Influence of User Perception, Security Needs, and Social Factors on Device Pairing Method Choices

Iulia Ion
Inst. for Pervasive Computing
ETH Zurich
8092 Zurich, Switzerland
iulia.ion@inf.ethz.ch

Marc Langheinrich
Faculty of Informatics
University of Lugano (USI)
6904 Lugano, Switzerland
langheinrich@acm.org

Ponnurangam Kumaraguru
IIIT Delhi
New Delhi, India
pk@iiitd.ac.in

Srdjan Čapkun
Department of Computer Science
ETH Zurich
8092 Zurich, Switzerland
capkuns@inf.ethz.ch

ABSTRACT

Recent years have seen a proliferation of secure device pairing methods that try to improve both the usability and security of today's de-facto standard – PIN-based authentication. Evaluating such improvements is difficult. Most comparative laboratory studies have so far mainly focused on completeness, trying to find the single best method among the dozens of proposed approaches – one that is both rated the most usable by test subjects, and which provides the most robust security guarantees. This search for the "best" pairing method, however, fails to take into account the variety of situations in which such pairing protocols may be used in real life. The comparative study reported here, therefore, explicitly situates pairing tasks in a number of more realistic situations. Our results indicate that people do not always use the easiest or most popular method – they instead prefer different methods in different situations, based on the sensitivity of data involved, their time constraints, and the social conventions appropriate for a particular place and setting. Our study also provides qualitative data on factors influencing the perceived security of a particular method, the users' mental models surrounding security of a method, and their security needs.

General Terms

Human Factors, Security.

Keywords

Device Pairing, User Studies, Authentication, Security, Usability, Social Factors.

1. INTRODUCTION

With the increasing proliferation of mobile devices – mobile phones, PDAs, netbooks, and tablet PCs – