providing performance metrics. Observations are utilized by recording participants' comments during the interaction with the prototype, offering additional qualitative insights into the usability and functionality of the system. This section outlines the methods used to gather data throughout the study. The data collection process consisted of three main components: an initial questionnaire, interaction with an interactive prototype, and a post-interaction questionnaire.

The rationale for using two questionnaires in this study arises from the need to address two distinct areas of inquiry while maintaining participant engagement. The initial and post-interaction questionnaires were carefully designed to prevent fatigue and boredom that could result from presenting a lengthy set of questions either before or after the prototype interaction. Additionally, it was essential that the System Usability Scale (SUS) questions, along with other design and usability-related inquiries, be administered immediately following the prototype interaction. This timing ensures that participants can accurately recall their experiences, thereby enhancing the reliability and validity of their responses.

3.2.1 Initial questionnaire

Participants complete an initial questionnaire to gather data on their online banking habits and basic knowledge of online banking. This questionnaire includes both multiple-choice and open-ended questions and also inquires about any experiences participants may have had with cyber attacks. The questionnaire is designed to address both groups: those with online banking experience and those without it. The complete initial questionnaire is provided in A.6 and can be accessed online [8].

3.2.2 Interactive prototype

Participants interact with a simulation of a TAN confirmation App. This interactive prototype presents them with both safe and hacked transaction scenarios. Participants must decide which transactions are safe based on the information provided in a preliminary instructional video. All slides are provided in A.8, A.9, A.10, A.11 and can be accessed online [9].

3.2.3 Post-Interaction questionnaire

After using the interactive prototype participants complete a second questionnaire. This questionnaire includes the System Usability Scale (SUS), a standardised questions for evaluations the system's usability. Additionally, the post-interaction questionnaire contains questions about their experience with the prototype and open-ended questions for providing feedback on the system. The complete post-interaction questionnaire is provided in A.7 and can be accessed online [10].

3.3 Development of the interactive prototype

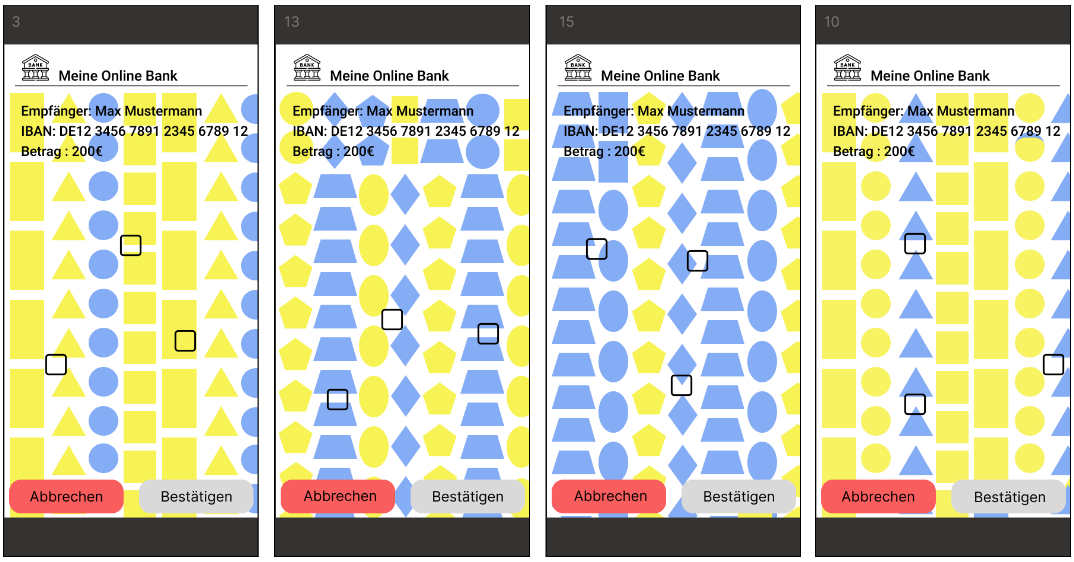
In the beginning phases of this study, it was clear that the creation of a prototype without considering the design looks is not fair to this method. Since the colors and shapes play a very deciding role in this system and this was the first time that this method was being tested with non-IT experts, this research tried to consider all the aspects related to the design of this interactive prototype which could influence the effectiveness and usability of this system. For designing and implementation of this prototype, the Figma [11] platform was used which is widely used by user interface and user experience professionals [19]. Figure 5 presents four examples of final slides, the design of which will be explained in this section of the thesis.

[8] https://forms.gle/2EMfA81mGHc6vdq57

[9] https://tinyurl.com/b4ds5vtr

[10] https://forms.gle/F8ajfdyWTFQsqDwE9

[11] https://www.figma.com/de-de/



Empfänger: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Betrag : 200€

Empfänger: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Betrag : 200€

Empfänger: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Betrag : 200€

Empfänger: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Betrag : 200€

Figure 5: Four different examples of final slides

3.3.1 Criteria for Choosing Colors

In this method the black text is mainly on shapes with two different colors and also partly on a white background. To be able to also create a usable screen that doesn't negatively influence the main goal of transaction confirmation. Different aspects were considered.

The first deciding point in choosing the colors was considering the population with color blindness or low vision. WCAG 2.0 [76] provides three levels of conformance based on luminance contrast: A, AA, and AAA. Each level indicates a different degree of accessibility compliance, with Level AA being the mid-range level that addresses the most common and impactful barriers for users with disabilities. The visual presentation of text and images of text at the AAA level must have a contrast ratio of at least 7:1.

Studies [80] also show that increasing luminance contrast significantly improves visual processing efficiency. For example, response times and the number of fixations required to identify items decrease as luminance contrast increases. Moreover, multiple researches [76, 47, 32] showing that people with normal vision can read at the same speed with high luminance contrast or high color contrast, but people with color blindness and low vision can read faster and better with higher luminance contrast.

After considering different color pallets [46] shown in figure 6, and creating different prototypes with such colors, the problem with lots of these pairs was that black text on two different good recognisable colors, was yet undetectable. This created a clue that the two colors chosen for a prototype should first serve as a proper background for a black text by having high luminance contrast against black color and at the same time be perfectly recognisable from each other to help all users to detect any changes in shapes while confirming a transaction.

An experiment conducted by Ming et al. [42] demonstrated that participants achieved

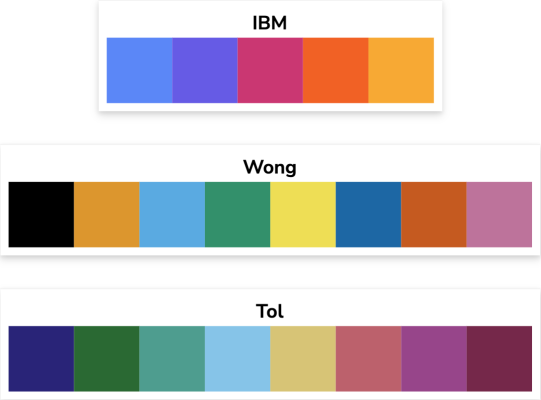
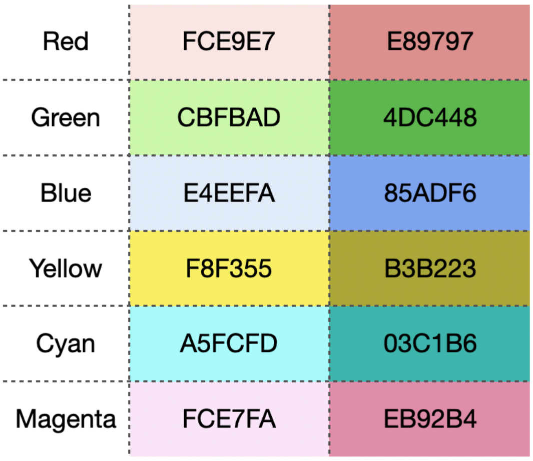


Figure 6: Three different colorblind friendly pallets [40]

their highest cognitive performance when processing and understanding icons with a luminance contrast ratio of 18:1. This finding provided important insights for developing specific color schemes for the prototype. To tackle the problem of reading the black text on different colors, the AAA standard from WCAG was far more moved up to around 9:1 ratio for creating bright colors, and another group of very bright color with approximately 18:1 ratio. These colors were created manually, using the WebAIM contrast checker [74]. For selecting the color, three main colors which are red, green, and blue as primary colors, and secondary colors created from the primary colors, magenta, cyan and yellow were chosen. Furthermore, each color was brought to a 9:1 and 18:1 contrast ratio creating twelve colors in two groups named bright and very bright colors, demonstrated in figure 7, guaranteeing good visibility of black text on these colors.

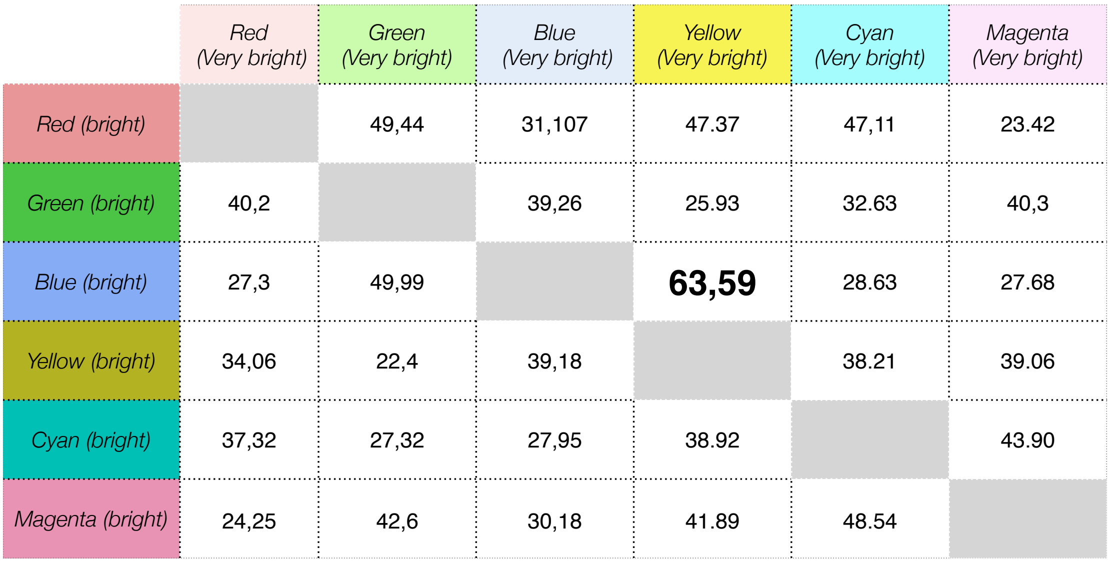


Red	FCE9E7	E89797
Green	CBFBAD	4DC448
Blue	E4EEFA	85ADF6
Yellow	F8F355	B3B223
Cyan	A5FCFD	03C1B6
Magenta	FCE7FA	EB92B4

Figure 7: Candidate colors with two different luminance contrast against black color created manually using WebAIM contrast checker [74] divided in two groups of bright and very bright.

The second problem was choosing the best pair of a bright and a very bright color that are also very distinguishable from each other. CIEDE2000 is a color difference formula developed by the International Commission on Illumination (CIE) to quantify the perceived difference between two colors [57]. It is an advanced metric used to quantify how different two colors appear to the human eye. It's a mathematical model designed to approximate human color perception more accurately than previous color difference formulas. The CIEDE2000 formula was developed through extensive research and validated by various studies. Luo et al. [36] demonstrated that CIEDE2000 had better performance in predicting perceived color differences compared to CIELAB and other formulas, particularly in applications like textiles, coatings, and digital imaging. For individuals with normal vision, CIEDE2000 offers significant improvements in distinguishing colors [37].

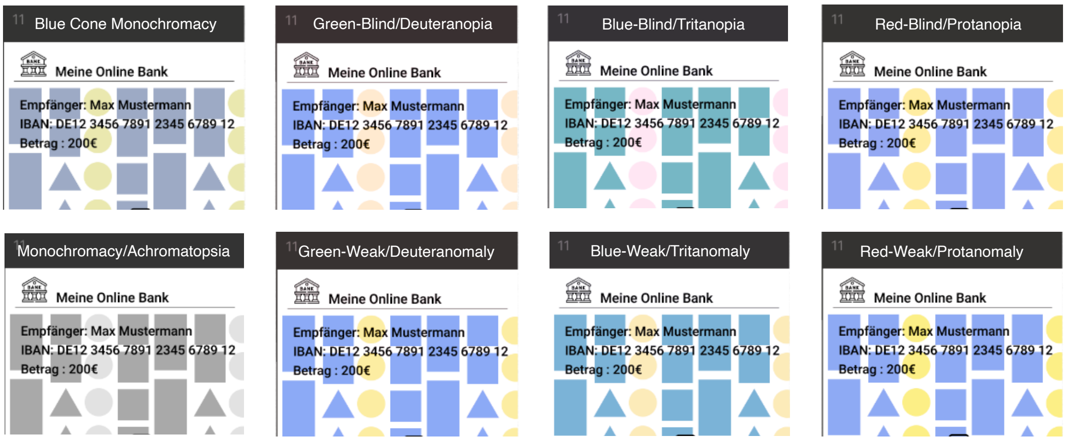
In the next step, 30 different none identical pairs of colors were extracted from the 12 colors in the first step and the CIEDE2000 difference was calculated for all these pairs in python using the colormath library [67] as demonstrated in figure 8. The pair of bright blue with the HEX code 85ADF6 and very bright Yellow with the HEX code F8F355 have the highest difference between all other pairs. In the end, this color pair was tested using multiple colorblind simulation tools[12] [13] [14] to ensure its visibility for the colorblind population. As planned, the high luminance contrast between the two colors guarantees their distinguishability. Even if users cannot perceive the colors themselves, the difference in brightness will still allow them to distinguish between the two. Figure 9 visualizes how the chosen colors are perceived by the human eye across different types of color blindness, demonstrating their distinguishability under these conditions created with colblindor[15] tool.



	Red (Very bright)	Green (Very bright)	Blue (Very bright)	Yellow (Very bright)	Cyan (Very bright)	Magenta (Very bright)
Red (bright)		49,44	31,107	47.37	47,11	23.42
Green (bright)	40,2		39,26	25.93	32.63	40,3
Blue (bright)	27,3	49,99		**63,59**	28.63	27.68
Yellow (bright)	34,06	22,4	39,18		38.21	39.06
Cyan (bright)	37,32	27,32	27,95	38.92		43.90
Magenta (bright)	24,25	42,6	30,18	41.89	48.54	

Figure 8: Thirty color pairs and their differences in human eye perception measured by CIEDE2000 formula[57] and created from 12 colors shown in figure 7 using colormath library[67] in python.

[12] https://pilestone.com/pages/color-blindness-simulator-1.
[13] https://www.farbsehschwaeche.de/en/color-blindness-simulator.
[14] https://www.color-blindness.com/coblis-color-blindness-simulator/.
[15] https://www.color-blindness.com/coblis-color-blindness-simulator/.



Empfänger: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Betrag : 200€

Figure 9: The perception of the selected colors as seen by individuals with various types of color blindness.

3.3.2 Criteria for choosing shapes

The human brain is highly adept at recognising patterns and shapes, and research indicates that shapes with fewer edges and greater symmetry are more easily and quickly detected. Symmetry is a fundamental feature that our visual system uses to parse and interpret the complex visual world. Symmetrical shapes are processed more efficiently due to their predictability and reduced cognitive load. Symmetry allows the brain to anticipate and complete patterns, making recognition faster and more accurate. This aligns with the brain's preference for simplicity and order, facilitating quicker and more efficient shape detection [68].

Moreover, shapes with fewer edges are generally easier for the brain to process. Each edge in a shape represents a point where the brain must decide on continuing the shape's outline, increasing cognitive load. Research shows that shapes with fewer edges are associated with lower visual processing thresholds, requiring less mental effort to recognise. This is because fewer edges result in simpler geometric structures that the brain can easily and quickly interpret, enhancing shape recognition efficiency [68].

Additionally, combining fewer edges with symmetry results in optimal shape recognition. Studies indicate that the brain's parallel processing capabilities are more effective when dealing with symmetrical shapes that also have fewer edges. These shapes provide clear, concise information that the visual system can process with minimal effort. Empirical studies consistently show faster reaction times and higher accuracy rates in recognising these shapes compared to more complex, asymmetrical ones, further evidencing the efficiency of this combination [68, 51].

Based on these findings, the shapes used in the interactive prototype were divided into two groups, as shown in figure 10. The first group, considered the easier group of shapes, includes those with more symmetrical axes and fewer edges. The second group consists of shapes with fewer symmetrical axes and more edges.

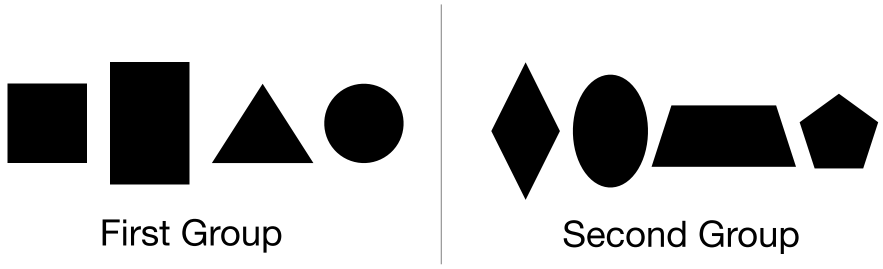


Figure 10: Two shape groups used in the prototype

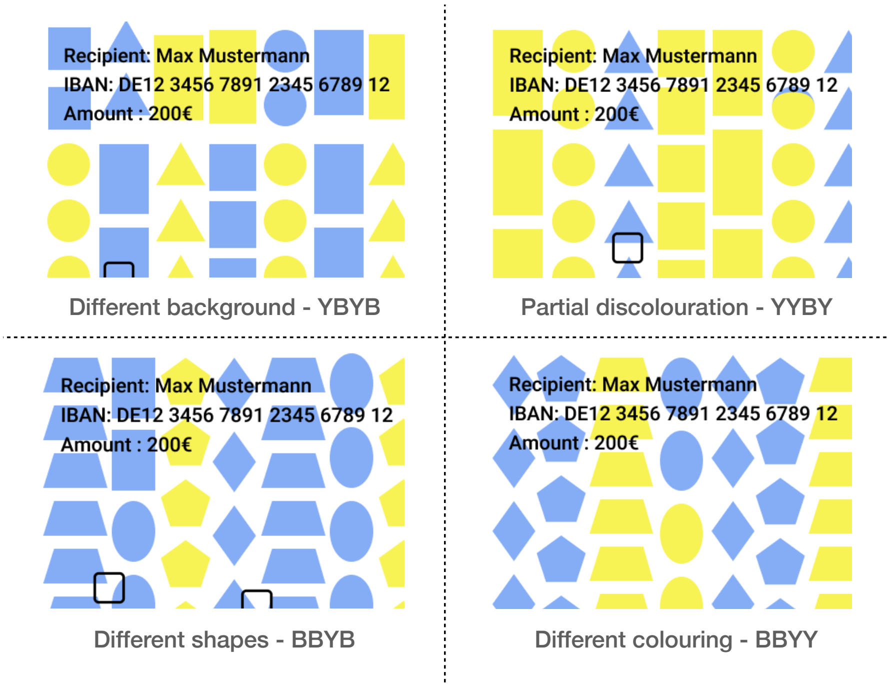
3.3.3 Attack scenarios and their difficulty

The prototype for this study consists of sixteen slides: eight representing safe transactions and eight indicating attacked transactions. To ensure the prototype remains unbiased, the slides are arranged with randomisation. This random ordering was created using Python's random library, selecting from ten different samples with specific criteria: There are no more than two consecutive slides of either safe or attacked transactions, and there are no consecutive slides featuring the same type of attack to employ counterbalancing and minimise order effect and potential biases.

As recommended by the creators of this method, the prototype includes four distinct indicators of overlay attacks: a background that significantly differs from most of the screen, discoloration of some shapes, the use of different shapes within the same column, and partial discoloration of a shape, as illustrated in figure 11. Furthermore, to reduce potential biases, the order of colors in the attack and safe slides was standardised. The same types of attacks across both shape groups follow identical color sequences: YBYB, BBYY, BBYB, and YYBY, where "Y" stands for yellow and "B" stands for blue, moreover each type of attack has the same coloring order for both shape groups as shown in an example in figure 11.

Additionally, the slide numbering, displayed in the left corner of each slide, follows this sequence: Slides 1 to 4 represent safe transactions, and slides 9 to 12 are attacked transactions, all created using group shape one. Slides 5 to 8 also represent safe transactions, while slides 13 to 16 are attacked transactions, created with group shape two.

This research assumes varying levels of difficulty for participants in recognizing different overlays. Slides with a partially different background are presumed to be the easiest to identify, while partial discoloration of a shape is considered the most difficult. The other two overlay types are regarded as having a normal level of difficulty. One of the objectives of this study, as stated in the Introduction, is to test these assumptions by analyzing the results. These assumptions are based on the extent of the abnormality they create in each transaction screen.



Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€

Different background - YBYB

Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€

Partial discolouration - YYBY

Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€

Different shapes - BBYB

Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€

Different colouring - BBYY

Figure 11: Four attack scenarios and color orders used in the prototype

3.3.4 Development of video tutorial

The initial idea for introducing this method to participants was to create an interactive app that would display notes on the screen explaining what participants should do to confirm or cancel a transaction and to explain the overlay attack with examples. The problem with this method was that it required more involvement from the researcher, and the text on the screens would occupy significant space, potentially covering large parts of the transaction background. Furthermore, it would demand more time for introduction phase, and create a realistic tool which could be used for a finished product

To create a more uniform introduction for everyone, two video tutorials were made in English and German, both containing the same commentary and graphics. Furthermore, each video includes a subtitle function designed to assist individuals with hearing impairments. The concept of an overlay attack is presented to the participants through two different animated movements. The process of confirming a transaction, which involves clicking three checkboxes, is also demonstrated twice in the video but is only commented on once during the first mention. The German language video is 3:47 minutes long, and the English version is 2:40 minutes long. The English comments made in the video are as follows:

First Tone: Suppose you are transferring money to your friend Max using your mobile phone: The transaction details will be displayed in the app for confirmation. You need to check the three boxes shown below and then click the "CONFIRM" button to authorise the transaction. (15 Seconds) (see figure 12)

First Animated action: presenting how a transaction is confirmed and the first presentation of what an overlay attack can look on the transaction. (10 Seconds) (see figure 12)

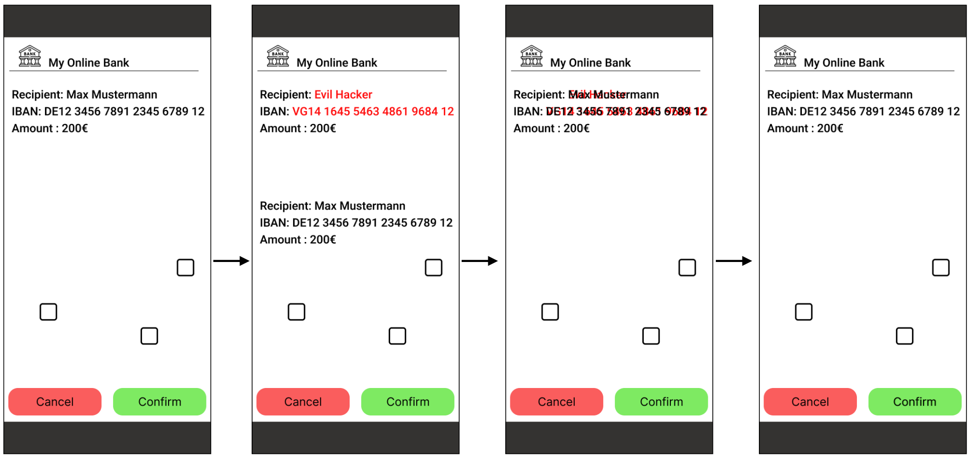

My Online Bank
Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€
Cancel
Confirm
My Online Bank
Recipient: Evil Hacker
IBAN: VG14 1645 5463 4861 9684 12
Amount : 200€
Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€
Cancel
Confirm
My Online Bank
Amount : 200€
Cancel
Confirm
My Online Bank
Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€
Cancel
Confirm

Figure 12: First tone and first animated action in introduction video

Seconds Tone: But be careful! Hackers are after you and want to redirect the money transfer to their own account. They have managed to infect your device and can manipulate the transaction details displayed to you by overlaying the top part of the screen. (14 Seconds) (see figure 13

Second animated action: Presenting again how an overlay can cover the actual transaction data underneath. (8 Seconds) (see figure 13)

Third Tone: While your bank assumes you are confirming the transfer to the hacker, it looks to you like the money is going to your friend Max! (First example slide appears) To help you recognise such attacks, the app has been given a special background that is difficult for hackers to replicate. You should only confirm the transaction if the background pattern is consistent across the entire screen: from top to bottom. As shown in this example, there should be no changes in either shape or color. If you see irregularities in the pattern, you should not confirm the transaction and instead cancel it. This is an example of a legitimate transaction. (36 Seconds) (see figure 13)

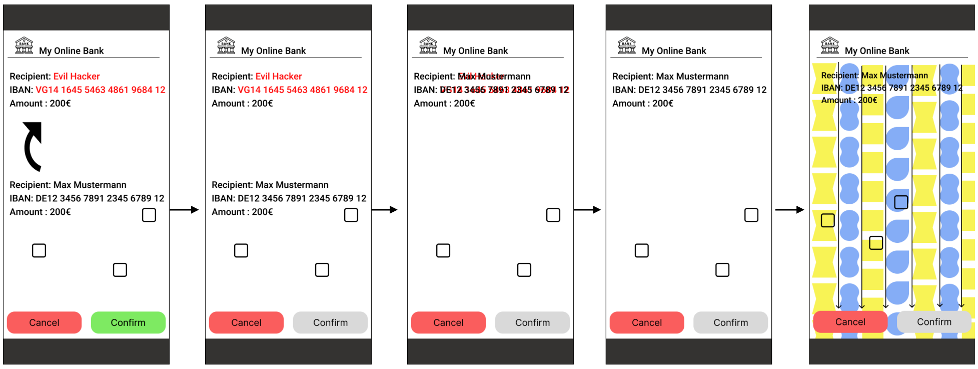

My Online Bank
Recipient: Evil Hacker
IBAN: VG14 1645 5463 4861 9684 12
Amount : 200€
Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€
Cancel
Confirm
My Online Bank
Recipient: Evil Hacker
IBAN: VG14 1645 5463 4861 9684 12
Amount : 200€
Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€
Cancel
Confirm
My Online Bank
Amount : 200€
Cancel
Confirm
My Online Bank
Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€
Cancel
Confirm
My Online Bank
Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€
Cancel
Confirm

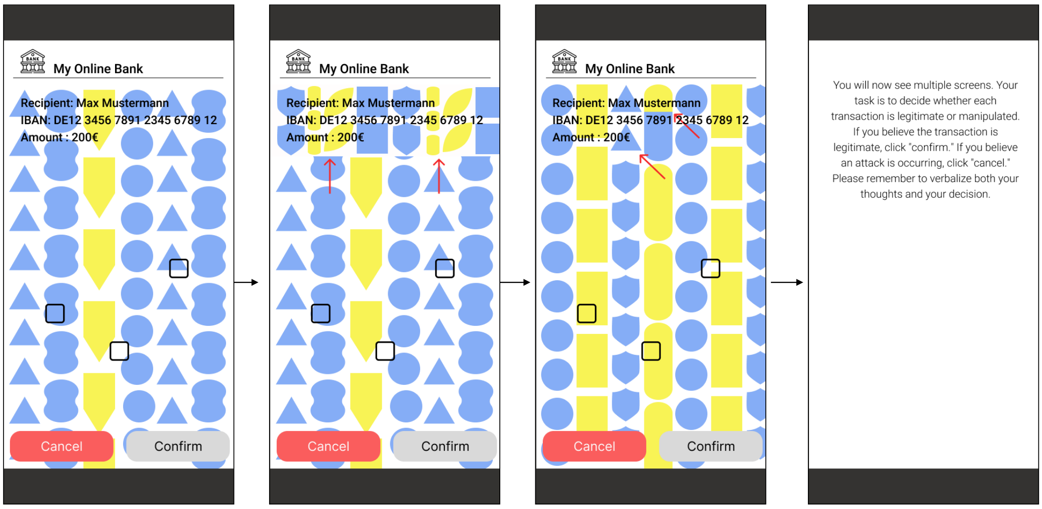
Figure 13: Second tone, and second animated action, and third tone in introduction video

Fourth Tone: (Second example slide appears) Here is another example of a legitimate transaction, where the pattern and colors remain unchanged from top to bottom. (7 Seconds) (see figure 14)

Fifth Tone: (Third example slide appears) This is an example of an attack. The background pattern in the upper part of the screen differs from the one in the lower part. The boundary between the different backgrounds is marked with red arrows. You must cancel the transaction. (13 Seconds) (see figure 14)

Sixth Tone: (Fourth example slide appears) Here is another example of an attack. One oval is a different color from the others, and the shape of the signs has been changed to triangles, while the color remains the same. These differences are marked with arrows. (12 Seconds) (see figure 14)

Seventh Tone: (The last slide with white background and text appears) You will now see multiple screens. Your task is to decide whether each transaction is legitimate or manipulated. If you believe the transaction is legitimate, click "confirm." If you believe an attack is occurring, click "cancel." Please remember to verbalise both your thoughts and your decision. (19 Seconds) (see figure 14)



Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€

Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€

Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€

You will now see multiple screens. Your task is to decide whether each transaction is legitimate or manipulated. If you believe the transaction is legitimate, click "confirm." If you believe an attack is occurring, click "cancel." Please remember to verbalize both your thoughts and your decision.

Figure 14: Fourth, fifth, sixth, and seventh tone in introduction video

3.4 Data analysis

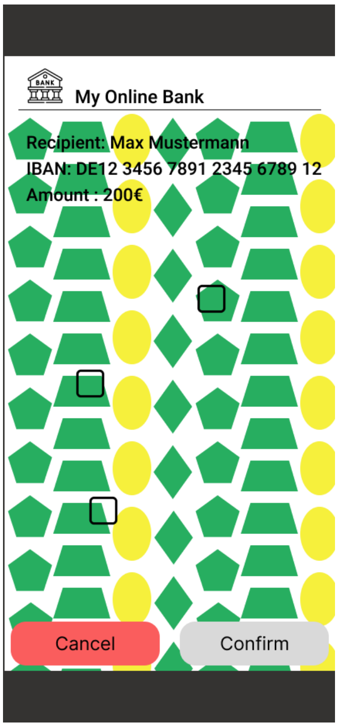
The data collected from the initial and post-interaction questionnaires, along with the results from the interactive prototype, will be analyzed using both quantitative and qualitative methods. Quantitative data from the questionnaires and prototype interactions will be statistically analyzed to identify trends and measure usability through the SUS scores. Qualitative data from open-ended questions and user feedback will be thematically analysed to gain insights into participants' experiences and perceptions [66].

4 User Study

To develop a usable transaction confirmation method that can be adopted by a diverse population of millions or even billions of users, it is crucial to undergo multiple stages of testing and refinement. This user study marks the first evaluation of the proposed confirmation method, aiming to enhance the user interface for transaction confirmations by increasing security and accessibility. This study serves as a foundational step that could pave the way for further research and user studies, ultimately leading to a market-ready product.

4.1 Pilot

Two pilot studies were conducted, each involving one participant, to test different aspects of the study.



Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€

Figure 15: An example of the version tested in first pilot study with different color pairs, using michelson contrast [41] as the measurement for color difference and smaller text font.

The first pilot study was conducted in person and aimed to examine the understandability of the questionnaire and video introduction with a non-native German or English speaker. Additionally, it helped to test the instruments and provided insight into the study's duration. This participant was asked to provide feedback on every detail of the study. The participant required 10 minutes to complete the interaction, and the entire study took approximately 45 minutes, which was deemed lengthy. Furthermore, the participant expressed negative feedback regarding the difficulty of the prototype and slides, as well as the lack of usability in this method. As a result of this pilot, a deeper investigation into the design was considered, and the criteria for choosing colors was changed, which was explained in section 3.3.1. Additionally, the clarity of both questionnaires was reviewed, and some were revised. There were also questions about

The second pilot study was conducted remotely with another non-native German or English speaker. Its primary purpose was to test the process and effectiveness of remote participation, as well as to evaluate the changes made to the questionnaires and prototype. For this participant, it took approximately 22 minutes to complete the study, and there were no comments regarding the clarity of the questions.

4.2 Participant selection

After the pilot studies, a target of 15 to 20 participants was suggested to gather sufficient data for this research. The recruitment process began within circles of friends, family, and college contacts, and was later extended to acquaintances met through work or daily activities. The recruitment process did not specifically target technical experts. The aim was to include a diverse range of participants, representing various backgrounds, education levels, professions, and covering different age groups and genders, without concentrating on any particular field or job category.

None of the participants were paid for their participation in this study. The sole requirement for participation was a willingness to engage in the user study related to testing an online banking system for a minimum of 30 minutes, without further explanation. For remote participants, there was an additional condition: they needed to be able to make a video call on their computer, and they were required to have a smartphone and a stable internet connection.

4.3 Procedure

The study was designed to be conducted either in person or remotely, moderated by a researcher. A remote study instruction file in PDF format (see A.1,A.2) was created for remote participants. This file was sent to participants at the beginning of the video call and was also used by the moderator during in-person study to access necessary links and follow study protocols.

The duration of the study varied among participants, but each was asked to commit at least 30 minutes. In-person studies were conducted one-on-one in various locations and not in a lab setting. Remote studies were also conducted one-on-one using Apple FaceTime [16]and Google Meet[17].

At the start, each participant has read and signed a consent form (see A.3, A.4) which explains sufficient introductory information about the study. Remote participants were instructed to install the Figma [18] app and create an account before the study began. Participants first completed the initial questionnaire and could ask questions

[16] https://apps.apple.com/de/app/facetime/id1110145091

[17] https://meet.google.com/landing?hs=197&pli=1&authuser=0

[18] https://www.figma.com/de-de/

if any were unclear, then watched the introduction video. At the end of the video they were asked if they understood the method and notified that they could watch the video multiple time if needed.

In the next step, participants interacted with the prototype and were asked to explain their choices from the beginning. Those who did not provide explanations were asked to do so again after canceling a transaction. Finally, each participant completed the post-interaction questionnaire and could ask questions if they needed further clarification.

4.4 Data Collection Instruments

The initial and post-interaction questionnaires were created in Google Forms[19], which automatically generates a Google Sheets[20] file for each form. In-person participants answered these questionnaires on a tablet, while remote participants were asked to complete them on their computers. Prototype interactions were recorded with smartphone screen recording for both remote and in-person participants. The results for each participant were later extracted manually by the researcher from these videos and entered into tables for further analysis in Apple Numbers[21].

[19] https://www.google.com/forms/about/

[20] https://www.google.com/sheets/about/

[21] https://www.apple.com/numbers/

5 Results

This chapter presents the findings from the various stages of the study. The results are organised into five main sections: The first section, Participant Demographics, provides detailed information about the individuals who participated in the study. The second section, Initial Questionnaire [22] Findings, summarises the participants' background information, including their online banking habits and prior experiences with cyber attacks. The third section, Prototype Interaction [23] Results, details the participants' performance and success rates in identifying overlay attacks during the interactive prototype sessions. The fourth section, Post-Interaction Questionnaire [24], offers insights into the participants' experiences and perceptions of the prototype, gathered through both quantitative and qualitative responses. Finally, the System Usability Scale (SUS) section evaluates the overall usability of the method, using standardised SUS metrics to quantify user satisfaction and usability. These results collectively provide a comprehensive understanding of the prototype's suitability and user interaction.

5.1 Participant Demographics

A total of 19 participants were interviewed for this study. The average age of the participants was thirty-four years old , ranging from 17 to 54 years old. The age distribution of the participants is shown in figure 16, illustrating the variety of age groups represented in the study.

One participant worked as a consultant in the IT security field, while the others were considered non-experts in this area based on their employment backgrounds. To protect the privacy of the participants, their specific job titles were categorised into different groups, as shown in figure 17. Additionally, the gender distribution of the participants is also visualised in figure 17.

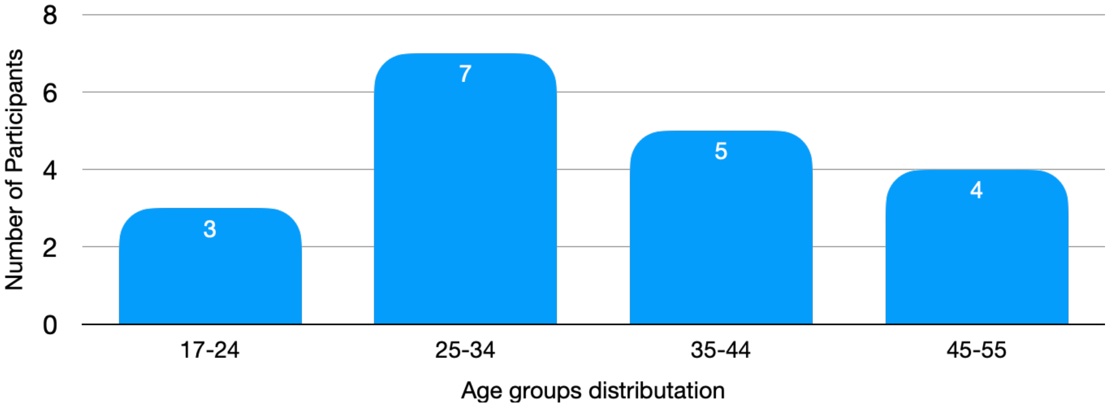


Figure 16: Participants' demographics based on their age groups

[22] https://forms.gle/2EMfA81mGHc6vdq57
[23] https://tinyurl.com/b4ds5vtr
[24] https://forms.gle/F8ajfdyWTFQsqDwE9

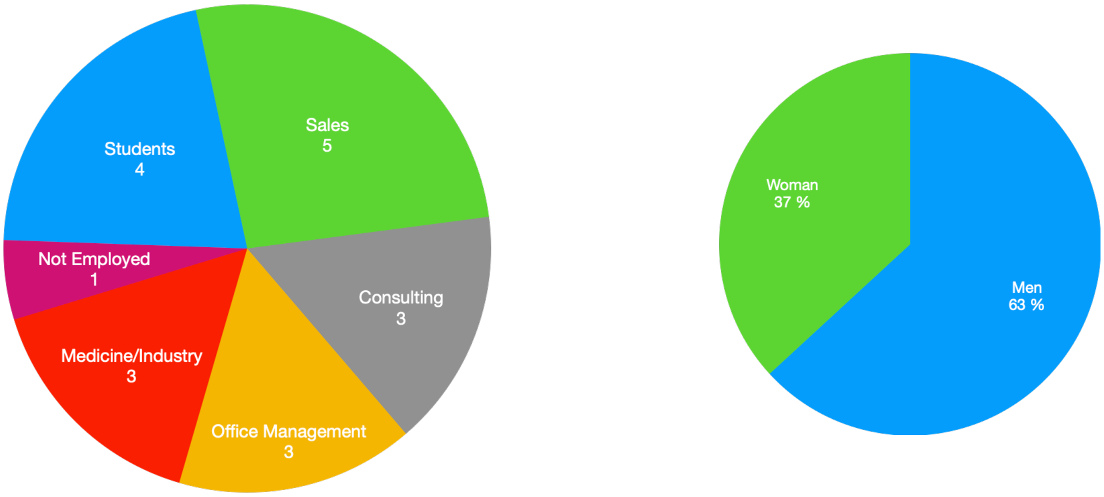


Figure 17: Distribution of participants by occupation and gender groups

5.2 Initial Questionnaire Findings

Among the 19 participants, 3 were not online banking users, while the remaining 16 had over a year of experience using online banking. Of these users, 61% had more than one online banking account, and half reported using online banking at least once a week to transfer money. The details of how frequently these users transfer money via online banking are illustrated in figure 18.

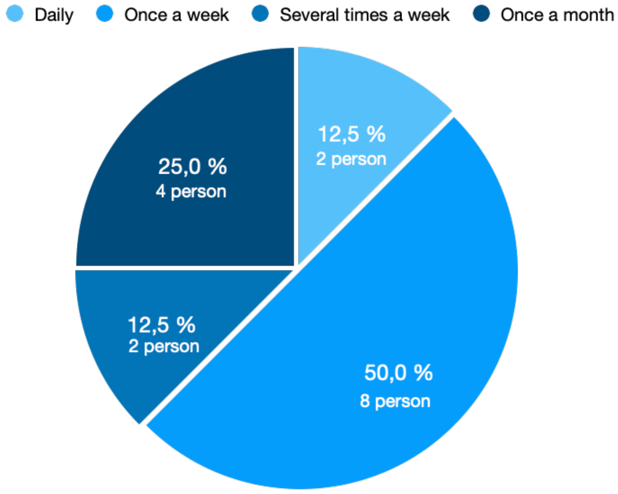


Figure 18: Frequency of Money Transfers via online banking by users

Additionally, all 16 participants used online banking on their smartphones, and 37% used both their PC and smartphone for online banking. Furthermore, 62% of the users transferred money on behalf of third parties, such as their workplace or other organisations. Participants engaged in a variety of online banking activities, including paying bills, transferring money, managing credits and loans, and creating standing orders, as shown in figure 19.

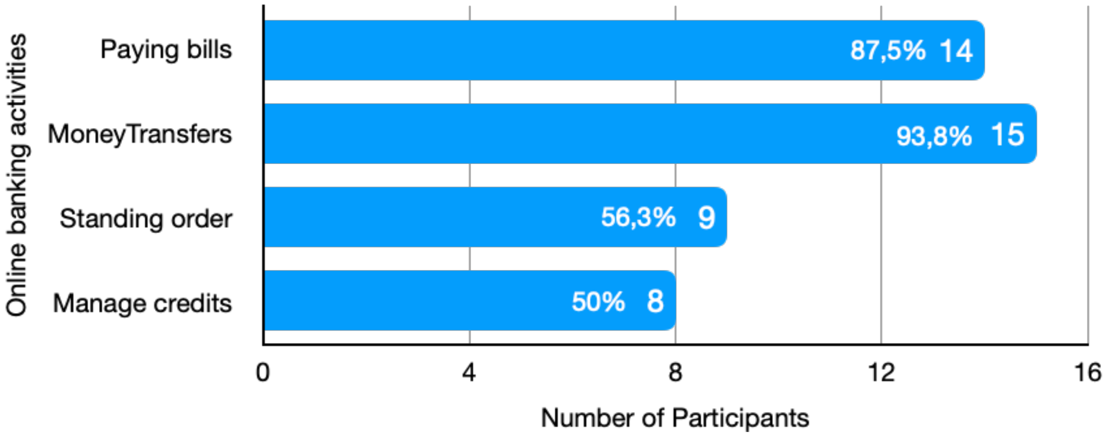


Figure 19: Frequency of money transfers via online banking by users

An interesting finding from this questionnaire was that 25% (4 persons) of the users experienced some form of attack or suspicious activity on their account. In response to the open-ended question, "Have you ever experienced any type of cyber attack on your online banking account or on your smartphone in general?", these 4 participants reported incidents such as attempts to hack their online banking accounts or unauthorised attempts to make purchases in online shops using their accounts. One participant mentioned that money was withdrawn from their account without their knowledge, and another experienced multiple unauthorised transactions. When asked if they reported these incidents to the authorities, all 4 participants responded affirmatively. Figure 20 illustrates how the 16 participants with online banking experience rated the security of their current online banking on a scale from 1 to 10, with 10 being the most secure. Another question aimed to assess the participants' general knowledge about online money transfers. Among the 16 participants, 68% mistakenly believed that, in addition to the IBAN, the recipient's name must also be correct for a transaction to be completed. However, the reality is that an online money transfer will execute as long as the IBAN corresponds to an existing bank account, regardless of the recipient's name. This finding could have important implications for the design of this method, which will be discussed in the following chapter.

Regarding usability and user-friendliness, the results of this questionnaire indicate that 12% of users have helped someone else with their online banking interactions once, and 69% of them have done so multiple times. Users also rated the current complexity of their online banking transaction confirmation procedures, as shown in figure 21. Furthermore, 62% of participants reported that they are unaware of the methods used to secure their online banking transactions. To understand how the users complete a money transfer online the question of "In which step of the transfer do you usually check the correctness of the recipient's IBAN?" were made, which the answers are demonstrated in table 1.

The 3 non-online banking users answered 3 specific questions, revealing that none of them had stopped using online banking; rather, they had never used it. Additionally,

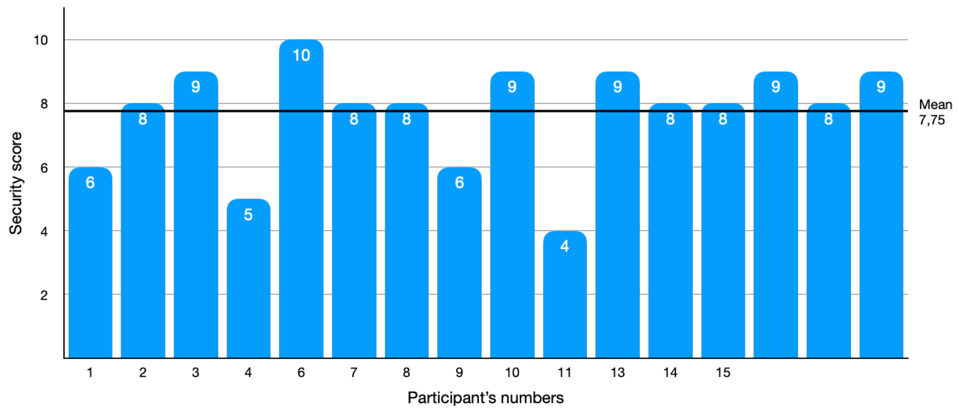


Figure 20: User ratings of online banking security

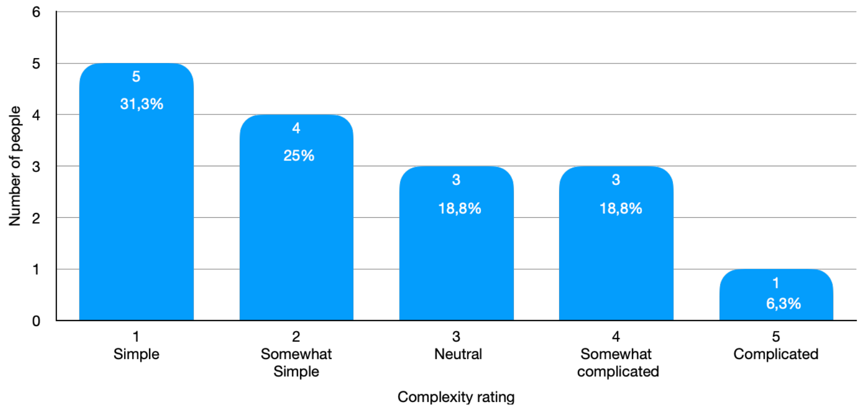


Figure 21: Participants' online banking transaction confirmation procedure complexity

none of them provided a reason for their decision. They were also asked, "How do you rate the security of online banking transactions?" on a scale from 1 to 10, and their responses with a mean of 6.6 out of 10. All the results of this questionnaire have been presented in text, figures, or both in this section. To view the complete questionnaire in German, see (A.6).

Response	**Number of Participants**	**Percent**
Not at all	0	0%
When i fill out the transfer order	10	62.5%
In the summary of the transfer (if my app provides it)	8	50%
In the TAN app before the confirmation	8	50%

Table 1: Steps in online banking where users check the correctness of the IBAN number

5.3 Prototype interaction Results

The specifications of the prototype were thoroughly detailed in the third chapter (see 3.3). In this section, various aspects of the results derived from the participants' interactions will be presented.

Out of 19 participants, 10 correctly identified all 16 transactions, including both safe and attacked ones. The mean success rate was 89%, with a median of 100%. There are 3 participants with a considerably lower detection rate than both the median and the mean in this group. The detailed success rates of all participants are illustrated in figure 22.

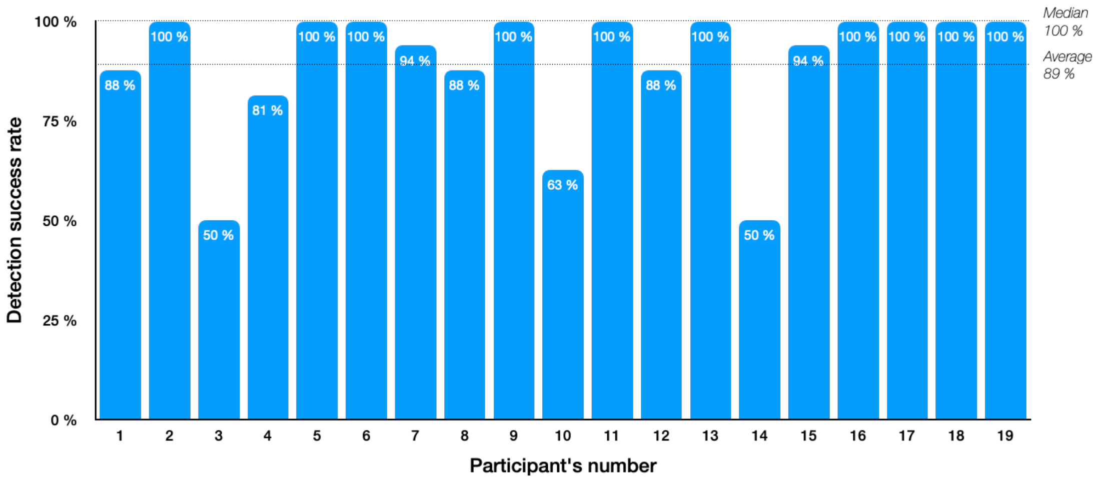


Figure 22: Success rate of each participants in detecting safe and attack scenarios

As mentioned in the third chapter (see 3.3.3), four types of attacks were generated for testing in this prototype. One of the research objectives was to evaluate the assumptions made regarding the difficulty of these attack scenarios. The detection success rate for each attack scenario is illustrated in the figure 23. For each attack type, the success rate is calculated as the mean between 2 different slides, with each slide created using a specific group of shapes.

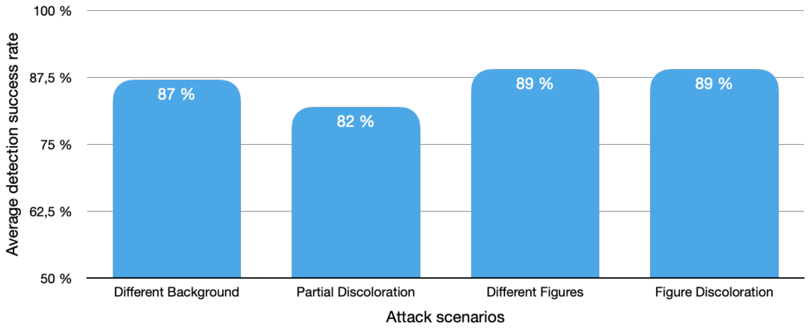


Figure 23: Mean detection success rate of participants by each attack scenarios

The detection success rate for safe transactions was 92%, while attacked transactions were successfully detected by 87%. Another assumption in this research was that the difficulty of detecting shapes varies based on their visual characteristics. An objective of this study was to test this assumption. The figure 24 illustrates the success rate of detecting transaction types, broken down by the different shape groups. The safe transaction slides with assumed easier group shapes have the highest detection rate among all other with 96% and the attacked transactions visualised by second shape group which is assumed to be more difficult to detect, have the lowest detection rate by 85%.

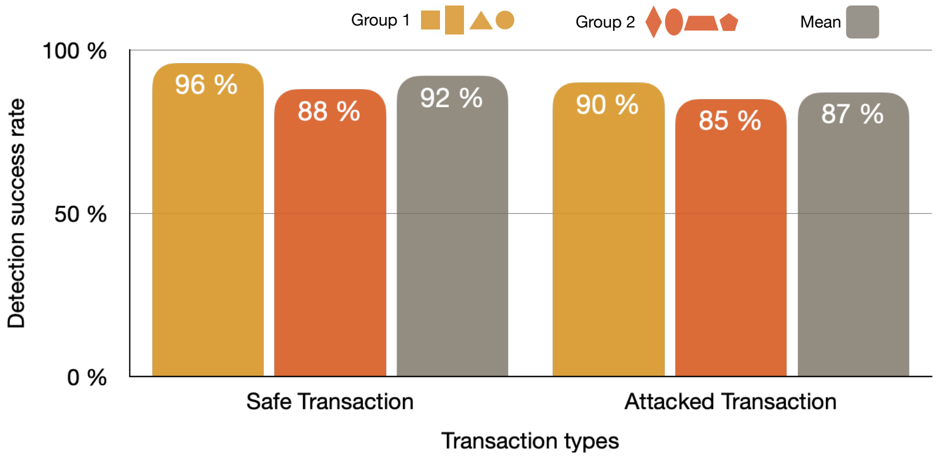


Figure 24: Mean detection success rate in attack and safe transaction and detection success rate categorised by shape groups

The 10 participants with a 100% success rate spent an average of 7.6 seconds per slide to make their decisions, while the other 9 participants had an average decision time of 6.3 seconds. Overall, the average decision time across all slides was approximately 7 seconds and slide 5 which was a safe transaction with group shape 2 had the highest average time spent with 14 seconds. To examine differences among participants based on their age groups, decision times were categorised into four age groups, along with their corresponding success rates, as visualised in figure 25.

To proof the existence of a correlation between age and detection success rate, a scatter plot of age and detection success rate was created, which is visualised in figure 26. Furthermore, The Pearson correlation coefficient and p-value were calculated using the pearsonr function from the SciPy library[25] in Python, which is widely used for scientific computing and provides robust tools for statistical analysis, including correlation calculations [72].

The Pearson correlation coefficient measures the linear relationship between two variables, ranging from -1 as perfect negative correlation to 1 indicating perfect positive correlation [53, 65]. A value close to 0 indicates a weak or no linear relationship between the variables. The Pearson correlation coefficient was calculated as -0.47.

The p-value assesses the statistical significance of the observed correlation. A p-value

[25] https://scipy.org/

below 0.05 generally suggests that the correlation is statistically significant [53, 65]. The p-value = 0.04 was calculated between age and detection success rate of the participants.

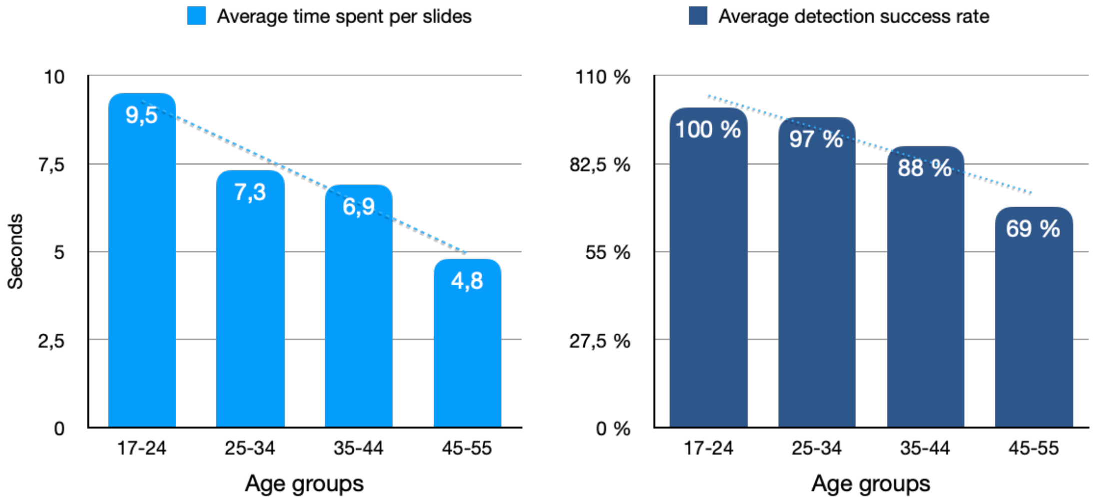


Figure 25: Comparison of decision times and success rates across different age groups

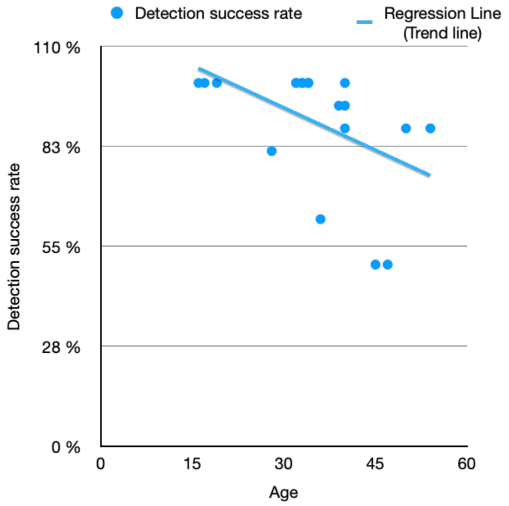


Figure 26: Scatter plot created between participant's Age and their detection success rate

Among 16 tested slides by participants the detection rates are varying between 74% and 100%. There is only one slide with a 100% detection rate, which present a safe transaction, created with shape group 1, which is assumed to be easier for participants to detect. The mean and median of detection per slide was 89%. The detailed detection success rate of each slide is shown in figure 27.

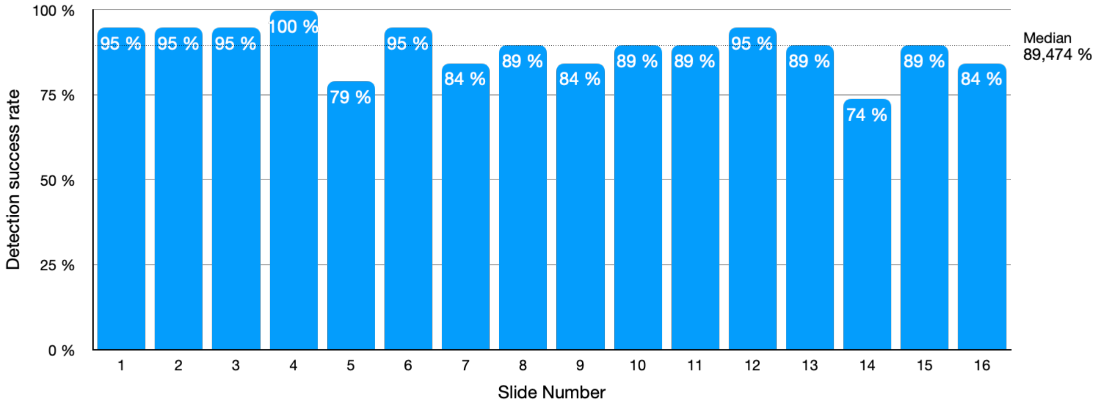


Figure 27: Detection success rate by each slide.

Slides 1-4 represent safe transactions with shape group 1, and slides 5-8 with shape group 2. Slides 9-12 represent attacked transactions with shape group 1, and slides 13-16 with shape group 2.

The quantitative data includes comments made by users during their interaction with the prototype. All participants were asked to explain their choices at the end of the introduction video, and the researcher also requested that they verbalise their reasoning at the beginning of the interaction. Some participants initially forgot to do so and were reminded to articulate their choices when they canceled their first transaction. The most frequently commented topic concerned the shapes with many corners, particularly trapezoids. Three participants felt that one or two slides containing trapezoids featured shapes that were not the same size or that the distance between the trapezoid and neighboring shapes was inconsistent. Slide number 5, which represented a safe transaction and involved group shape 1, was especially problematic, particularly in relation to the trapezoids. Additionally, one participant experienced a similar issue with the polygons on slide 10, which represented an attack. Another design comment was that 2 participants felt the shapes appeared to have a darker blue color when black text was placed on them.

Due to the nature of the task, many participants focused on completing the interaction and were less inclined to verbalise their thoughts. Sometimes, they simply pointed out issues to the researcher rather than describing them, as they seemed to consider the problems self-evident. The youngest age group, comprising three participants who all achieved a 100% success rate, provided clear commentary on their choices.

Another design issue was the habit of checking the checkboxes. Two participants detected an attack but, out of habit, still checked the boxes and confirmed the transaction twice in a row and each time they realised their mistake afterward. Multiple participants asked, upon confirming the first transaction, whether they needed to check all the boxes to ensure that was the correct way to confirm a transaction. Additionally, 2 other participants either did not remember or did not understand the functionality of the checkboxes from the introduction video. The most significant issue with the checkboxes occurred with a participant from the oldest age group, who achieved a 50% success rate. This participant misunderstood the concept entirely, confirming every transaction where at least one checkbox was fully within a shape and canceling the transaction if all 3 checkboxes were partially outside the shapes. Another participant from this age group did not grasp how to confirm a transaction. After the first slide, when the procedure was explained, they confirmed every transaction as safe. To avoid biasing the study, the researcher refrained from further explanation or asking them to re-watch the video.

The final observation in this quantitative section concerned a participant who, after the third slide, began to simply check each column for the required points before deciding on the transaction. He was the only participant who clearly expressed using this method.

5.4 Post-Interaction Questionnaire

In this questionnaire, alongside the SUS questions, participants answered 11 additional questions. The last 4 were demographic questions, with the results of 3 (age, gender, and occupation) discussed earlier in this chapter. The fourth demographic question involved a self-assessment of the participants' about experience in either IT or mobile security. 32% (six participants) responded positively to this question, indicating they consider themselves experienced in this field. To further assess the participants' actual knowledge in security, and as an additional measure of their experience, four technical terms were presented, before the participant's response to the question about their experience in IT or mobile security experience. Participants were asked to indicate which of these terms they could define beyond merely recognising the name. The responses to this question are illustrated in figure 28.

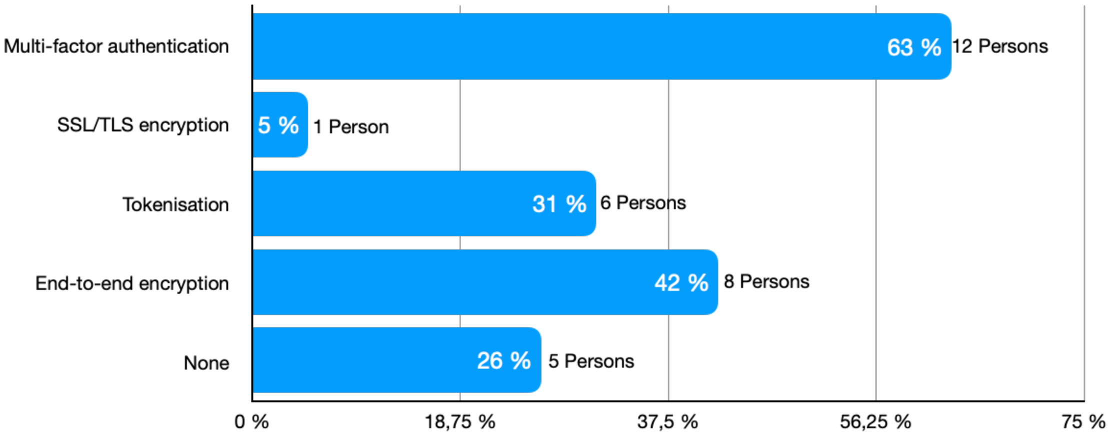


Figure 28: Number of participants who could define selected technical terms in the mobile security field

Participants responded to various questions regarding the design and their experience with the prototype. Additionally, they were asked a background question about whether they had any diagnosed colorblindness, which all of the participants responded with none. One of the design-related questions focused on the shape groups, with 18 participants indicating that they found the first group of shapes (circles, polygons, rectangles, and squares) to be the easiest to analyse.

In addition to the blue-yellow color pair used in this prototype, whose selection process was thoroughly explained in chapter three (see 3.3.1), the next two best color pairs based on the CIEDE2000 metric were chosen for comparison in a question illustrated in Figure 29. 62% of the participants selected the blue-yellow pair, already used in the prototype, as the most distinguishable, while 26% preferred the blue-green pair, and 11% favored the red-green pair.

The final design-related question, which also aimed to evaluate the usability of this method, asked participants whether they had difficulty recognising patterns of shapes

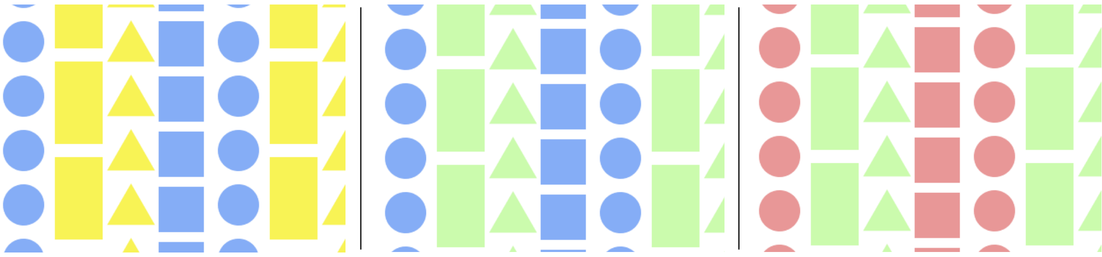


Figure 29: Best three color pairs with the highest CIEDE2000 difference from left to right

that repeated from top to bottom. Participants responded on a scale of 1 to 5, with 5 indicating complete agreement. 11 participants disagreed or completely disagreed that they had difficulty, while 3 responded neutrally. The detailed responses are shown in Figure 30.

Another important question asked participants to evaluate whether they find this method safer than their current online banking method on a scale of 1 to 5. All participants responded to this question, including those who do not use online banking, as they were introduced to the TAN confirmation method primarily used in Germany (see figure 3) through the introductory video. 58% of participants indicated that they believe the new method is safer than their current online banking method, 25% felt that their current methods are safer, and the remaining participants responded neutrally.

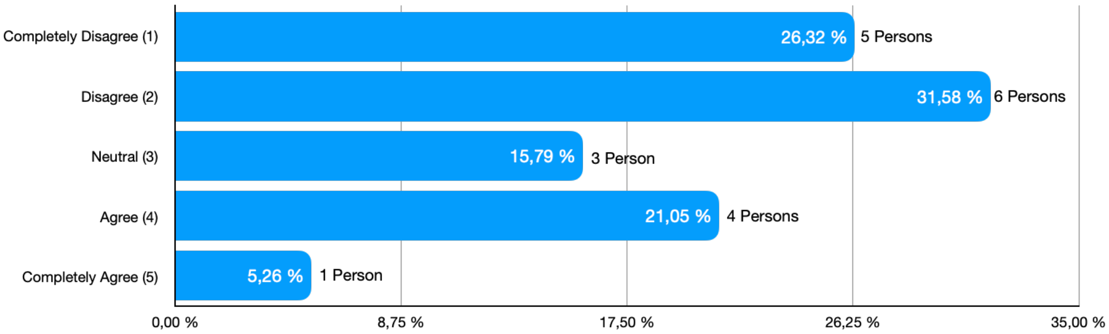


Figure 30: Participants' responses regarding the difficulty of recognising patterns of shapes that repeat from top to bottom

The quantitative section of the post-interaction questionnaire included an open-ended question asking participants for their feedback and critique of the system. Out of 19 participants, 11 had no specific feedback or criticism to offer. One participant stated that they found the system to be normal. However, the remaining comments varied significantly. One participant felt that the system was too time-consuming, while another found it too complex. Additionally, a participant remarked that this method placed too much responsibility on the customer. Another participant suggested displaying only 6 columns on the screen to avoid partial columns on the right side.

One participant in the pilot study criticised the system for interfering with the primary purpose of the application, which is to easily check the IBAN and other information, rather than complicating the process with shapes and colors. The participant who misunderstood the task and confirmed every transaction remarked that they did not believe the system would be widely accepted. Meanwhile, the participant who checked each column individually suggested that using lines instead of shapes might be easier, though they did not fully articulate their thoughts. Another participant, who only accepted transactions where a checkbox was fully inside a shape, commented that the video should have provided more detailed information. Finally, one participant noted that some transactions were faster to detect than others.

5.5 System Usability Scala

The SUS scores were calculated for each participant, with individual scores ranging from 42 to 100. The overall mean SUS score for the prototype was 76, which corresponds to a grade B on the grading scale developed by Lewis et al. [34], and falls within the "good" range according to the adjective rating scale developed by Bangor et al. [2]. The details of the SUS score for each participant are demonstrated in figure 31. The numbering of the participants is in the order they participated during the study. The participants with a 100% detection success rate had a mean SUS score of 73%, while the other 9 participants who did not achieve a perfect detection score had a mean SUS score of 79.

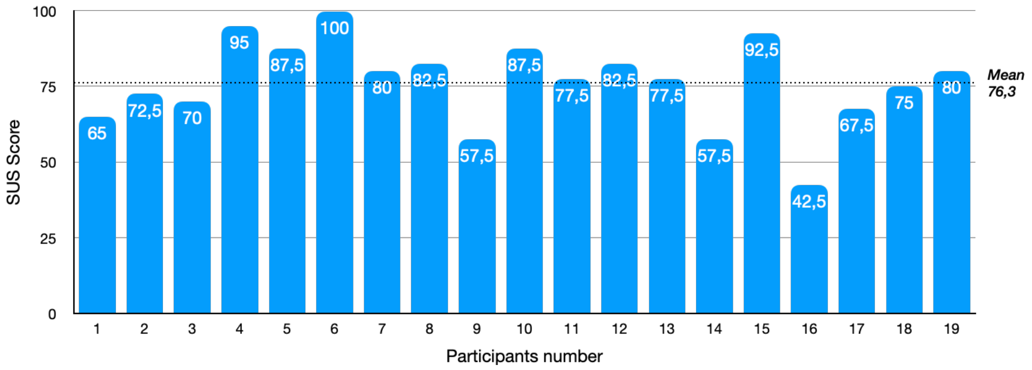


Figure 31: SUS score of each participant

Different age groups were also observed by their SUS points. The highest SUS score belonged to the age group 35-44, with a mean of 89 points, and the lowest SUS score belonged to the 17-24 age group, with 59. Participants aged 35-44 rated the usability of the system with a mean of 88 points, while the largest age group, 25-34, including 7 people, rated the system with a mean SUS score of 79. In the figure 32 the SUS score of different age groups are demonstrated and each bar also show the rating of the score based on Lewis et al.[34] and Bangor et al.[2]. This figure also demonstrates a scatter plot showing the relationship between SUS scores and the age of the participants. To answer the first research question, a scatter plot of SUS score and detection success rate was created, which is visualised in figure 33. Furthermore, The Pearson correlation coefficient and p-value were calculated using the pearsonr function from the SciPy library[26] in Python. The Pearson correlation coefficient between SUS score and detection success rate was reported as 0.083. Furthermore, the of p-value = 0.73 was calculated between these two variables.

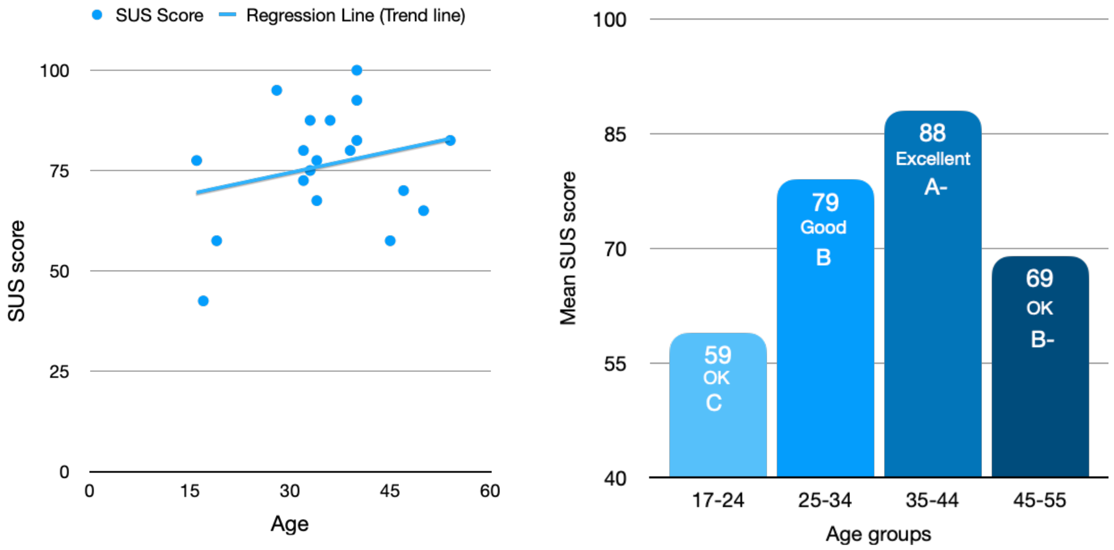


Figure 32: On the left: Scatter plot created between SUS score and participant's Age. On the right: SUS scores by age group, highlighting usability ratings according to Lewis et al. [34] and Bangor et al. [2]

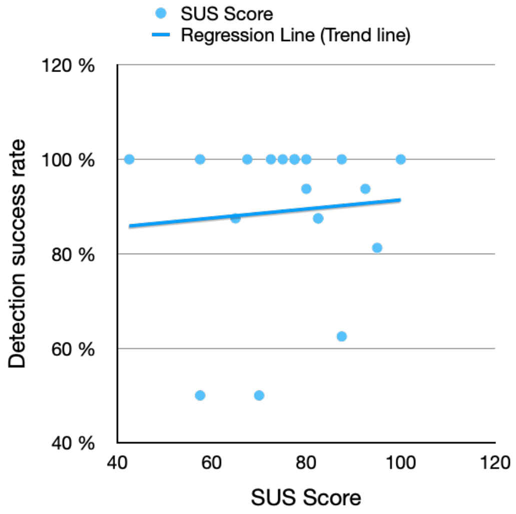


Figure 33: Scatter plot created between SUS score and detection success rate of participants.

[26] https://scipy.org/

6 Discussion

This chapter synthesises the findings of the study, focusing on the key design presumptions, user feedback and suggestions, prototype interaction results, attack hypothesis testing, and the answers to the research questions. It examines the assumptions made during the design process, evaluates the user feedback to suggest improvements, and interprets the user's success in detecting overlay attacks.

6.1 Design presumptions

The selection of colors for the interactive prototype was thoroughly examined in the research methodology (see 3.3.1), with a focus on using the most suitable metrics for human visual perception and the people with color-blindness. As demonstrated in the results section (see section 5.4, figure 29), 62% of participants preferred the color pair used in the prototype as the most distinguishable. This preference establishes a strong foundation for this method in color selection. Additionally, the second most favorable pair, as determined by the metrics, was chosen by 26% of participants, further reinforcing the blue-yellow pair as the clear leader in user-friendliness and distinguishability. As noted in the user study section (see 4.1), the first pilot participant found the visibility of black text on the chosen color pairs problematic. However, none of the later participants reported this issue after selecting the new color pairs using new metrics. This indicates that using higher contrast standards than those recommended by WCAG 2.0 [76] proved to be effective. Additionally, the CIEDE2000 color difference metric identified the most distinguishable pairs among the 30 tested, highlighting its effectiveness in selecting color pairs that are well-suited for this method. This metric can also be utilised to identify additional pairs of colors that are optimally distinguishable.

As outlined in the research methodology (see 3.3.2), two groups of shapes were selected for this prototype based on their symmetry and the number of edges. It is important to note that the detection rate of shapes used in this method must be validated through multiple user studies to determine those that provide the highest usability and security for the system. To maintain the integrity of the user study, it was essential to test specific shapes systematically, establishing a solid foundation for future research in this method. Randomly selecting shapes during the research phase would have undermined the scientific rigor and fairness of the prototype. The fact that participants spend less time on average, analysing safe transaction with group 1 shapes (see 24), and their response in questionnaire, with 94% of participants finding the shapes in group 1 easier to analyse, strongly indicates that more symmetrical shapes with fewer edges are particularly well suited for this method of confirmation. This suggests that in future studies, other factors influencing human visual perception of shapes can be mixed with these two characteristics.

Overall, the results indicate that the assumptions regarding color and shapes made in this study were accurate. These findings suggest that in future implementations, these assumptions can be fully leveraged to enhance key design elements. By focusing on the elements identified as most effective, we can potentially increase the accuracy of detecting overlays by users.

6.2 User Feedback and Suggestions

The qualitative part of the study includes comments of the participant during the interaction with prototype (see section 5.3) and the result of an open ended question in post-interaction questionnaire (see section 5.4) asking for users critic and feedback.

The participants' feedback on the design elements, particularly the trapezoid shape, highlighted their concerns. According to six the participants, the inconsistent size and spacing of trapezoids, especially on slide five (see A.9), may have distracted or confused users during the decision-making process. This issue was also mentioned on slide 10 (see A.10), indicating that the issue with this specific shape was not just incidents but repeated throughout the interaction. The comments on the color brightness, where shapes with black text on them appeared darker for multiple participants despite being the same color as other shapes, further highlight the importance of carefully considering visual design choices in this method.

The behavior observed with the checkboxes points to a usability issue. Some participants confirmed transactions out of habit, even after detecting an attack, suggesting that the checkboxes might encourage automatic interaction without fully considering the consequences. This habitual checking could lead to errors, particularly in real-life situations where users are busy or distracted. This raises the question of whether checkboxes are necessary, as they might contribute to unintentional mistakes, especially when users are not fully focused. Eliminating the use of checkboxes could simplify the interaction process and reduce the chances of these errors. Furthermore, the designer of this method did not clearly explain the purpose of the checkboxes; they are likely intended to be placed randomly on the screen to cover different areas and reduce the chance of a successful overlay attack.

These findings underscore the need for a thoughtful and user-centered approach to design. Addressing the issues with shape consistency, color contrast, and checkbox functionality is crucial for minimising user confusion and preventing errors. Clear communication of design intentions, particularly for elements like checkboxes, is essential to ensure users understand their purpose and can interact with the interface effectively. By refining these aspects, the prototype can be made more intuitive and accessible, ultimately leading to a better user experience.

The post-interaction feedback reveals mixed reactions, highlighting key areas for improvement in the system's design and usability. While eleven participants provided no feedback in this part , others expressed concerns about the complexity, time consumption, and the responsibility this method demands from users. These concerns suggest that the design may need simplification to reduce cognitive load and better align with user expectations.

One notable critique was the suggestion to implement a fixed number of six columns that would adjust to the screen size, as the current layout might make it difficult for users to detect attacks. Ensuring that all columns remain fully visible, without any partial columns extending off the right side of the screen, could improve user interaction and satisfaction. However, a fixed number of columns might make it easier for attackers to create an overlay, potentially compromising the system's security. Allowing the number of columns to vary randomly for each transaction could enhance security while ensuring that no columns extend beyond the visible screen area.

Feedback on the clarity of instructions and the system's purpose from a user who did not understand the task from the introduction video suggests a need for better communication of the system's goals in the video. It also suggests that the study's methodology could include a step before the interaction where participants are asked to explain what they understood from the video and only those who correctly identify all critical points would then proceed to interact with the prototype.

The user feedback and suggestions indicate that there are areas that require refinement to improve usability and user experience. Key issues include the need for use of user friendly visual elements, better communication of the system's purpose in introduction phase, and a simplification of the interaction process to reduce cognitive load. Addressing these concerns through clearer instructions, adaptable background design, the exclusive use of user-friendly shapes with better detection performance, and consideration of user habits with checkboxes, will be critical in creating a more intuitive design and protecting users from overlay attacks.

6.3 Interpretation of prototype interaction results

As shown in results (see section 5.3, figure 22), the detection success rate of the participants had a mean of 89% and a median of 100%. While the mean of 89% is statistically significant, interpreting this as an indicator of performance for an overlay attack detection method requires further context. There are no universally established benchmarks for what represents a high or suitable detection rate when detection is performed by users themselves. Given this lack of benchmarks, it becomes crucial to establish robust standards, especially since anomaly detection often relies on machines with human oversight to refine the results [39, 52]. This method positions users as the final line of defense against overlay attacks. Therefore, it is advisable to aim for a very high success rate, such as 98% or more, to ensure robust protection, particularly in scenarios where other technical detection measures have failed to prevent an overlay attack.

The difference between the mean and median detection rates indicates that while more than half of the participants detected all slides correctly, a group of participants significantly lowered the mean. This may indicate either a lack of full comprehension of the task, a high cognitive load imposed by the method, or variability in visual acuity among the participants. As shown in figure 25, a clear downward trend in detection success rate is observed as participants age increases. The two younger age groups had a mean detection rate of 98%, which is near perfect, while the older groups had a mean rate of 80%. This suggests a need for changes in both the method's design and the introduction process. Notably, two participants from the oldest age group (45-55) had a detection rate of just 50%. Due to the system's design, a participant who marks every slide as safe would achieve a 50% success rate due to the even distribution of safe and attacked transactions. One of these participants followed this pattern, while the other misinterpreted the task and used the positioning of the checkboxes as a deciding factor for confirming the transactions. As presented in the results section (see 5.3), figure 26 illustrates a clear negative trend line between age and detection success rate. The Pearson correlation coefficient of -0.47, and a p-value of 0.04, further confirms a moderate negative correlation between these two variables, indicating that the decline in detection success rate with increasing age is statistically significant. This suggests that the observed downward trend is unlikely to be due to random chance.

This issue could be addressed by adding a step in the methodology where participants explain their understanding of the introduction video. Only those who demonstrate a correct understanding would proceed to interact with the prototype. It is important to note that this prototype was not designed to be the most optimised version of the method, but rather to serve as an unbiased testing tool that reflects varying levels of difficulty. The introduction phase was carefully crafted to simulate a real-life scenario where personalised explanations are not feasible, using a video with minimal necessary information to prevent bias. In a market-ready product, the approach would be significantly different, focusing on maximising user detection rates. As this work now lays the foundation for future design iterations, incorporating user feedback and insights from this research should concentrate on developing a more effective introduction phase and refining design elements, particularly for older age groups, with the goal of achieving a near 100% detection rate.

The mean SUS score of 76.3, as mentioned in section 5.5, falls within the "good" range of usability scores according to Bangor et al. [2], although it is slightly below the 80% benchmark commonly observed in the industry, as highlighted by Lewis et al. [34]. Several factors could have contributed to this SUS score. First, participant feedback (see figure 30) revealed that 26% experienced difficulty recognising shape patterns, with an additional 16% responding neutrally. This suggests that the design choice of using shape patterns significantly impacted the system's usability, as reflected in the lower mean SUS score of 67.5 for this group. These results highlight the importance of carefully selecting design elements, particularly shapes, as they influence task complexity and, consequently, the usability of the method. This observation aligns with findings by Bangor et al. [3], who noted that complex interfaces tend to lower usability ratings. Participants who reported "no difficulty" by recognising shape patterns had a mean SUS score of 82, further emphasizing the relationship between perceived interface difficulty and overall usability.

Secondly, before interacting with the prototype, participants rated their current online banking transaction confirmation method. 25% of users found these methods at least "somewhat complicated" (see figure 21), while around 19% responded neutrally regarding the complexity of their current online banking confirmation procedures. Additionally, participants with online banking experience indicated that 69% of them had helped others with their online banking multiple times (see section 5.2). These observations suggest that a considerable portion of the population may generally perceive online banking and its transaction confirmation procedures as complicated, regardless of the specific methods being used.

As detailed in results (see figure 32) different age groups exhibited varying mean SUS scores, with no strong evidence of a correlation between age and SUS score. However, the lower mean SUS scores observed in the oldest and youngest age groups, compared to the other two groups, could not be definitively explained by this study and warrant further research. One significant factor may be the low-fidelity design of the prototype, particularly for the youngest group. A higher-fidelity design could potentially improve the SUS score by providing a more polished and realistic user experience, and this hypothesis could be explored in future iterations of prototype testing.

From another perspective, despite the youngest group fully understanding the method, achieving a 100% detection rate, and spending the most time per slide on average (see 25), they reported a mean SUS score of 59. This suggests that the cognitive load required by the method may have negatively impacted their perception of usability. Similarly, the oldest group may have also experienced a high cognitive load. With a mean detection success rate of 69% and a mean SUS score of 69, this group faced considerable challenges. As mentioned earlier, two participants from this group struggled to grasp the task from the introduction video, while the other two participants achieved a success rate of 88%.

As previously noted, there are well-established benchmarks for SUS scores that evaluate a system's usability. Based on these benchmarks [34, 2], this method is not yet ready for a market product and requires further development. However, it has achieved a good level of usability, suggesting that with additional effort and refinement, this method could meet or even exceed industry-standard SUS scores.

6.4 Hypothesis Testing on Attack Detection

The criteria for developing different attack scenarios were detailed in the research methodology (see section 3.3.3), with efforts made to ensure that all four types of attack interfaces (see figure 11) appeared as similar as possible by both group shapes. As illustrated in figure 23, the "partial discoloration" of figures was detected less frequently than in other scenarios, supporting the assumption that this type of attack would be the most difficult to identify. The similarity in mean detection rates between attacks involving "different figures" and those involving "figure discoloration" suggests that these two scenarios are equally challenging, as was initially assumed in the methodology. However, based on the overall detection rates for each attack scenario, both of these types might actually be considered easier than expected. For instance, the "different background" attacks, originally presumed to be easy, were detected by 87% of participants, suggesting that this scenario might better align with what could be considered a normal difficulty level.

By closer examination, the data reveals that two participants who did not fully understand the task had a significant impact on the results of this attack scenario. When the data from these two participants are excluded, the detection rates for the "different background" and "discoloration of figures" scenarios rise to 100%, with the different figures scenario close behind at 97%. Meanwhile, partial discoloration remains the most challenging scenario, with an 85% detection rate.

These findings strongly confirm that partial discoloration is indeed the most difficult scenario among the four, while the other three scenarios appear to have similar difficulty levels.

6.5 Answering the research Questions

The first research question aimed to identify the presence of a correlation between users' SUS score and their performance in detecting overlay attacks and identifying safe transactions. As shown in the results sections (see 5.5) the pearson correlation coefficient with 0.083 and p-value of 0.73 were calculated and the visualisation of the SUS score and detection success rate of the participants on a scatter plot (see figure 33) was also demonstrated to have a better overview of the relation between this two rating. The pearson correlation coefficient indicates a very weak positive correlation between SUS scores and detection success rates, meaning that higher usability ratings are only marginally related to better detection performance. Furthermore, the p-value of 0.73 suggests that this relationship is not statistically significant, implying that the observed correlation could have occurred by chance. This lack of statistical significance suggests that SUS scores may not be a reliable predictor of a user's ability to detect overlay attacks.

The second research question is aiming to assess whether the information provided in the introduction video was helpful for understanding the task. As mentioned in the methodology section (see section 3.3.4), the initial idea for the introduction was an interactive tutorial, which was later simplified to a video to reduce complexity and time consumption and create an introduction method with more familiarity for people with different demographic. Of the 21 individuals who watched the video, including 19 study participants and 2 pilot study participants, 2 did not fully grasp the task after viewing the video, both achieving only a 50% detection success rate and not understanding the confirmation method. One other participants, who took part remotely, did not watch the video with sound. This became evident to the researcher when the participant, upon confirming the third slide, was unable to justify their decision when questioned.

As documented in the methodology section (see section 3.3.4), the video intentionally does not emphasize the new function of the three checkboxes, to gather insights on how participants interact with checkboxes and perceive their function. This approach led to some confusion among participants mentioned in results section (see section 5.3), with some of them asking if they should click all three checkboxes before confirming the first slide to be more certain and two other participants either did not recall or failed to notice the functionality of the checkboxes from the introduction video. Given this outcome, if the three checkboxes are retained in this method, the introduction video should explicitly introduce and explain their purpose to avoid any further confusion.

Overall, the video effectively informed 18 out of the 21 participants, helping them to understand the confirmation method and the concept of an overlay attack, which indicate that 85% of the participants were well-informed by the video introduction. For future improvements to the introduction phase by further research, incorporating a short quiz with multiple-choice questions could further enhance the informative value of this method. This quiz would ensure that users fully comprehend the video content by requiring them to correctly answer key points before proceeding to the next step. If this method is eventually adopted for use in online banking, integrating a video with such a quiz in the introduction phase of the application could guarantee an effective information process.

7 Conclusion

This study aimed to evaluate the suitability and usability of a novel transaction confirmation method designed to combat overlay attacks in online banking. Using a mixed-methods approach that combined user interaction with a prototype and post-interaction surveys, we explored the validity of design assumptions, assessed the success of attack hypothesis testing, evaluated the participants' detection rates, gathered participants' feedback, and measured the method's usability using the System Usability Scale (SUS).

The System Usability Scale (SUS) results revealed an overall positive reception, with a mean score of 76.3, placing the method within the "good" usability range based on Bangor et al.[2] and a grade B based on Lewis et al.[34] benchmarks, close to industry standard of 80 score. Furthermore, Participants who found the system challenging had a significantly lower mean SUS score of 67.5, suggesting that the complexity of the design could hinder its practical adoption. The study also uncovered demographic variations in usability perceptions, with younger and older age groups reporting lower SUS scores compared to middle-aged participants. This suggests that the method's design may need to be tailored to accommodate different user demographics more effectively.

By evaluating the detection success rate, we observed a mean success rate of 89% and a median of 100%. Although the majority of participants achieved a perfect detection rate, the study highlighted the challenges faced by the older age groups (35 to 55 years) in detecting both safe and fraudulent transactions. Furthermore, statistical analysis suggests a high probability of a negative correlation between age and detection success rate.

The feedback and suggestions gathered through interaction with the prototype and post-interaction questionnaires provided valuable insights into user preferences and highlighted areas for improvement, such as the potential need for a more intuitive interface and clearer instructional materials.

In conclusion, while the proposed method demonstrates potential as a robust tool against overlay attacks for the majority of users, its current iteration requires enhancements to ensure its suitability and usability for a broad user base.

7.1 Recommendations for Future Research

Future research should prioritise refining the design elements to enhance usability across diverse user groups. This includes not only improving the visual design but also exploring additional features that could further strengthen the system's security and detection rates. In particular, larger-scale studies are recommended to validate the current findings and optimise the method for real-world application. These studies should focus on a broader participant base, particularly individuals aged 35 to 55, with an emphasis on creating a more intuitive design that reduces cognitive load.

Additionally, future studies should investigate factors influencing human visual perception of shapes, especially in relation to aging and vision weaknesses beyond colorblindness, which was considered in this work. The role of checkboxes in the confirmation process also warrants re-evaluation. The observed tendency of users to confirm transactions out of habit suggests that the current implementation may not be optimal.

To further enhance user understanding, future iterations of an introduction video could incorporate an interactive tutorial or a short quiz with multiple-choice questions during the introduction phase. This addition would ensure that participants fully grasp the critical points before interacting with the prototype, potentially leading to higher detection rates.

Finally, as a suggestion for simplifying this method, minimising the size of the special background area could reduce cognitive load and increase user's focus. Additionally, incorporating more randomisation could enhance security. Based on the results from a question in the initial questionnaire (see Section 5.2), 68% of online banking users were unaware that the only necessary information for executing an online money transfer is the correct IBAN and the accuracy of the recipient's name or BIC is not required. In the event of an overlay attack, hackers would need to successfully conceal the IBAN along with other details. To counter this, the method could randomly select 25% to 35% of the screen for each transaction, placing the IBAN within this area using a random visible text size, while only this portion of the screen has the proper background for overlay detection. This approach could increase the difficulty for attackers by introducing variability in the size of the background, font, and positioning of the IBAN, while the smaller special background would reduce the cognitive load for the user by minimising the focus required for detecting an overlay.

References

[1] Atruvia AG. *Screenshot from the VR SecureGo plus App*. Screenshot obtained from the Apple App Store: https://apps.apple.com/de/app/vr-securego-plus/id1535422059. Accessed: Aug. 18, 2024. 2024.

[2] Aaron Bangor, Phil Kortum, and James Miller. "Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale". In: *J. Usability Stud.* 4 (Apr. 2009). Available: https://www.researchgate.net/publication/228368593_Determining_What_Individual_SUS_Scores_Mean_Adding_an_Adjective_Rating_Scale, Accessed: Aug. 18, 2024, pp. 114–123.

[3] Aaron Bangor, Phil Kortum, and T. Philip Miller. "The System Usability Scale (SUS): an Empirical evaluation". In: *International Journal of Human-Computer Interaction* 24 (Aug. 2008), pp. 574–592. DOI: 10.1080/10447310802205776.

[4] Bank Policy Institute. *Multifactor Authentication: Opportunities and Challenges*. Accessed: Aug. 18, 2024. 2023. URL: https://bpi.com/multifactor-authentication-opportunities-and-challenges/.

[5] John Brooke. "SUS: a "quick and dirty" usability scale". In: *Usability evaluation in industry* 189 (1996), pp. 4–7.

[6] Michele Bugliesi et al. "CookiExt: Patching the Browser Against Session Hijacking Attacks". In: *Journal of Computer Security* 23 (Sept. 2015), pp. 1–20. DOI: 10.3233/JCS-150529.

[7] Commerzbank. *TAN Verfahren*. Accessed: Jul. 26, 2023. URL: https://www.commerzbank.de/konten-zahlungsverkehr/service/tan-verfahren/.

[8] Commerzbank AG. *Screenshot from the Commerzbank photoTAN App*. Screenshot obtained from the Apple App Store: https://apps.apple.com/de/app/commerzbank-phototan/id577752083. Accessed: Aug. 18, 2024. 2024.

[9] John W. Creswell and Vicki L. Plano Clark. *Designing and Conducting Mixed Methods Research*. Thousand Oaks, CA: SAGE Publications, 2017.

[10] Deutsche Postbank AG. *Screenshot from the Postbank BestSign App*. Screenshot obtained from the Apple App Store: https://apps.apple.com/de/app/postbank-bestsign/id1442251022. Accessed: Aug. 18, 2024. 2024.

[11] Android Developers. *Manifest.permission: BIND_ACCESSIBILITY_SERVICE*. Online. Available: https://bit.ly/46R26tc, Accessed: Jul. 27, 2024. 2023.

[12] Android Developers. *Manifest.permission: SYSTEM_ALERT_WINDOW*. Online. Available: https://bit.ly/4dNEXKu, Accessed: Jul. 27, 2024. 2023.

[13] K. Dunham. *Mobile Malware Attacks and Defense*. Burlington, MA: Syngress/Elsevier, 2009. ISBN: 9780080949192.

[14] Duo Security. *The 2021 State of the Auth Report: 2FA Climbs, Password Managers & Biometrics Trend*. Accessed: Jul. 26, 2023. URL: https://duo.com/blog/the-2021-state-of-the-auth-report-2fa-climbs-password-managers-biometrics-trend.

[15] M. A. Elakrat and J. C. Jung. "Development of field programmable gate array–based encryption module to mitigate man-in-the-middle attack for nuclear power plant data communication network". In: *Nuclear Engineering and Technology* 50.5 (June 2018), pp. 780–787. DOI: 10.1016/j.net.2018.01.018.

[16] European Union Agency for Cybersecurity (ENISA). *ENISA Threat Landscape 2023*. Tech. rep. Available: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2023, Accessed: Jul. 27, 2024. ENISA, Oct. 2023, p. 15.

[17] Eurostat. *Online banking penetration in the European Union and in the Euro area from 2010 to 2022 [Graph]*. Online. Available: https://www.statista.com/statistics/1310965/online-banking-penetration-in-the-european-union/. Sept. 2023.

[18] FIDO Alliance. *Barometer 2023*. Accessed: Jul. 26, 2023. 2023. URL: https://fidoalliance.org/barometer-2023/.

[19] Interaction Design Foundation. *UX Design Tools — Definitive Guide*. Online. Available: https://www.interaction-design.org/literature/article/ux-design-tools-definitive-guide, Accessed: Aug. 18, 2024. 2024.

[20] Yanick Fratantonio et al. "Cloak and Dagger: From Two Permissions to Complete Control of the UI Feedback Loop". In: *Proceedings of the 2017 IEEE Symposium on Security and Privacy (SP)*. May 2017, pp. 1041–1057. DOI: 10.1109/SP.2017.39.

[21] GuardSquare. *Mobile App Security Research Center: Overlay Attacks*. Accessed: Jul. 28, 2023. July 2023. URL: https://www.guardsquare.com/mobile-app-security-research-center/malware/overlay-attacks.

[22] Chaonian Guo et al. "Fraud Risk Monitoring System for E-Banking Transactions". In: *2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing, and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech)*. 2018, pp. 100–105. DOI: 10.1109/DASC/PiCom/DataCom/CyberSciTec.2018.00030.

[23] D. Hillman, Y. Harel, and E. Toch. "Evaluating organizational phishing awareness training on an enterprise scale". In: *Computers & Security* 132 (2023), p. 103364. ISSN: 0167-4048. DOI: 10.1016/j.cose.2023.103364.

[24] ING. *Aufträge freigeben*. Accessed: Jul. 26, 2023. URL: https://www.ing.de/hilfe/auftraege-freigeben/.

[25] ING-DiBa Banking App Article IT Finanzmagazin. *ING-DiBa entwickelt mit 5.000 Kunden bequeme Banking-App mit erstaunlich wenig Funktionen*. Online. Available: https://www.it-finanzmagazin.de/ing-diba-entwickelt-mit-5-000-kunden-bequeme-banking-app-mit-erstaunlich-wenig-funktionen-49535/, Accessed: Aug. 18, 2024. 2024.

[26] K. Jansson and R. von Solms. "Phishing for phishing awareness". In: *Behaviour & Information Technology* 32.6 (2011), pp. 584–593. ISSN: 0144-929X. DOI: 10.1080/0144929X.2011.632650.

[27] Kaspersky. *Financial Threat Report 2023: Phishing, PC and Mobile Malware*. Online. Available: https://securelist.com/financial-threat-report-2023/112526/, Accessed: Jul. 27, 2024. 2023.

[28] Kaspersky. *Kaspersky Reports on New Mobile APT Campaign Targeting iOS Devices*. Accessed: Jul. 28, 2023. 2023. URL: https://www.kaspersky.com/about/press-releases/2023_kaspersky-reports-on-new-mobile-apt-campaign-targeting-ios-devices.

[29] Kaspersky. *What is a Zero-day Exploit? - Definition and Explanation*. Online. Available: https://usa.kaspersky.com/resource-center/definitions/zero-day-exploit,Accessed:Jul.27,2024. 2023.

[30] Laith Khrais. "Highlighting the Vulnerabilities of Online Banking System". In: *The Journal of Internet Banking and Commerce* 20 (2015). DOI: 10.4172/1204-5357.1000120.

[31] Kaspersky Lab. *Worldwide organizations most targeted by phishing attacks in 2023, by industry [Graph]*. Online. Available: https://www.statista.com/statistics/420442/organizations-most-affected-by-phishing-by-industry/,Accessed:Jul.27,2024. Mar. 2024.

[32] Gordon Legge et al. "Psychophysics of reading. XI. Comparing color contrast and luminance contrast". In: *Journal of the Optical Society of America. A, Optics and image science* 7 (Nov. 1990), pp. 2002–2010. DOI: 10.1364/JOSAA.7.002002.

[33] Dr. Hansjörg Leichsenring. *Die Top 10 Kreditinstitute in Deutschland*. Online. Available: https://www.der-bank-blog.de/top10-kreditinstitute-deutschland/retail-banking/37685303/, Accessed: Jul. 28, 2024. 2022.

[34] J.R. Lewis and J. Sauro. “Item benchmarks for the system usability scale”. In: *Journal of Usability Studies* 13.3 (2018), pp. 158–167.

[35] SEON Technologies Ltd. *Global Banking Fraud Index 2023*. Online. Accessed: Jul. 27, 2024. 2023. URL: https://seon.io/resources/global-banking-fraud-index/.

[36] Ming Luo, Guihua Cui, and B. Rigg. “The development of the CIE 2000 colour-difference formula: CIEDE2000”. In: *Color Research & Application* 26 (Oct. 2001), pp. 340–350. DOI: 10.1002/col.1049.

[37] Ming R. Luo, Guihua Cui, and Brian Rigg. “Colour Difference Formulae: Past, Present and Future”. In: *Color Research & Application* 26.5 (2001), pp. 307–318. URL: https://www.researchgate.net/profile/Ming-Luo-13/publication/254682236_Colour_Difference_Formulae_Past_Present_and_Future/links/541e8d330cf241a65a18b591/Colour-Difference-Formulae-Past-Present-and-Future.pdf.

[38] Malwarebytes. *What is Spyware | Spyware Removal and Protection*. Online. Available: https://www.malwarebytes.com/de/spyware, Accessed: Aug. 18, 2024. 2024.

[39] Arham Masood, Shujhat Khan, and Wasif Afzal. “A Comprehensive Survey of Evolutionary Machine Learning”. In: *Artificial Intelligence Review* 55.8 (2022), pp. 7157–7206. DOI: 10.1007/s10462-022-10246-w. URL: https://link.springer.com/article/10.1007/s10462-022-10246-w.

[40] David Mathlogic. *Colorblind Color Palette Generator*. Online. Available: https://davidmathlogic.com/colorblind/#%23000000-%23E69F00-%2356B4E9-%23009E73-%23F0E442-%230072B2-%23D55E00-%23CC79A7, Accessed: Aug. 18, 2024. 2024.

[41] Albert A. Michelson. “Studies in optics”. In: *The University of Chicago Press* 1 (1927), pp. 1–192.

[42] Chengye Ming et al. “The Effects of Luminance Contrast and Color Combination on Icon Cognitive Performance”. In: *Color Research & Application* 47.4 (2021), pp. 564–576. DOI: 10.1002/col.22734. URL: https://www.researchgate.net/publication/349636930_The_effects_of_color_combinations_luminance_contrast_and_area_ratio_on_icon_visual_search_performance.

[43] I. Mitic. *35+ Insightful Mobile Banking & Online Banking Statistics for 2024*. Accessed: Jul. 27, 2024. 2023. URL: https://fortunly.com/statistics/online-mobile-banking-statistics/.

[44] money.co.uk. *Average monetary loss per reported fraud and cybercrime incident in the United Kingdom (UK) in 2nd quarter 2023, by type of crime (in GBP) [Graph]*. Online. Available: https://www.statista.com/statistics/1425881/uk-cybercrime-and-fraud-cases-average-loss-by-type/,Accessed:Jul.27,2024. Sept. 2023.

[45] money.co.uk. *Number of reported fraud and cybercrime incidents in the United Kingdom (UK) in 2nd quarter 2023, by type [Graph]*. Online. Available: https://www.statista.com/statistics/1425815/uk-cybercrime-and-fraud-cases-by-type/,Accessed:Jul.27,2024. Sept. 2023.

[46] David Nicholas. *Color Blindness Simulator*. Online. Available: https://davidmathlogic.com/colorblind/#D81B60-#1E88E5-#FFC107-#004D40, Accessed: Jul. 29, 2024. 2024.

[47] Helena Ojanpää and Risto Näsänen. "Effects of luminance and colour contrast on the search of information on display devices". In: *Displays* 24 (Dec. 2003), pp. 167–178. DOI: 10.1016/j.displa.2004.01.003.

[48] Postbank. *BestSign*. Accessed: Jul. 26, 2023. URL: https://www.postbank.de/privatkunden/services/online-banking/bestsign.html.

[49] Juniper Research. *Digital Banking Users to Exceed 3.6 Billion Globally by 2024, as Digital-Only Banks Catalyse Market*. Accessed: Jul. 27, 2024. 2020. URL: https://www.juniperresearch.com/press/digital-banking-users-to-exceed-3-6-billion.

[50] Juniper Research. *Number of active online banking users worldwide in 2020 with forecasts from 2021 to 2024, by region (in millions) [Graph]*. Online. Available: https://www.statista.com/statistics/1228757/online-banking-users-worldwide/. Mar. 2021.

[51] G. Rhodes et al. "Face Configuration Processing in the Human Brain: The Role of Symmetry". In: *Cerebral Cortex* (2005). URL: https://academic.oup.com/cercor/article/15/9/1276/299379.

[52] Anne Schumann et al. "Reducing Discrimination Through Bias-Aware Resampling and Decoupling Approaches". In: *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2023, Turin, Italy, September 18–22, 2023, Proceedings, Part IV*. Ed. by Nicolo Cesa-Bianchi et al. Springer, 2024, pp. 309–324. DOI: 10.1007/978-3-031-46452-2_20.

[53] Scribbr. *Pearson Correlation Coefficient*. Online. 2023. URL: %5Curl%7Bhttps://www.scribbr.com/statistics/pearson-correlation-coefficient/%7D.

[54] Heimdal Security. *What Is Session Hijacking. Session Hijacking Types and Prevention*. Online. Available: https://heimdalsecurity.com/blog/session-hijacking/, Accessed: Aug. 18, 2024. 2024.

[55] Keeper Security. *Types of Multi-Factor Authentication (MFA)*. Accessed: Jul. 26, 2023. June 2023. URL: https://www.keepersecurity.com/blog/2023/06/27/types-of-multi-factor-authentication-mfa/.

[56] Amazon Web Services. *What is Multi-Factor Authentication (MFA)?* Online. Available: https://aws.amazon.com/what-is/mfa/, Accessed: Jul. 28, 2024. 2024.

[57] Gaurav Sharma, Wencheng Wu, and Edul N. Dalal. "The CIEDE2000 Color-Difference Formula: Implementation Notes, Supplementary Test Data, and Mathematical Observations". In: *Color Research & Application* 30.1 (2005), pp. 21–30. DOI: 10.1002/col.20070. URL: https://hajim.rochester.edu/ece/sites/gsharma/ciede2000/.

[58] Jeevesh Sharma and Suhasini Verma. "Mounting Cases of Cyber-Attacks and Digital Payment". In: Available: https://www.researchgate.net/publication/371255743_Mounting_Cases_of_Cyber-Attacks_and_Digital_Payment, Accessed: Jul. 27, 2024. June 2023.

[59] Federico Sinigaglia et al. "A survey on multi-factor authentication for online banking in the wild". In: *Computers & Security* 95 (2020), p. 101745. ISSN: 0167-

4048. DOI: 10.1016/j.cose.2020.101745. URL: https://www.sciencedirect.com/science/article/pii/S0167404820300316.

[60] Sander Smets. *Protecting Against Android Overlay Attacks*. Accessed: Jul. 28, 2023. Mar. 2023. URL: https://www.guardsquare.com/blog/protecting-against-android-overlay-attacks-guardsquare.

[61] Sparkasse. *S-pushTAN*. Accessed: Jul. 26, 2023. URL: https://www.sparkasse.de/pk/produkte/konten-und-karten/finanzen-apps/s-pushtan.html.

[62] StatCounter. *Market share of mobile operating systems worldwide from 2009 to 2024, by quarter [Graph]*. Online. Available: https://www.statista.com/statistics/272698/global-market-share-held-by-mobile-operating-systems-since-2009/, Accessed: Jul. 27, 2024. May 2024.

[63] Statista. *Penetration rate of online banking worldwide in 2023, by country [Graph]*. Online. Available: https://www.statista.com/forecasts/1169529/online-banking-penetration-by-country,Accessed:Jul.27,2024. July 2024.

[64] Statista. *Neobanking - Worldwide*. Online. Available: https://www.statista.com/outlook/dmo/fintech/neobanking/worldwide#transaction-value, Accessed: Jul. 27, 2024.

[65] Statistics LibreTexts. *Testing the Significance of the Correlation Coefficient*. Online. Available: https://stats.libretexts.org/Bookshelves/Introductory_Statistics/Introductory_Statistics_1e_(OpenStax)/12%3A_Linear_Regression_and_Correlation/12.05%3A_Testing_the_Significance_of_the_Correlation_Coefficient, Accessed: Aug. 18, 2024. 2023.

[66] A. Tashakkori and C. Teddlie, eds. *SAGE Handbook of Mixed Methods in Social & Behavioral Research*. SAGE Publications, 2010.

[67] Gregory Taylor. *colormath: A Python package for color science*. Online. Available: https://github.com/gtaylor/python-colormath. 2018.

[68] Matthias Sebastian Treder. "Behind the Looking-Glass: A Review on Human Symmetry Perception". In: *Symmetry* 2.3 (2010), pp. 1510–1543. DOI: 10.3390/sym2031510. URL: https://doi.org/10.3390/sym2031510.

[69] United Nations, Department of Economic and Social Affairs, Population Division. *World Population Prospects 2024: Summary of Results*. 2024. URL: https://population.un.org/wpp/Publications/.

[70] Verimatrix. *Deconstructing a Mobile Banking App Overlay Heist*. Online. Available: https://www.verimatrix.com/cybersecurity/cybersecurity-insights/deconstructing-a-mobile-banking-app-overlay-heist/,Accessed:Jul.27,2024. 2023.

[71] Verimatrix. *Screen Spoofing: Dangerous Mobile App Overlay Attacks on the Rise*. Online. Available: https://www.verimatrix.com/cybersecurity/cybersecurity-insights/screen-spoofing-dangerous-mobile-app-overlay-attacks-on-the-rise/,Accessed:Jul.27,2024. 2024.

[72] Pauli Virtanen et al. "SciPy 1.0: Fundamental algorithms for scientific computing in Python". In: *Nature methods* 17.3 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

[73] VR Bank. *TAN Verfahren*. Accessed: Jul. 26, 2023. URL: https://www.vr.de/privatkunden/unsere-produkte/was-ist-ein-girokonto/tan-verfahren.html.

[74] WebAIM. *Contrast Checker*. Online. Available: https://webaim.org/resources/contrastchecker/, Accessed: Aug. 18, 2024. 2024.

[75] World Economic Forum. "Global Cybersecurity Outlook 2024". In: *World Economic Forum* (2024). Available: https://www.weforum.org/publications/global-cybersecurity-outlook-2024/, p. 15.

[76] World Wide Web Consortium. *Web Content Accessibility Guidelines (WCAG) 2.0*. Accessed: Jul. 28, 2024. 2008. URL: https://www.w3.org/TR/WCAG20/.

[77] Worldometers. *Europe Population (2023)*. Online. 2023. URL: https://www.worldometers.info/world-population/europe-population/.

[78] Worldometers. *Latin America and the Caribbean Population (2023)*. Online. 2023. URL: https://www.worldometers.info/world-population/latin-america-and-the-caribbean-population/.

[79] Yuxuan Yan et al. "Understanding and Detecting Overlay-based Android Malware at Market Scales". In: *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security (CCS '19)*. Nov. 2019, pp. 168–179. DOI: 10.1145/3319535.3354214.

[80] X. Zang et al. "Influences of luminance contrast and ambient lighting on visual context learning and retrieval". In: *Attention, Perception, & Psychophysics* 82 (2020). DOI: 10.3758/s13414-020-02106-y.

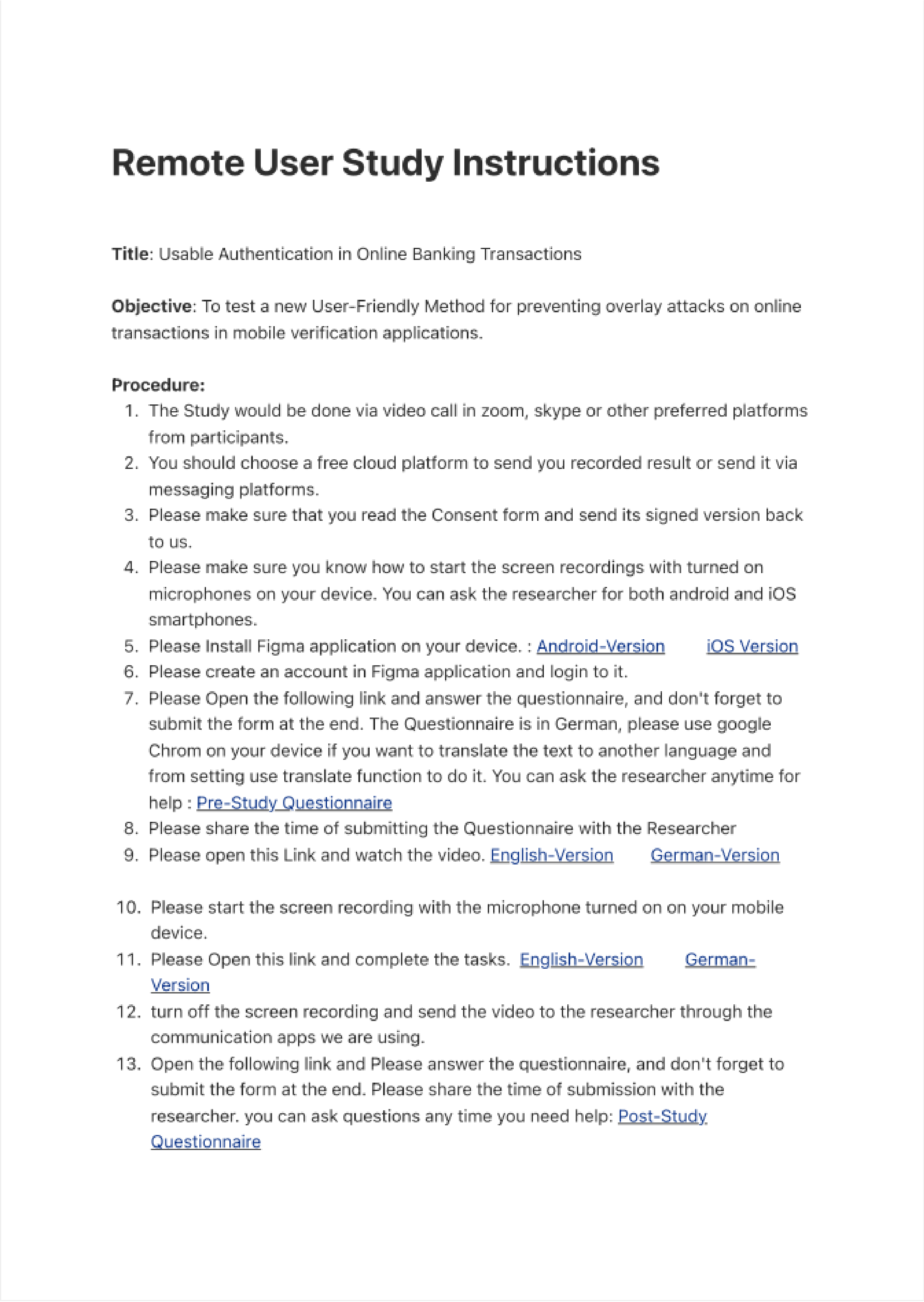
[81] Hao Zhou et al. "Beyond the Surface: Uncovering the Unprotected Components of Android Against Overlay Attack". In: Jan. 2024. DOI: 10.14722/ndss.2024.24035.

[82] Zimperium. *2023 Mobile Banking Heists Report*. Online. Available: https://get.zimperium.com/mobile-banking-heists-2023/,Accessed:Jul.27,2024. 2023.

[83] Zimperium. *Banking malware families detected worldwide in 2023, by the number of known variants [Graph]*. Online. Available: https://www.statista.com/statistics/1450078/banking-malware-families-by-known-variants/,Accessed:Jul.27,2024. Jan. 2024.

A Appendicces

A.1 Remote User Study Instructions English



Remote User Study Instructions

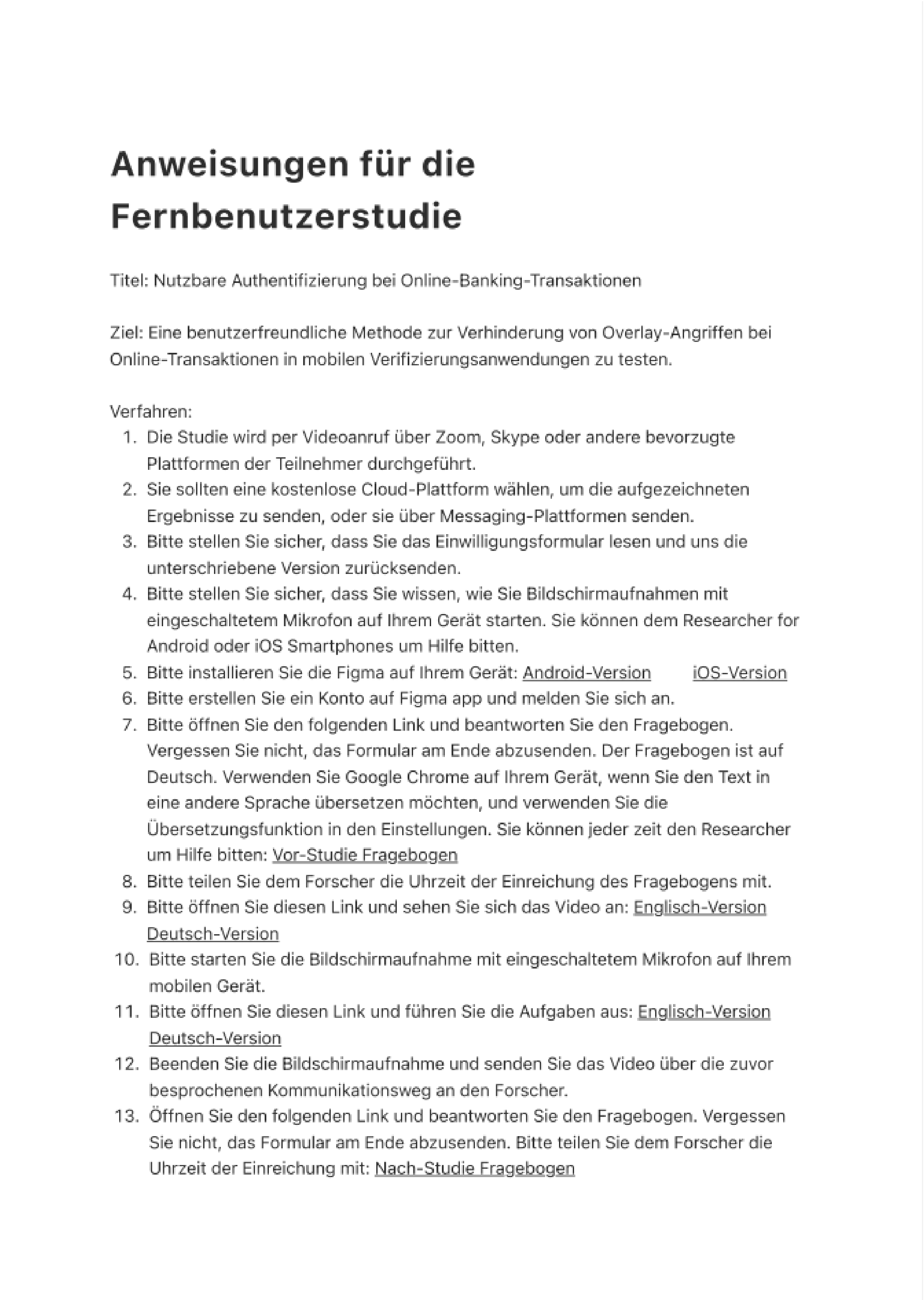
Title: Usable Authentication in Online Banking Transactions

Objective: To test a new User-Friendly Method for preventing overlay attacks on online transactions in mobile verification applications.

Procedure:

1. The Study would be done via video call in zoom, skype or other preferred platforms from participants.
2. You should choose a free cloud platform to send you recorded result or send it via messaging platforms.
3. Please make sure that you read the Consent form and send its signed version back to us.
4. Please make sure you know how to start the screen recordings with turned on microphones on your device. You can ask the researcher for both android and iOS smartphones.
5. Please Install Figma application on your device. : Android-Version iOS Version
6. Please create an account in Figma application and login to it.
7. Please Open the following link and answer the questionnaire, and don't forget to submit the form at the end. The Questionnaire is in German, please use google Chrom on your device if you want to translate the text to another language and from setting use translate function to do it. You can ask the researcher anytime for help : Pre-Study Questionnaire
8. Please share the time of submitting the Questionnaire with the Researcher
9. Please open this Link and watch the video. English-Version German-Version
10. Please start the screen recording with the microphone turned on on your mobile device.
11. Please Open this link and complete the tasks. English-Version German-Version
12. turn off the screen recording and send the video to the researcher through the communication apps we are using.
13. Open the following link and Please answer the questionnaire, and don't forget to submit the form at the end. Please share the time of submission with the researcher. you can ask questions any time you need help: Post-Study Questionnaire

A.2 Remote User Study Instructions Deutsch



Anweisungen für die Fernbenutzerstudie

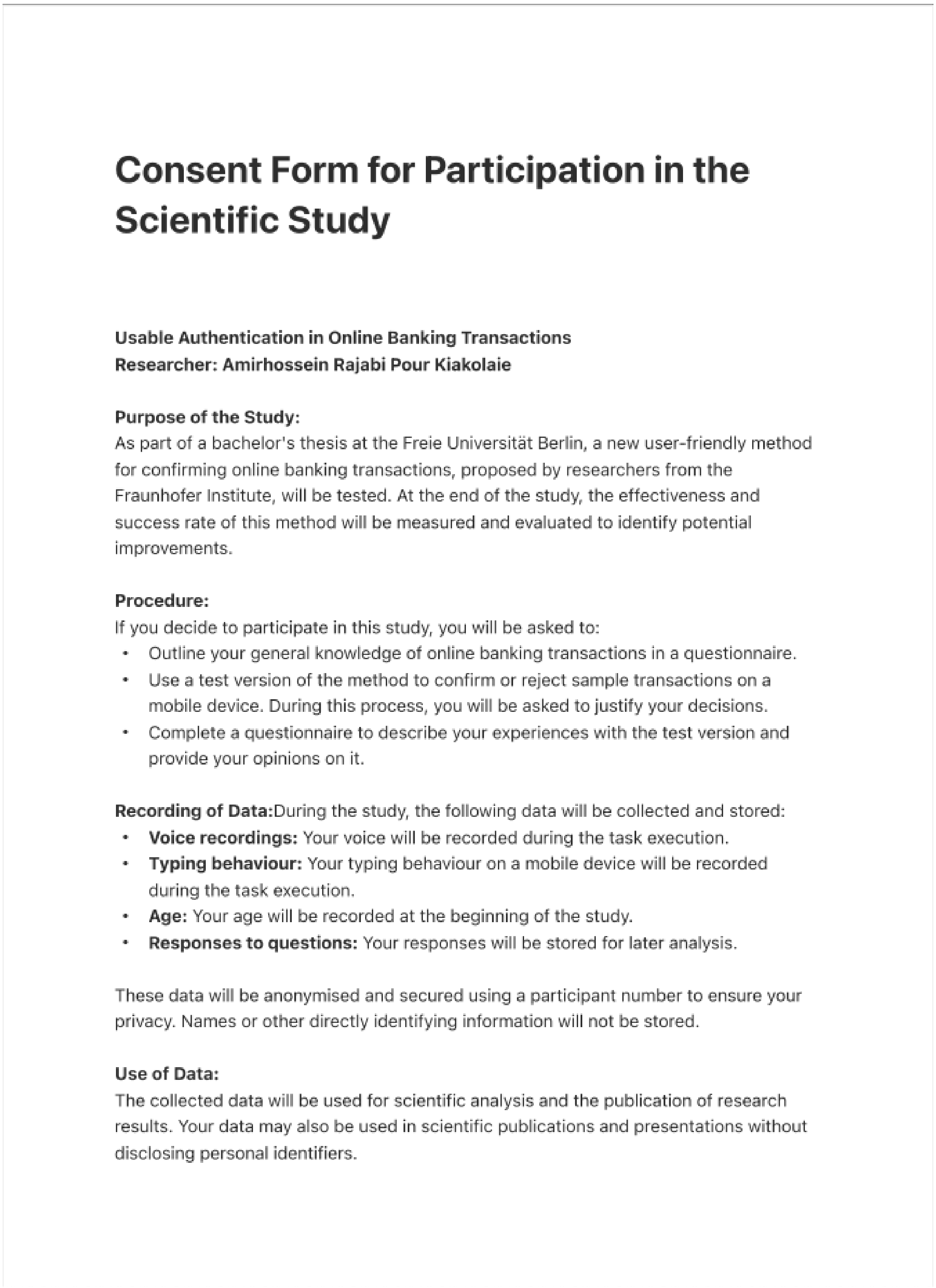
Titel: Nutzbare Authentifizierung bei Online-Banking-Transaktionen

Ziel: Eine benutzerfreundliche Methode zur Verhinderung von Overlay-Angriffen bei Online-Transaktionen in mobilen Verifizierungsanwendungen zu testen.

Verfahren:

1. Die Studie wird per Videoanruf über Zoom, Skype oder andere bevorzugte Plattformen der Teilnehmer durchgeführt.
2. Sie sollten eine kostenlose Cloud-Plattform wählen, um die aufgezeichneten Ergebnisse zu senden, oder sie über Messaging-Plattformen senden.
3. Bitte stellen Sie sicher, dass Sie das Einwilligungsformular lesen und uns die unterschriebene Version zurücksenden.
4. Bitte stellen Sie sicher, dass Sie wissen, wie Sie Bildschirmaufnahmen mit eingeschaltetem Mikrofon auf Ihrem Gerät starten. Sie können dem Researcher for Android oder iOS Smartphones um Hilfe bitten.
5. Bitte installieren Sie die Figma auf Ihrem Gerät: Android-Version iOS-Version
6. Bitte erstellen Sie ein Konto auf Figma app und melden Sie sich an.
7. Bitte öffnen Sie den folgenden Link und beantworten Sie den Fragebogen. Vergessen Sie nicht, das Formular am Ende abzusenden. Der Fragebogen ist auf Deutsch. Verwenden Sie Google Chrome auf Ihrem Gerät, wenn Sie den Text in eine andere Sprache übersetzen möchten, und verwenden Sie die Übersetzungsfunktion in den Einstellungen. Sie können jeder zeit den Researcher um Hilfe bitten: Vor-Studie Fragebogen
8. Bitte teilen Sie dem Forscher die Uhrzeit der Einreichung des Fragebogens mit.
9. Bitte öffnen Sie diesen Link und sehen Sie sich das Video an: Englisch-Version Deutsch-Version
10. Bitte starten Sie die Bildschirmaufnahme mit eingeschaltetem Mikrofon auf Ihrem mobilen Gerät.
11. Bitte öffnen Sie diesen Link und führen Sie die Aufgaben aus: Englisch-Version Deutsch-Version
12. Beenden Sie die Bildschirmaufnahme und senden Sie das Video über die zuvor besprochenen Kommunikationsweg an den Forscher.
13. Öffnen Sie den folgenden Link und beantworten Sie den Fragebogen. Vergessen Sie nicht, das Formular am Ende abzusenden. Bitte teilen Sie dem Forscher die Uhrzeit der Einreichung mit: Nach-Studie Fragebogen

A.3 Consent Form English



Consent Form for Participation in the Scientific Study

Usable Authentication in Online Banking Transactions
Researcher: Amirhossein Rajabi Pour Kiakolaie

Purpose of the Study:
As part of a bachelor's thesis at the Freie Universität Berlin, a new user-friendly method for confirming online banking transactions, proposed by researchers from the Fraunhofer Institute, will be tested. At the end of the study, the effectiveness and success rate of this method will be measured and evaluated to identify potential improvements.

Procedure:
If you decide to participate in this study, you will be asked to:

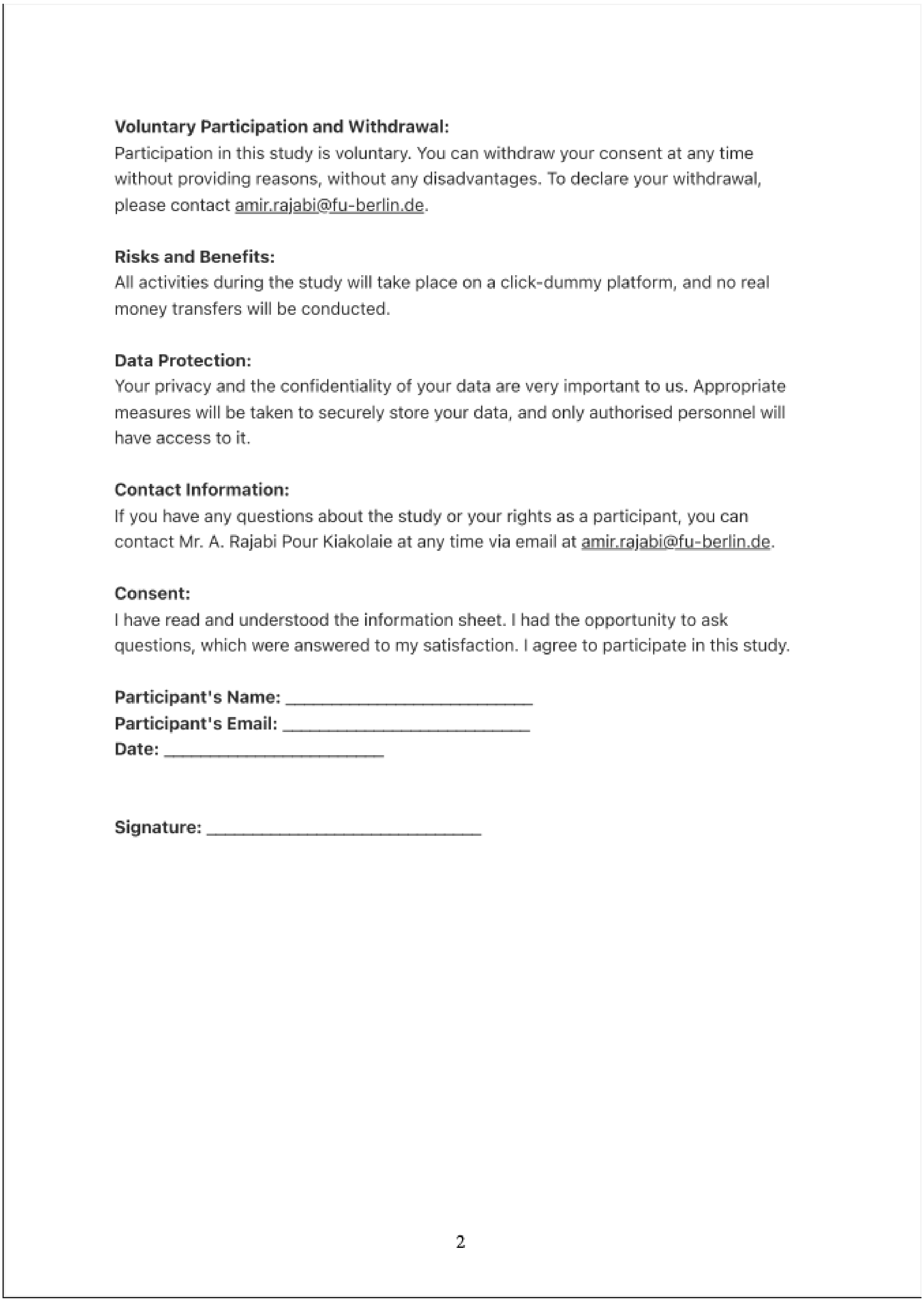
- Outline your general knowledge of online banking transactions in a questionnaire.
- Use a test version of the method to confirm or reject sample transactions on a mobile device. During this process, you will be asked to justify your decisions.
- Complete a questionnaire to describe your experiences with the test version and provide your opinions on it.

Recording of Data: During the study, the following data will be collected and stored:

- **Voice recordings:** Your voice will be recorded during the task execution.
- **Typing behaviour:** Your typing behaviour on a mobile device will be recorded during the task execution.
- **Age:** Your age will be recorded at the beginning of the study.
- **Responses to questions:** Your responses will be stored for later analysis.

These data will be anonymised and secured using a participant number to ensure your privacy. Names or other directly identifying information will not be stored.

Use of Data:
The collected data will be used for scientific analysis and the publication of research results. Your data may also be used in scientific publications and presentations without disclosing personal identifiers.



Voluntary Participation and Withdrawal:
Participation in this study is voluntary. You can withdraw your consent at any time without providing reasons, without any disadvantages. To declare your withdrawal, please contact amir.rajabi@fu-berlin.de.

Risks and Benefits:
All activities during the study will take place on a click-dummy platform, and no real money transfers will be conducted.

Data Protection:
Your privacy and the confidentiality of your data are very important to us. Appropriate measures will be taken to securely store your data, and only authorised personnel will have access to it.

Contact Information:
If you have any questions about the study or your rights as a participant, you can contact Mr. A. Rajabi Pour Kiakolaie at any time via email at amir.rajabi@fu-berlin.de.

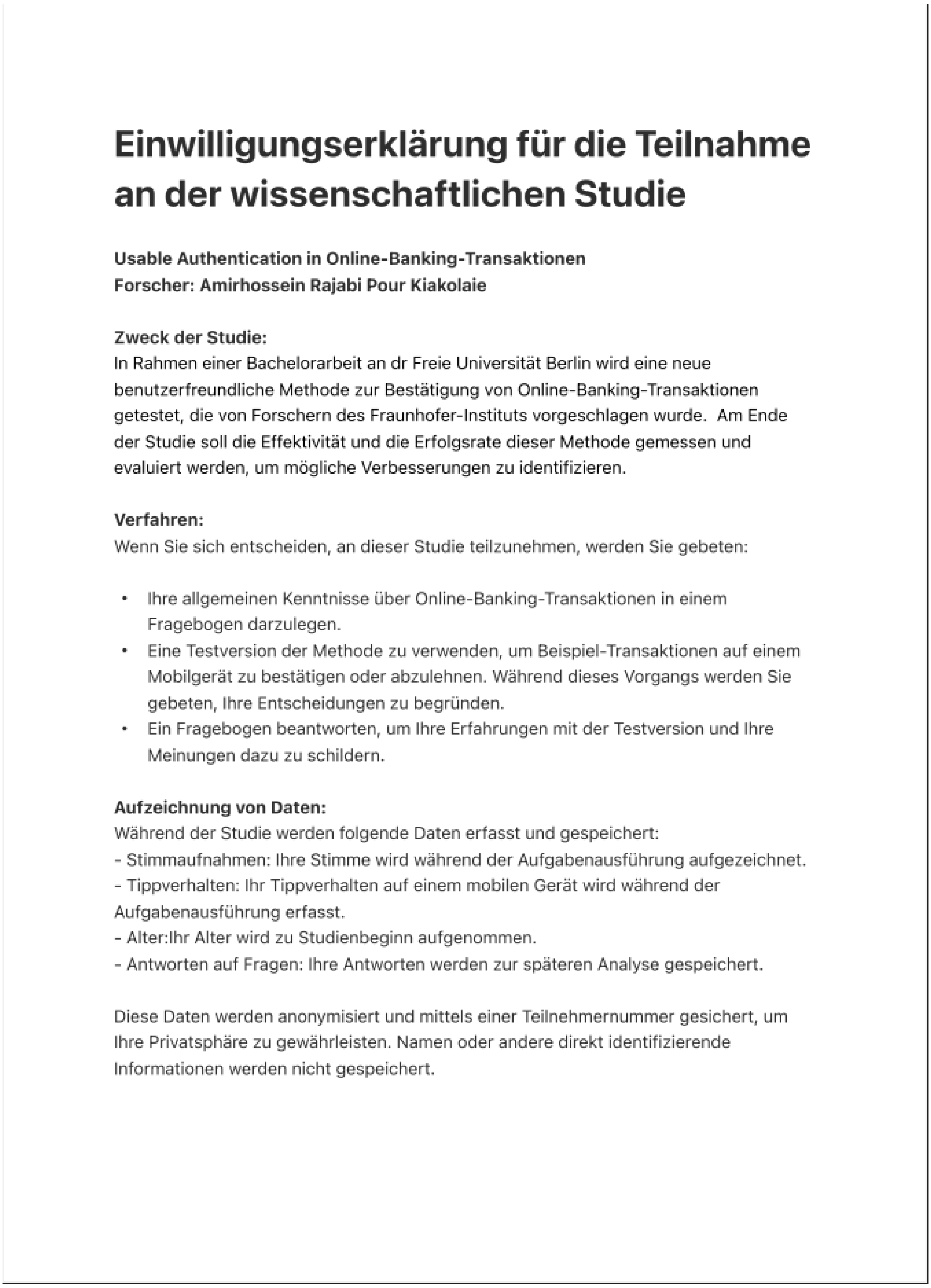
Consent:
I have read and understood the information sheet. I had the opportunity to ask questions, which were answered to my satisfaction. I agree to participate in this study.

Participant's Name: ___________________________
Participant's Email: ___________________________
Date: ________________________

Signature: ______________________________

2

A.4 Consent Form Deutsch



Einwilligungserklärung für die Teilnahme an der wissenschaftlichen Studie

Usable Authentication in Online-Banking-Transaktionen
Forscher: Amirhossein Rajabi Pour Kiakolaie

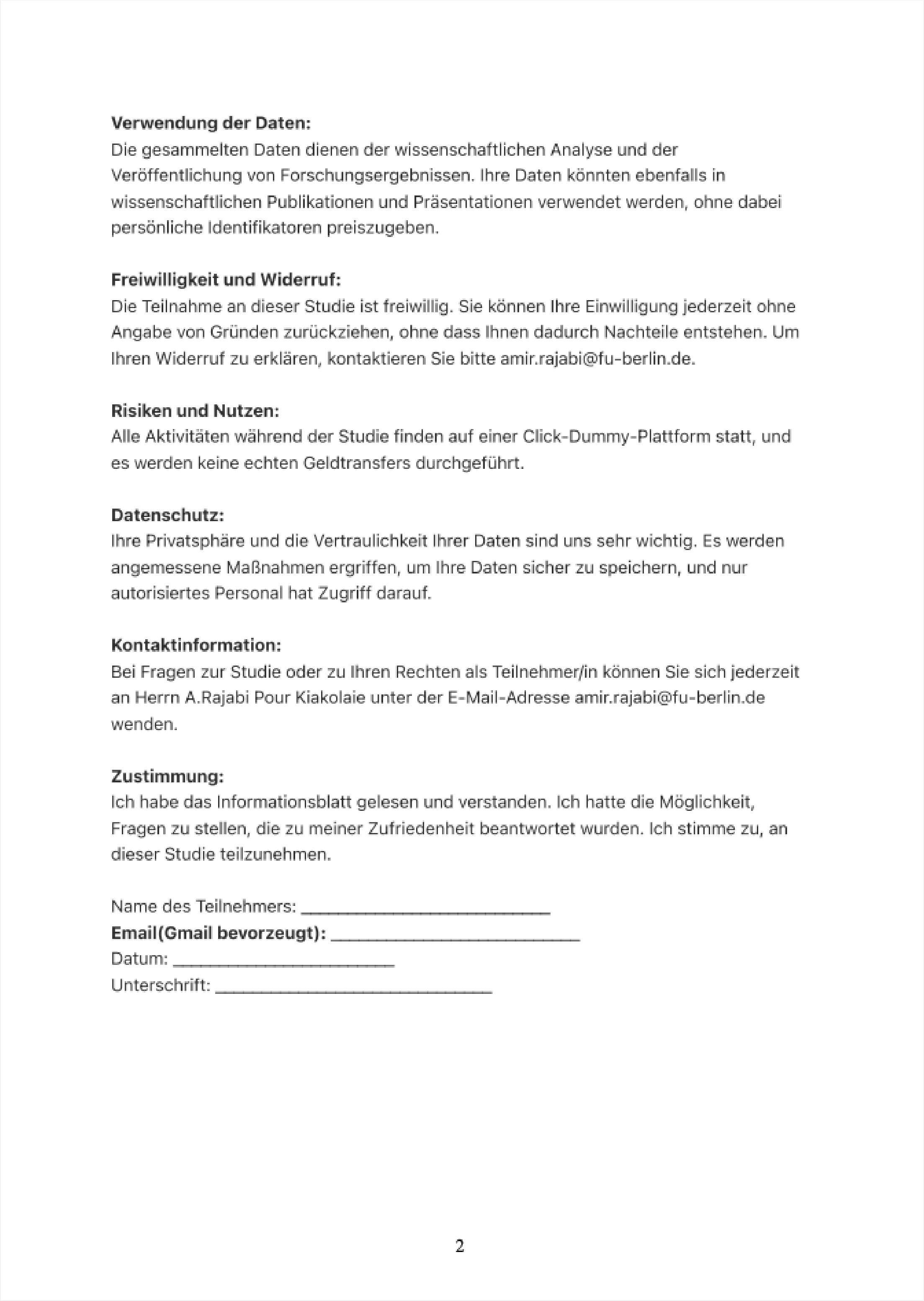
Zweck der Studie:
In Rahmen einer Bachelorarbeit an dr Freie Universität Berlin wird eine neue benutzerfreundliche Methode zur Bestätigung von Online-Banking-Transaktionen getestet, die von Forschern des Fraunhofer-Instituts vorgeschlagen wurde. Am Ende der Studie soll die Effektivität und die Erfolgsrate dieser Methode gemessen und evaluiert werden, um mögliche Verbesserungen zu identifizieren.

Verfahren:
Wenn Sie sich entscheiden, an dieser Studie teilzunehmen, werden Sie gebeten:

- Ihre allgemeinen Kenntnisse über Online-Banking-Transaktionen in einem Fragebogen darzulegen.
- Eine Testversion der Methode zu verwenden, um Beispiel-Transaktionen auf einem Mobilgerät zu bestätigen oder abzulehnen. Während dieses Vorgangs werden Sie gebeten, Ihre Entscheidungen zu begründen.
- Ein Fragebogen beantworten, um Ihre Erfahrungen mit der Testversion und Ihre Meinungen dazu zu schildern.

Aufzeichnung von Daten:
Während der Studie werden folgende Daten erfasst und gespeichert:
- Stimmaufnahmen: Ihre Stimme wird während der Aufgabenausführung aufgezeichnet.
- Tippverhalten: Ihr Tippverhalten auf einem mobilen Gerät wird während der Aufgabenausführung erfasst.
- Alter:Ihr Alter wird zu Studienbeginn aufgenommen.
- Antworten auf Fragen: Ihre Antworten werden zur späteren Analyse gespeichert.

Diese Daten werden anonymisiert und mittels einer Teilnehmernummer gesichert, um Ihre Privatsphäre zu gewährleisten. Namen oder andere direkt identifizierende Informationen werden nicht gespeichert.



Verwendung der Daten:
Die gesammelten Daten dienen der wissenschaftlichen Analyse und der Veröffentlichung von Forschungsergebnissen. Ihre Daten könnten ebenfalls in wissenschaftlichen Publikationen und Präsentationen verwendet werden, ohne dabei persönliche Identifikatoren preiszugeben.

Freiwilligkeit und Widerruf:
Die Teilnahme an dieser Studie ist freiwillig. Sie können Ihre Einwilligung jederzeit ohne Angabe von Gründen zurückziehen, ohne dass Ihnen dadurch Nachteile entstehen. Um Ihren Widerruf zu erklären, kontaktieren Sie bitte amir.rajabi@fu-berlin.de.

Risiken und Nutzen:
Alle Aktivitäten während der Studie finden auf einer Click-Dummy-Plattform statt, und es werden keine echten Geldtransfers durchgeführt.

Datenschutz:
Ihre Privatsphäre und die Vertraulichkeit Ihrer Daten sind uns sehr wichtig. Es werden angemessene Maßnahmen ergriffen, um Ihre Daten sicher zu speichern, und nur autorisiertes Personal hat Zugriff darauf.

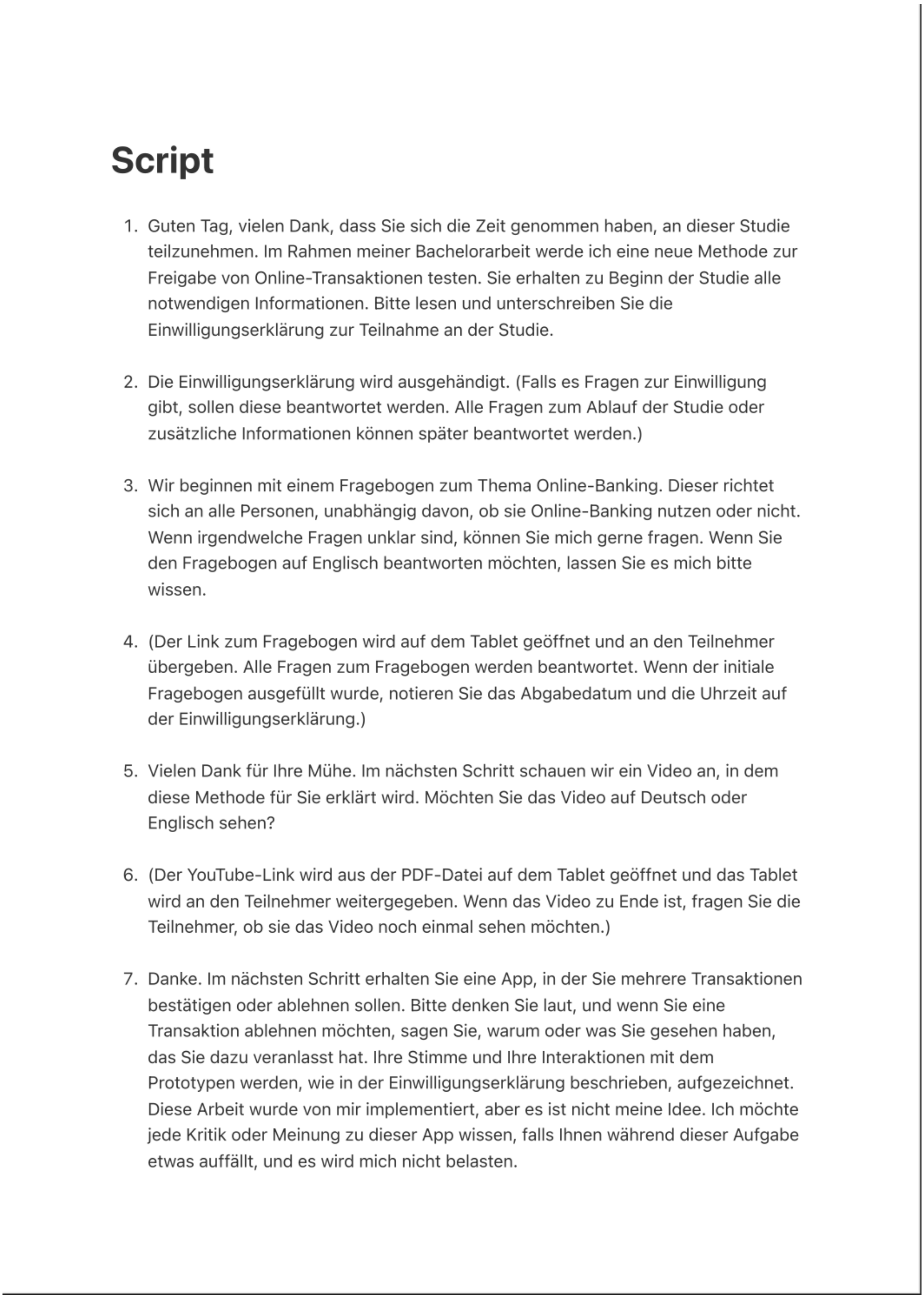
Kontaktinformation:
Bei Fragen zur Studie oder zu Ihren Rechten als Teilnehmer/in können Sie sich jederzeit an Herrn A.Rajabi Pour Kiakolaie unter der E-Mail-Adresse amir.rajabi@fu-berlin.de wenden.

Zustimmung:
Ich habe das Informationsblatt gelesen und verstanden. Ich hatte die Möglichkeit, Fragen zu stellen, die zu meiner Zufriedenheit beantwortet wurden. Ich stimme zu, an dieser Studie teilzunehmen.

Name des Teilnehmers: ___________________________
Email(Gmail bevorzeugt): ___________________________
Datum: ________________________
Unterschrift: ______________________________

2

A.5 Study script Deutsch



Script

1. Guten Tag, vielen Dank, dass Sie sich die Zeit genommen haben, an dieser Studie teilzunehmen. Im Rahmen meiner Bachelorarbeit werde ich eine neue Methode zur Freigabe von Online-Transaktionen testen. Sie erhalten zu Beginn der Studie alle notwendigen Informationen. Bitte lesen und unterschreiben Sie die Einwilligungserklärung zur Teilnahme an der Studie.

2. Die Einwilligungserklärung wird ausgehändigt. (Falls es Fragen zur Einwilligung gibt, sollen diese beantwortet werden. Alle Fragen zum Ablauf der Studie oder zusätzliche Informationen können später beantwortet werden.)

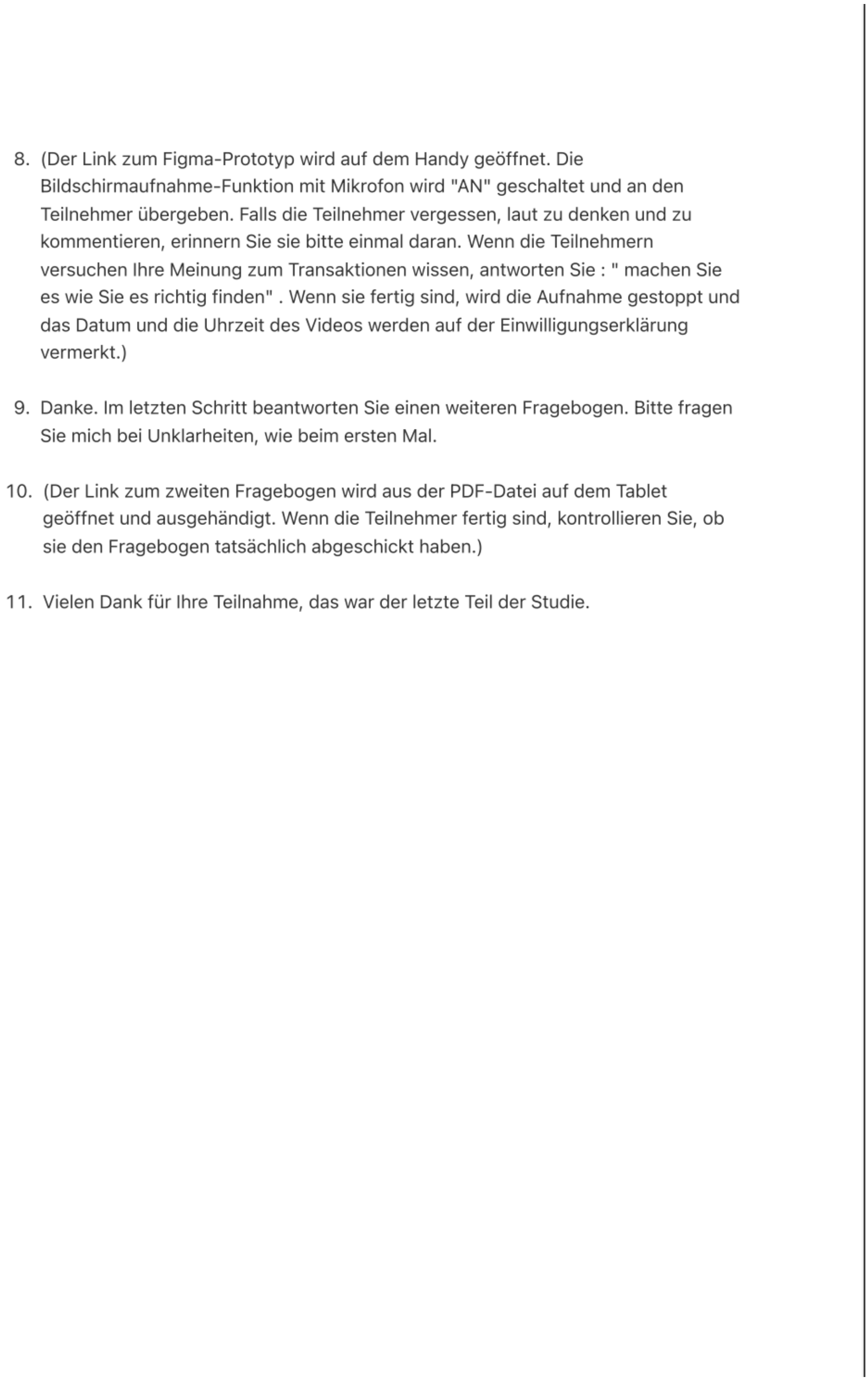
3. Wir beginnen mit einem Fragebogen zum Thema Online-Banking. Dieser richtet sich an alle Personen, unabhängig davon, ob sie Online-Banking nutzen oder nicht. Wenn irgendwelche Fragen unklar sind, können Sie mich gerne fragen. Wenn Sie den Fragebogen auf Englisch beantworten möchten, lassen Sie es mich bitte wissen.

4. (Der Link zum Fragebogen wird auf dem Tablet geöffnet und an den Teilnehmer übergeben. Alle Fragen zum Fragebogen werden beantwortet. Wenn der initiale Fragebogen ausgefüllt wurde, notieren Sie das Abgabedatum und die Uhrzeit auf der Einwilligungserklärung.)

5. Vielen Dank für Ihre Mühe. Im nächsten Schritt schauen wir ein Video an, in dem diese Methode für Sie erklärt wird. Möchten Sie das Video auf Deutsch oder Englisch sehen?

6. (Der YouTube-Link wird aus der PDF-Datei auf dem Tablet geöffnet und das Tablet wird an den Teilnehmer weitergegeben. Wenn das Video zu Ende ist, fragen Sie die Teilnehmer, ob sie das Video noch einmal sehen möchten.)

7. Danke. Im nächsten Schritt erhalten Sie eine App, in der Sie mehrere Transaktionen bestätigen oder ablehnen sollen. Bitte denken Sie laut, und wenn Sie eine Transaktion ablehnen möchten, sagen Sie, warum oder was Sie gesehen haben, das Sie dazu veranlasst hat. Ihre Stimme und Ihre Interaktionen mit dem Prototypen werden, wie in der Einwilligungserklärung beschrieben, aufgezeichnet. Diese Arbeit wurde von mir implementiert, aber es ist nicht meine Idee. Ich möchte jede Kritik oder Meinung zu dieser App wissen, falls Ihnen während dieser Aufgabe etwas auffällt, und es wird mich nicht belasten.



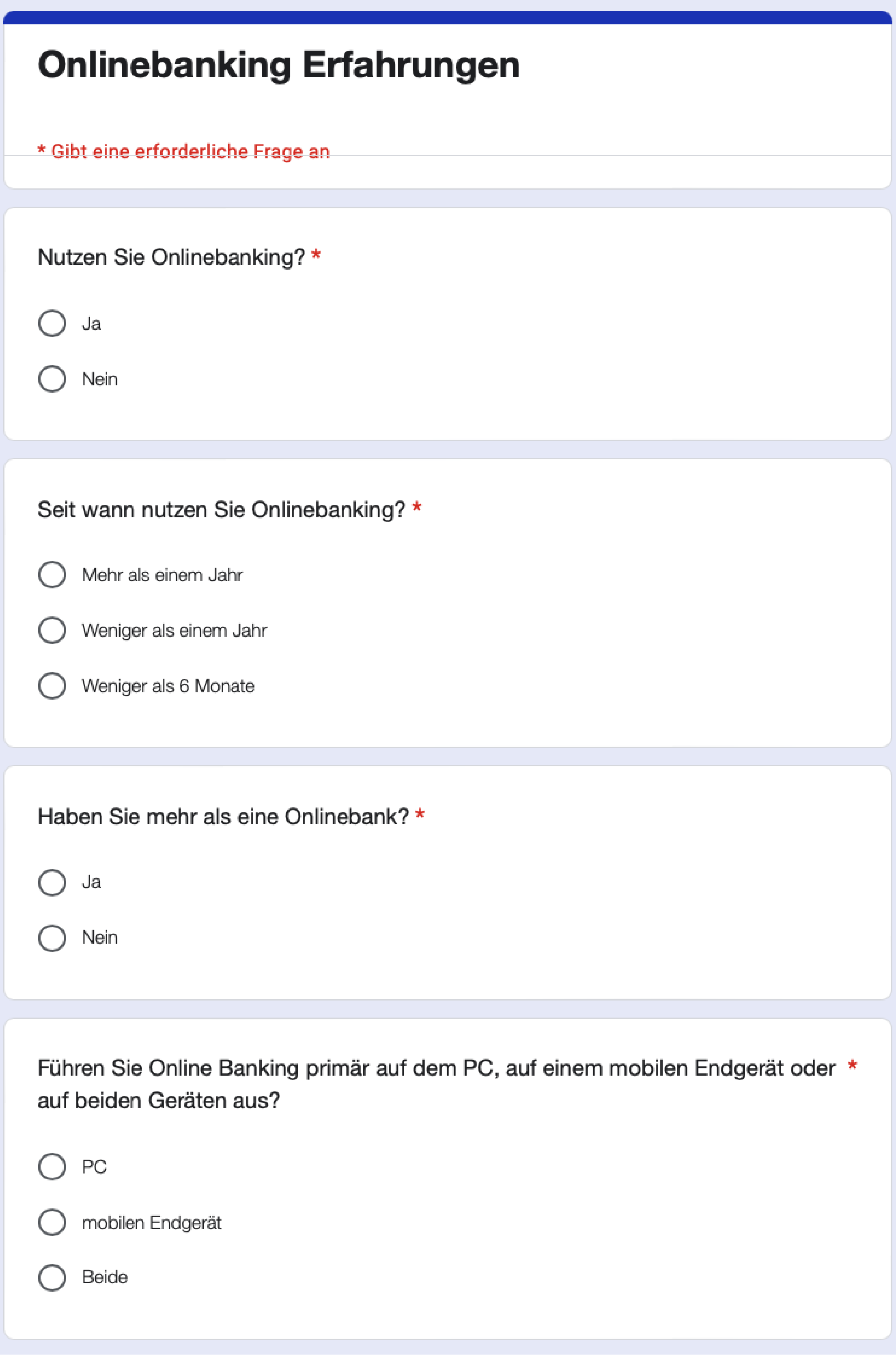
8. (Der Link zum Figma-Prototyp wird auf dem Handy geöffnet. Die Bildschirmaufnahme-Funktion mit Mikrofon wird "AN" geschaltet und an den Teilnehmer übergeben. Falls die Teilnehmer vergessen, laut zu denken und zu kommentieren, erinnern Sie sie bitte einmal daran. Wenn die Teilnehmern versuchen Ihre Meinung zum Transaktionen wissen, antworten Sie : " machen Sie es wie Sie es richtig finden" . Wenn sie fertig sind, wird die Aufnahme gestoppt und das Datum und die Uhrzeit des Videos werden auf der Einwilligungserklärung vermerkt.)

9. Danke. Im letzten Schritt beantworten Sie einen weiteren Fragebogen. Bitte fragen Sie mich bei Unklarheiten, wie beim ersten Mal.

10. (Der Link zum zweiten Fragebogen wird aus der PDF-Datei auf dem Tablet geöffnet und ausgehändigt. Wenn die Teilnehmer fertig sind, kontrollieren Sie, ob sie den Fragebogen tatsächlich abgeschickt haben.)

11. Vielen Dank für Ihre Teilnahme, das war der letzte Teil der Studie.

A.6 Initial Questionnaire Deustch



Onlinebanking Erfahrungen

* Gibt eine erforderliche Frage an

Nutzen Sie Onlinebanking? *

○ Ja

○ Nein

Seit wann nutzen Sie Onlinebanking? *

○ Mehr als einem Jahr

○ Weniger als einem Jahr

○ Weniger als 6 Monate

Haben Sie mehr als eine Onlinebank? *

○ Ja

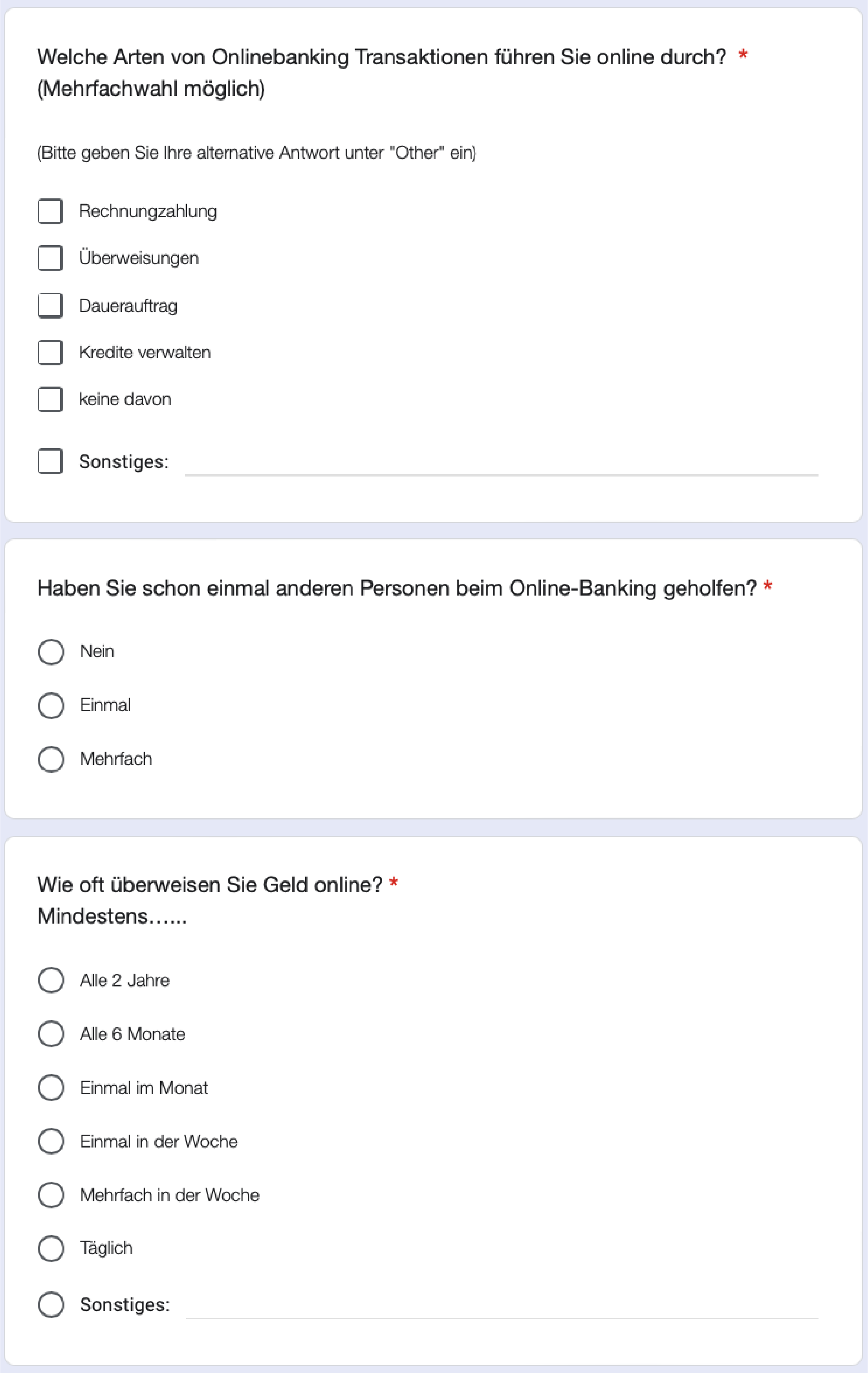
○ Nein

Führen Sie Online Banking primär auf dem PC, auf einem mobilen Endgerät oder auf beiden Geräten aus? *

○ PC

○ mobilen Endgerät

○ Beide



Welche Arten von Onlinebanking Transaktionen führen Sie online durch? *
(Mehrfachwahl möglich)

(Bitte geben Sie Ihre alternative Antwort unter "Other" ein)

□ Rechnungzahlung

□ Überweisungen

□ Dauerauftrag

□ Kredite verwalten

□ keine davon

□ Sonstiges: ____________________

Haben Sie schon einmal anderen Personen beim Online-Banking geholfen? *

○ Nein

○ Einmal

○ Mehrfach

Wie oft überweisen Sie Geld online? *
Mindestens......

○ Alle 2 Jahre

○ Alle 6 Monate

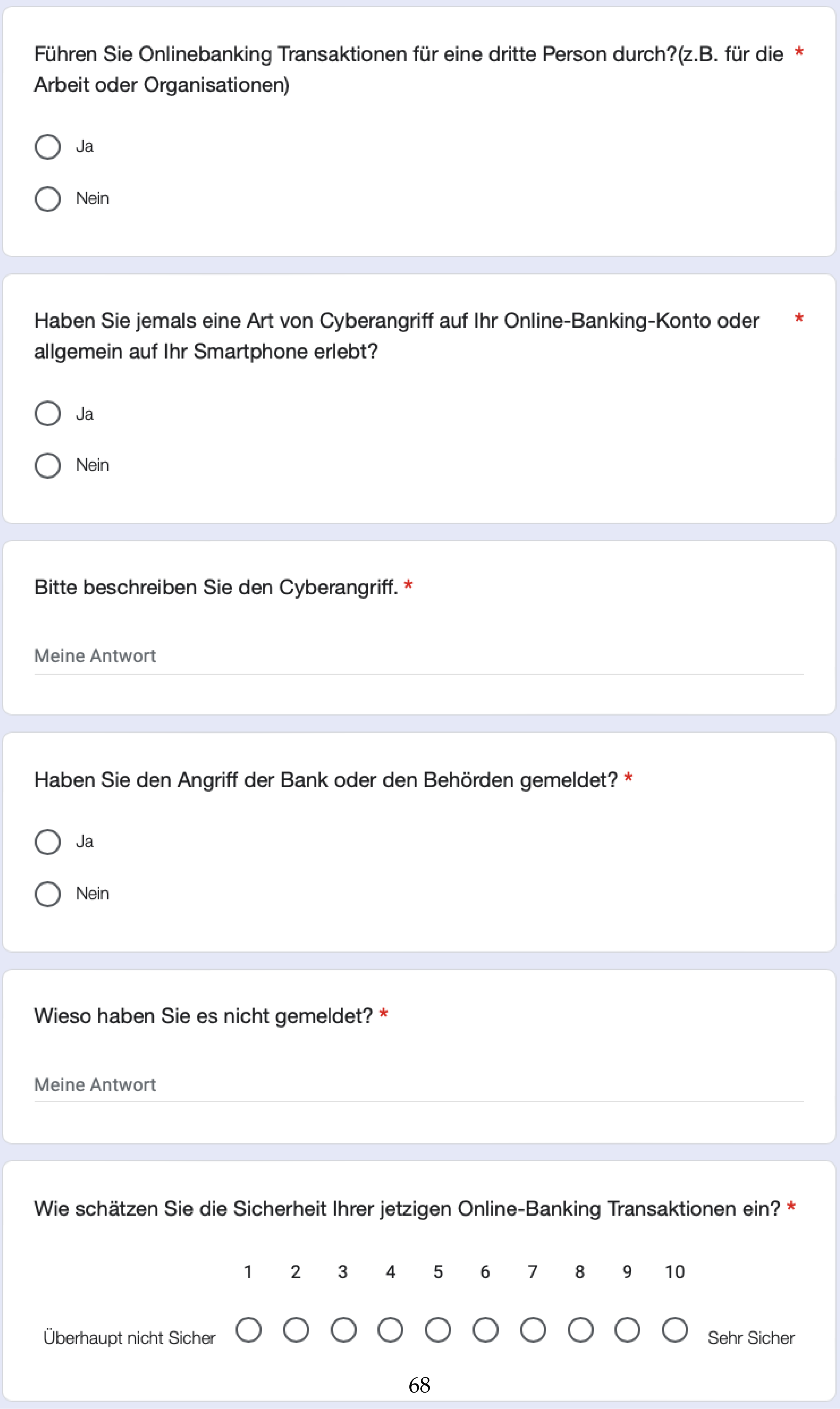
○ Einmal im Monat

○ Einmal in der Woche

○ Mehrfach in der Woche

○ Täglich

○ Sonstiges: ____________________



Führen Sie Onlinebanking Transaktionen für eine dritte Person durch?(z.B. für die Arbeit oder Organisationen) *

○ Ja

○ Nein

Haben Sie jemals eine Art von Cyberangriff auf Ihr Online-Banking-Konto oder allgemein auf Ihr Smartphone erlebt? *

○ Ja

○ Nein

Bitte beschreiben Sie den Cyberangriff. *

Meine Antwort

Haben Sie den Angriff der Bank oder den Behörden gemeldet? *

○ Ja

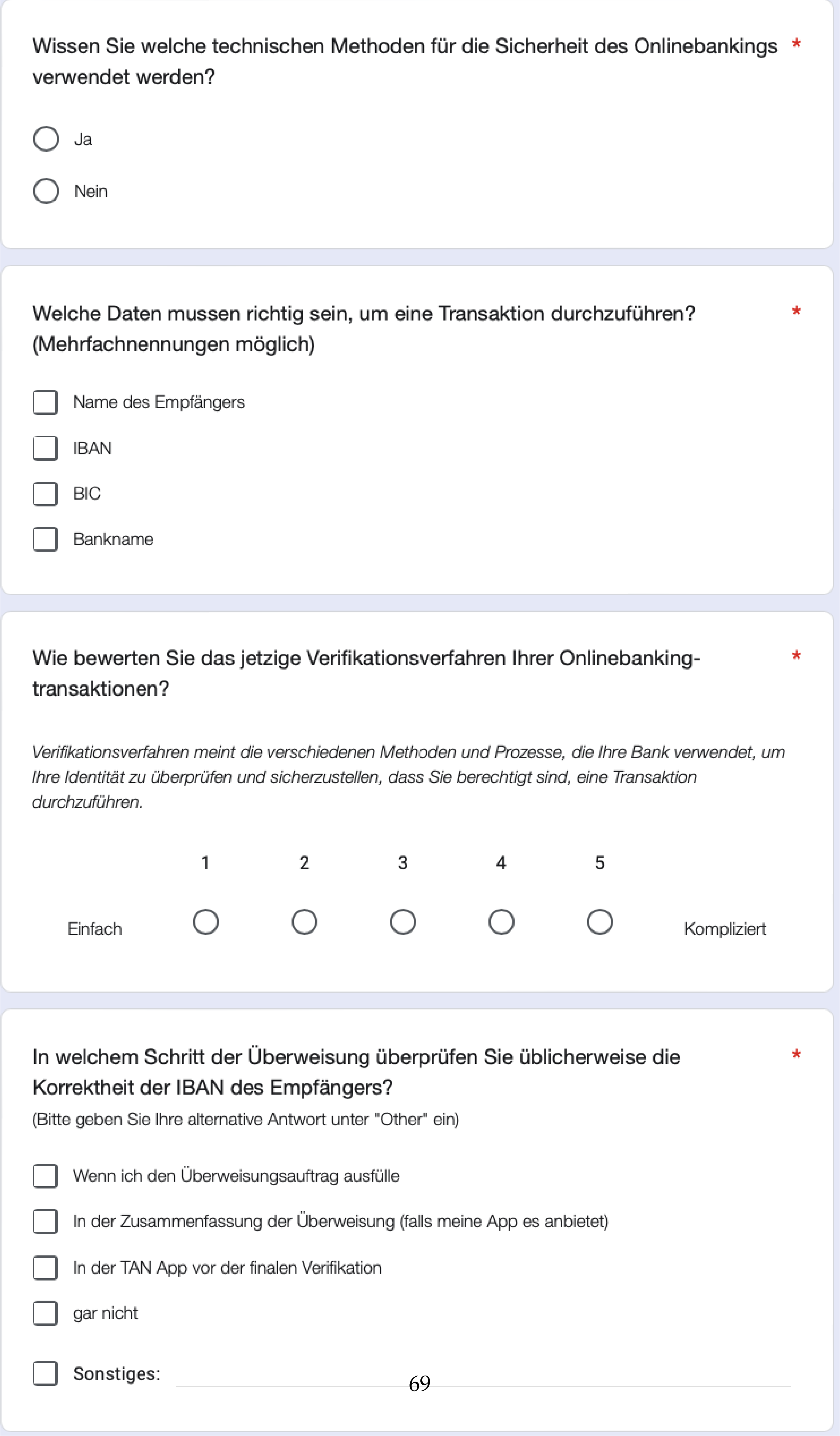
○ Nein

Wieso haben Sie es nicht gemeldet? *

Meine Antwort

Wie schätzen Sie die Sicherheit Ihrer jetzigen Online-Banking Transaktionen ein? *

	1	2	3	4	5	6	7	8	9	10	
Überhaupt nicht Sicher	○	○	○	○	○	○	○	○	○	○	Sehr Sicher



Wissen Sie welche technischen Methoden für die Sicherheit des Onlinebankings verwendet werden? *

○ Ja

○ Nein

Welche Daten mussen richtig sein, um eine Transaktion durchzuführen? (Mehrfachnennungen möglich) *

□ Name des Empfängers

□ IBAN

□ BIC

□ Bankname

Wie bewerten Sie das jetzige Verifikationsverfahren Ihrer Onlinebanking-transaktionen? *

Verifikationsverfahren meint die verschiedenen Methoden und Prozesse, die Ihre Bank verwendet, um Ihre Identität zu überprüfen und sicherzustellen, dass Sie berechtigt sind, eine Transaktion durchzuführen.

	1	2	3	4	5	
Einfach	○	○	○	○	○	Kompliziert

In welchem Schritt der Überweisung überprüfen Sie üblicherweise die Korrektheit der IBAN des Empfängers? *

(Bitte geben Sie Ihre alternative Antwort unter "Other" ein)

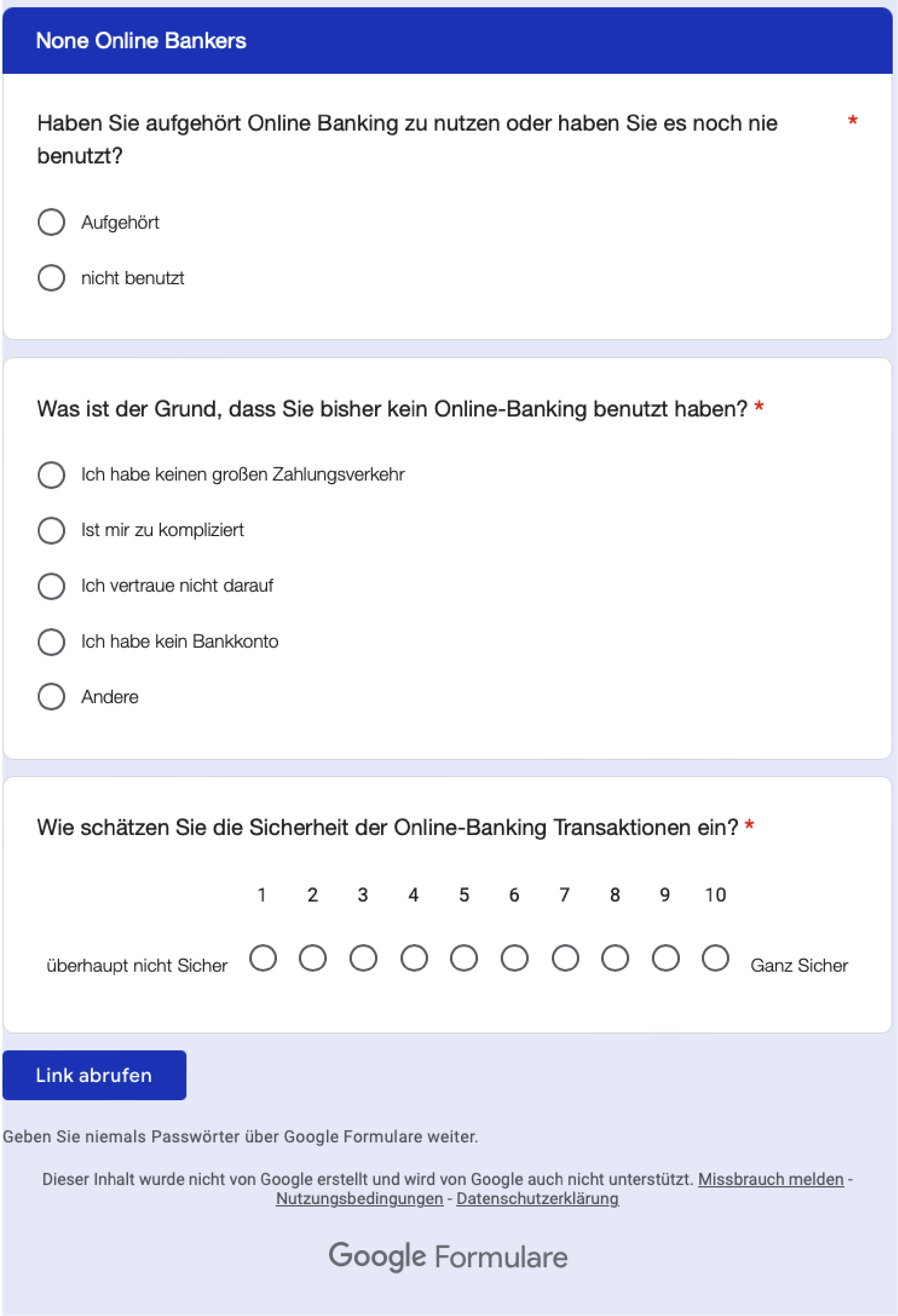
□ Wenn ich den Überweisungsauftrag ausfülle

□ In der Zusammenfassung der Überweisung (falls meine App es anbietet)

□ In der TAN App vor der finalen Verifikation

□ gar nicht

□ Sonstiges: ____________________



None Online Bankers

Haben Sie aufgehört Online Banking zu nutzen oder haben Sie es noch nie benutzt? *

○ Aufgehört

○ nicht benutzt

Was ist der Grund, dass Sie bisher kein Online-Banking benutzt haben? *

○ Ich habe keinen großen Zahlungsverkehr

○ Ist mir zu kompliziert

○ Ich vertraue nicht darauf

○ Ich habe kein Bankkonto

○ Andere

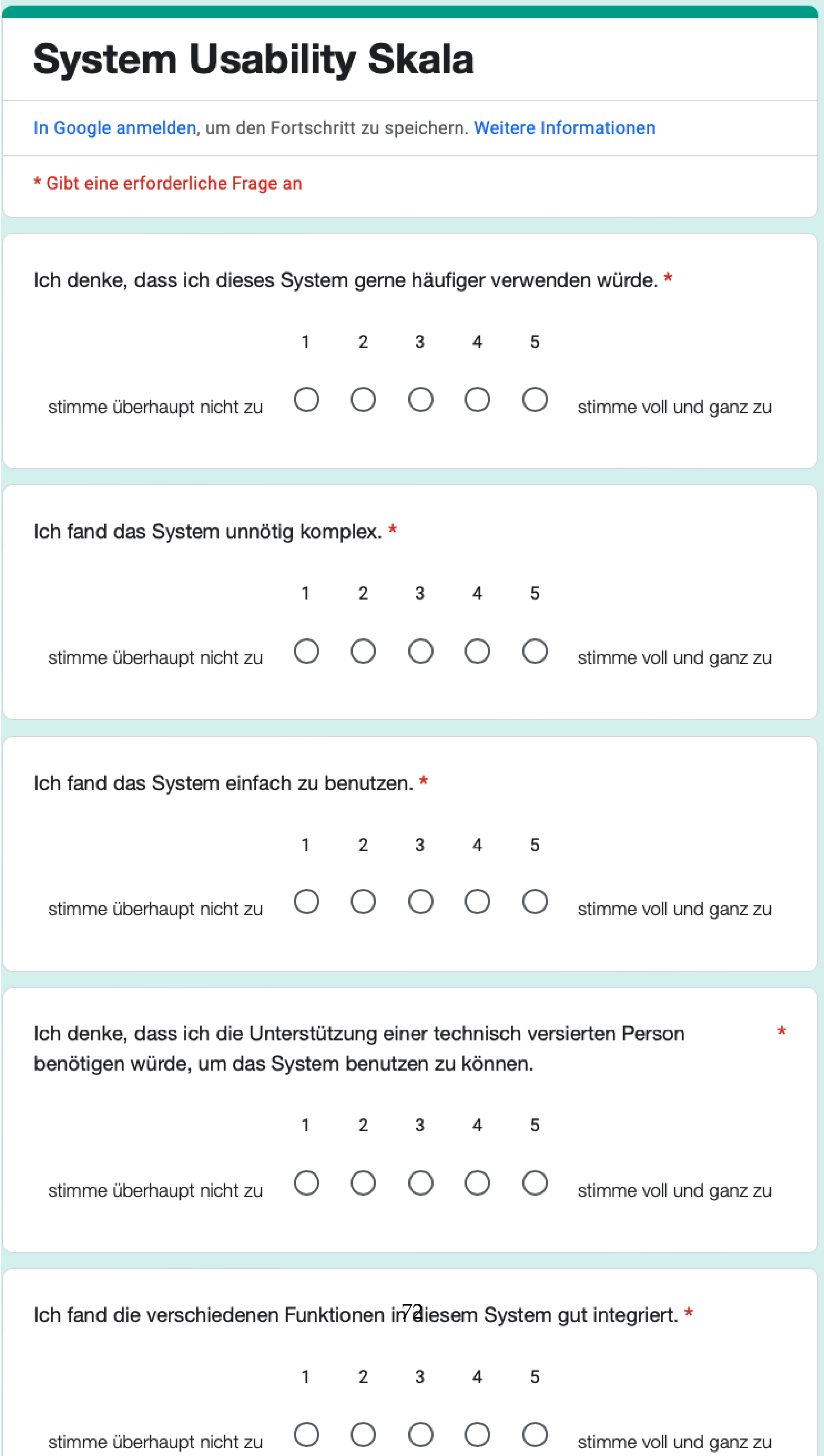
Wie schätzen Sie die Sicherheit der Online-Banking Transaktionen ein? *

	1	2	3	4	5	6	7	8	9	10	
überhaupt nicht Sicher	○	○	○	○	○	○	○	○	○	○	Ganz Sicher

Geben Sie niemals Passwörter über Google Formulare weiter.

Dieser Inhalt wurde nicht von Google erstellt und wird von Google auch nicht unterstützt. Missbrauch melden - Nutzungsbedingungen - Datenschutzerklärung

A.7 Post-Interaction Questionnaire Deustch



System Usability Skala

In Google anmelden, um den Fortschritt zu speichern. Weitere Informationen

* Gibt eine erforderliche Frage an

Ich denke, dass ich dieses System gerne häufiger verwenden würde. *

	1	2	3	4	5	
stimme überhaupt nicht zu	○	○	○	○	○	stimme voll und ganz zu

Ich fand das System unnötig komplex. *

	1	2	3	4	5	
stimme überhaupt nicht zu	○	○	○	○	○	stimme voll und ganz zu

Ich fand das System einfach zu benutzen. *

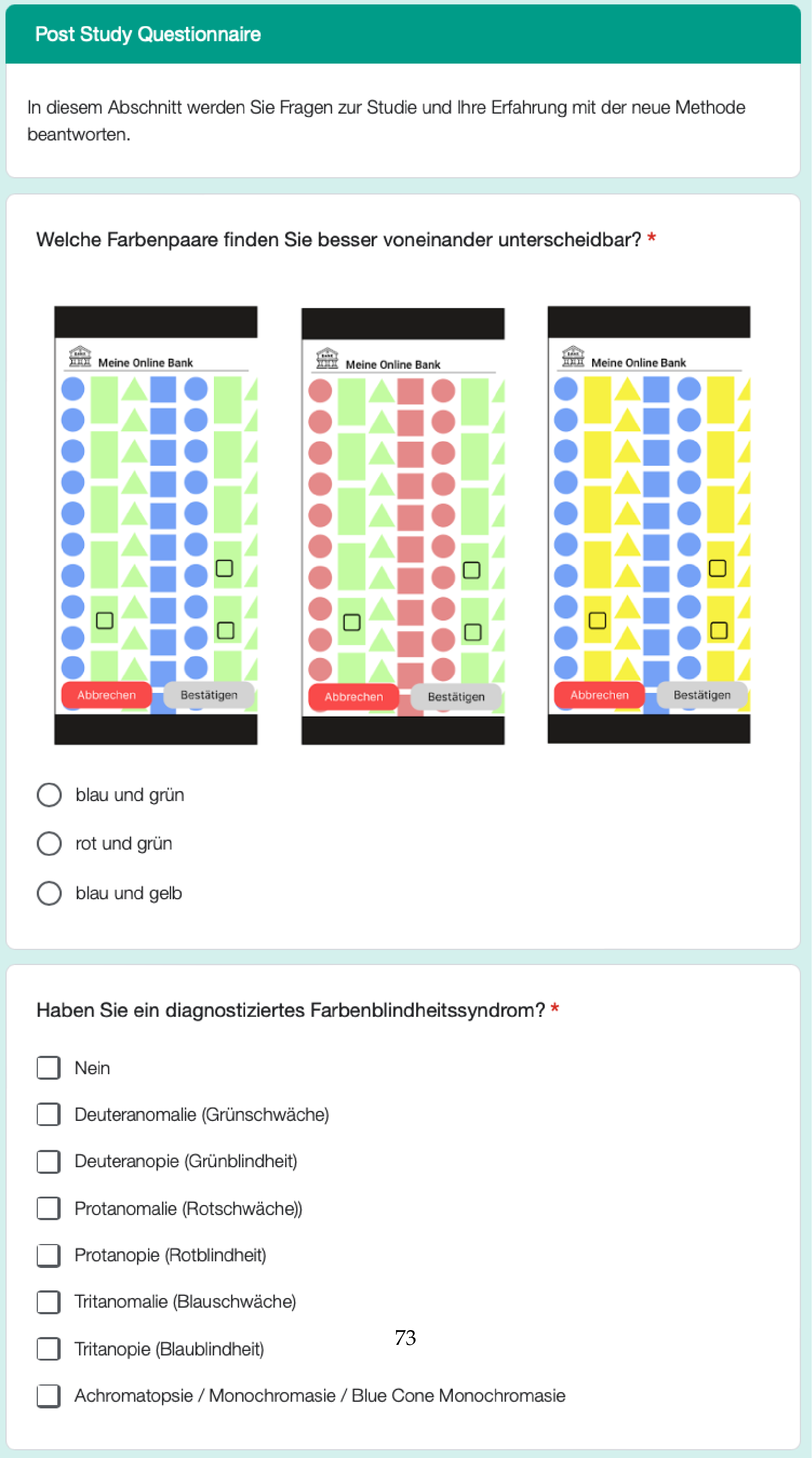
	1	2	3	4	5	
stimme überhaupt nicht zu	○	○	○	○	○	stimme voll und ganz zu

Ich denke, dass ich die Unterstützung einer technisch versierten Person benötigen würde, um das System benutzen zu können. *

	1	2	3	4	5	
stimme überhaupt nicht zu	○	○	○	○	○	stimme voll und ganz zu

Ich fand die verschiedenen Funktionen in diesem System gut integriert. *

	1	2	3	4	5	
stimme überhaupt nicht zu	○	○	○	○	○	stimme voll und ganz zu



Post Study Questionnaire

In diesem Abschnitt werden Sie Fragen zur Studie und Ihre Erfahrung mit der neue Methode beantworten.

Welche Farbenpaare finden Sie besser voneinander unterscheidbar? *

○ blau und grün

○ rot und grün

○ blau und gelb

Haben Sie ein diagnostiziertes Farbenblindheitssyndrom? *

□ Nein

□ Deuteranomalie (Grünschwäche)

□ Deuteranopie (Grünblindheit)

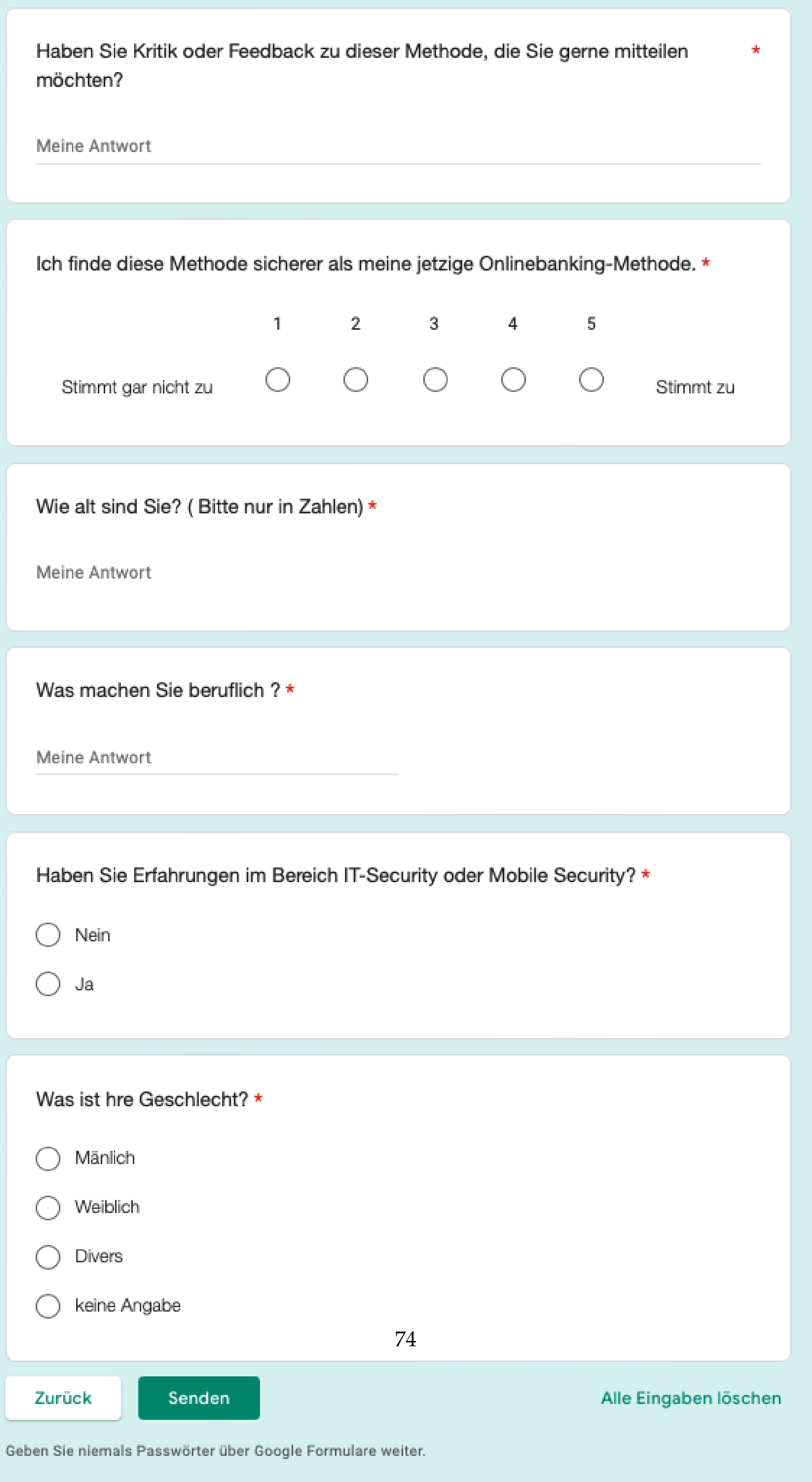
□ Protanomalie (Rotschwäche))

□ Protanopie (Rotblindheit)

□ Tritanomalie (Blauschwäche)

□ Tritanopie (Blaublindheit)

□ Achromatopsie / Monochromasie / Blue Cone Monochromasie



Haben Sie Kritik oder Feedback zu dieser Methode, die Sie gerne mitteilen möchten? *

Meine Antwort

Ich finde diese Methode sicherer als meine jetzige Onlinebanking-Methode. *

	1	2	3	4	5	
Stimmt gar nicht zu	○	○	○	○	○	Stimmt zu

Wie alt sind Sie? (Bitte nur in Zahlen) *

Meine Antwort

Was machen Sie beruflich ? *

Meine Antwort

Haben Sie Erfahrungen im Bereich IT-Security oder Mobile Security? *

○ Nein

○ Ja

Was ist hre Geschlecht? *

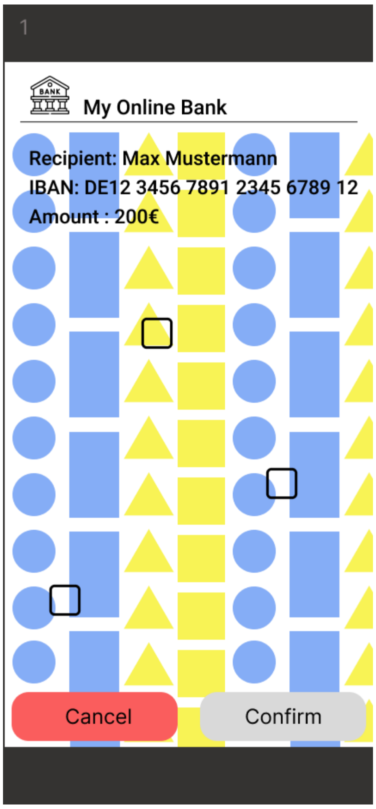
○ Mänlich

○ Weiblich

○ Divers

○ keine Angabe

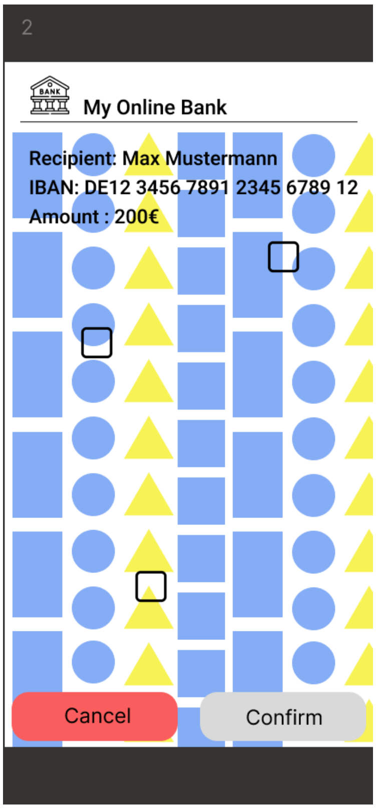
A.8 Prototype slides 1-4



1

My Online Bank

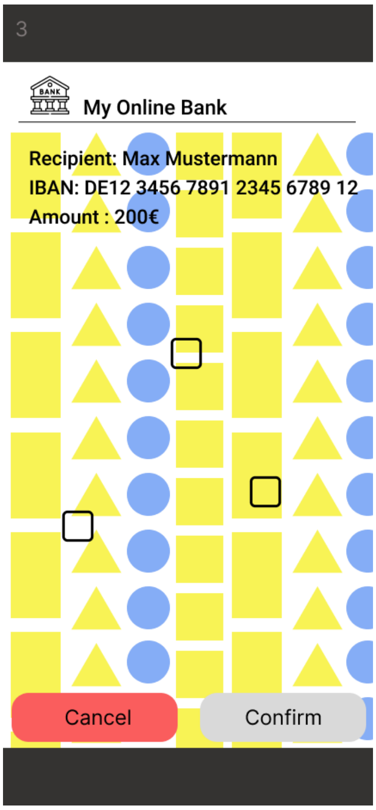
Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€



2

My Online Bank

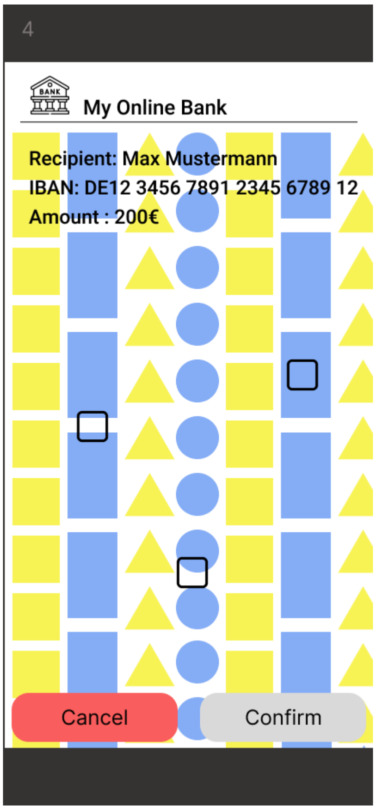
Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€



3

My Online Bank

Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€

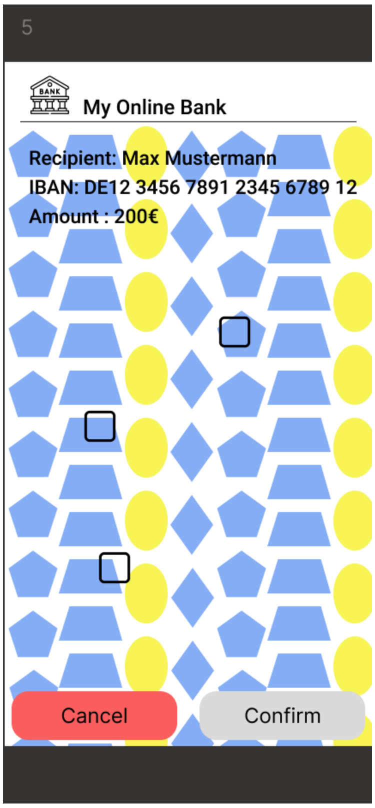


4

My Online Bank

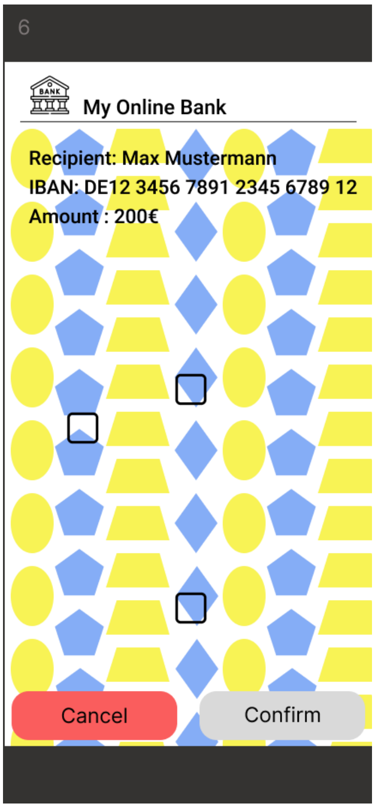
Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€

A.9 Prototype slides 5-8



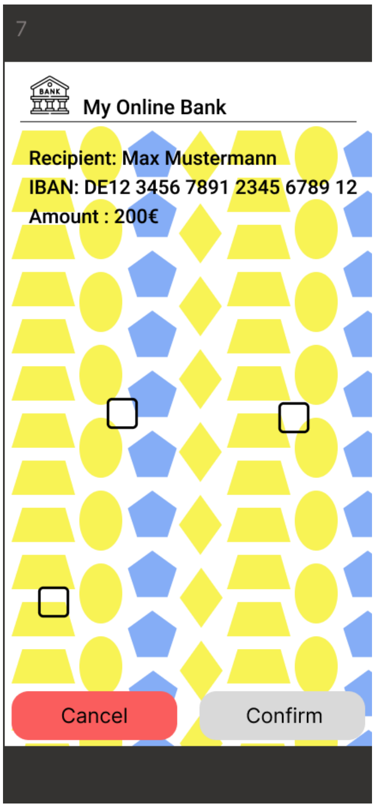
My Online Bank

Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€



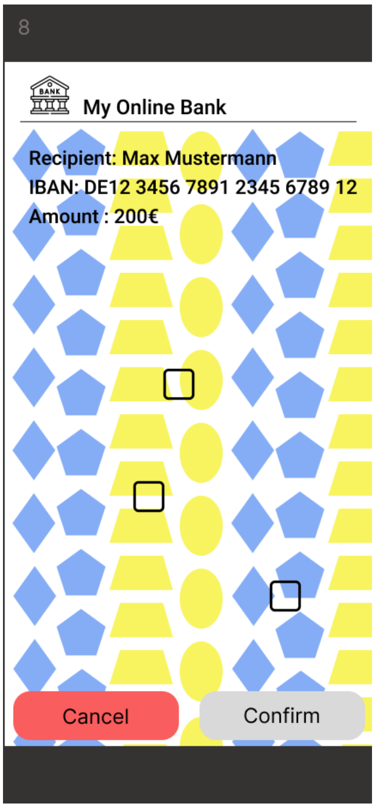
My Online Bank

Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€



My Online Bank

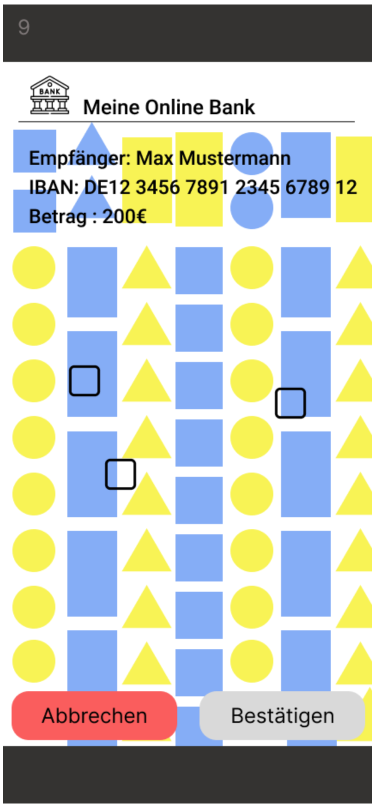
Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€



My Online Bank

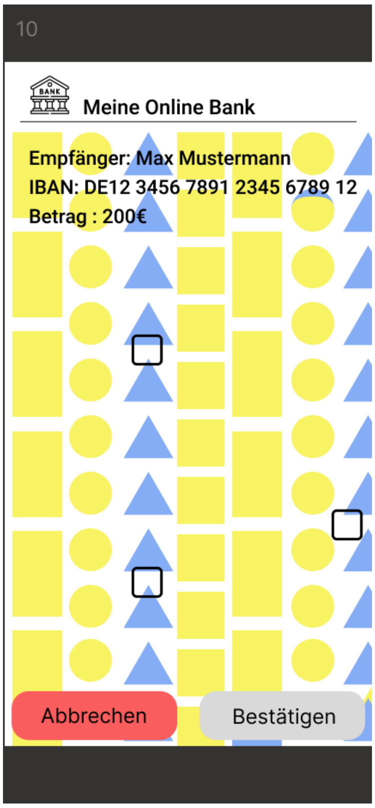
Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€

A.10 Prototype slides 9-12



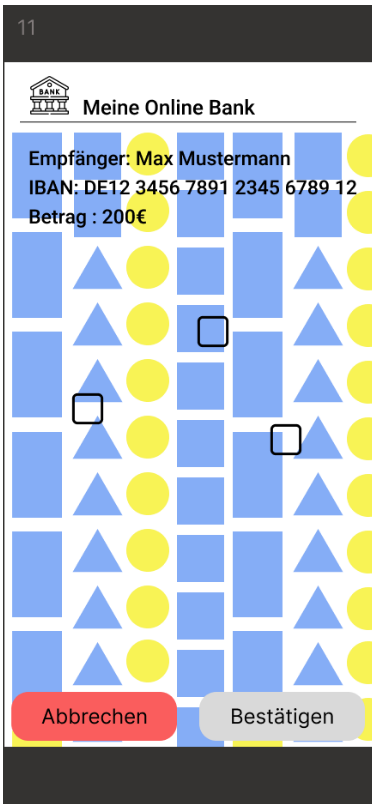
Meine Online Bank

Empfänger: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Betrag : 200€



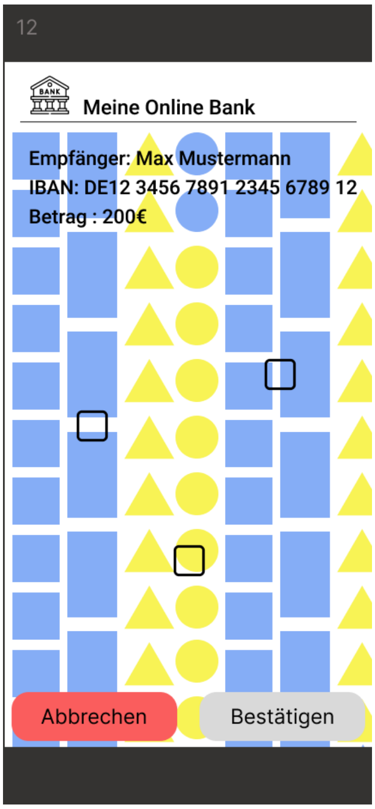
Meine Online Bank

Empfänger: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Betrag : 200€



Meine Online Bank

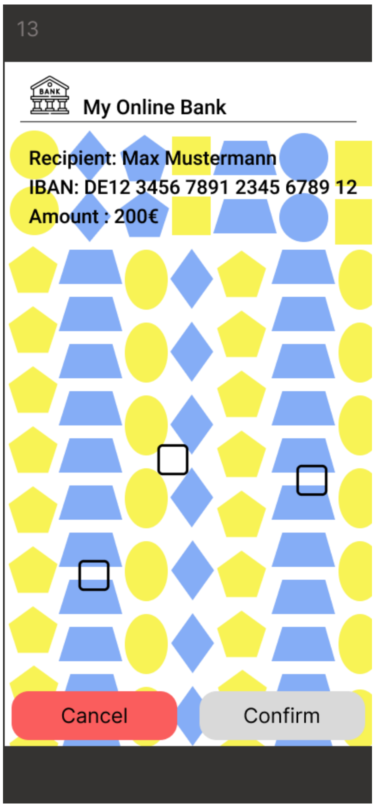
Empfänger: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Betrag : 200€



Meine Online Bank

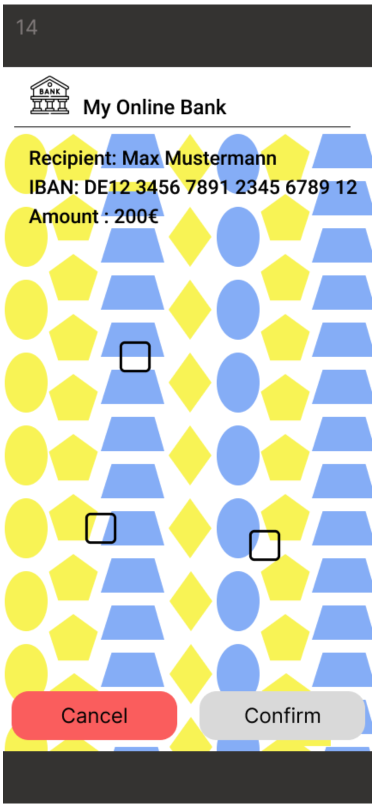
Empfänger: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Betrag : 200€

A.11 Prototype slides 13-16



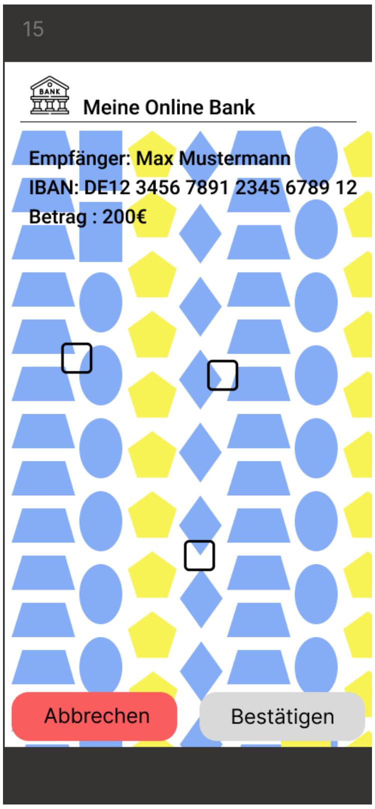
My Online Bank

Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€



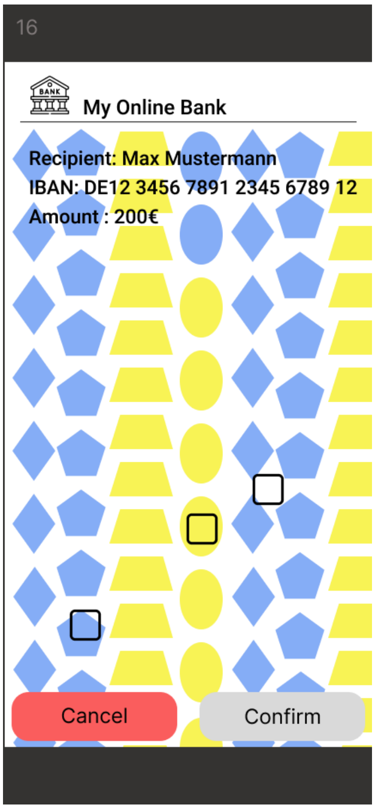
My Online Bank

Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€



Meine Online Bank

Empfänger: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Betrag : 200€



My Online Bank

Recipient: Max Mustermann
IBAN: DE12 3456 7891 2345 6789 12
Amount : 200€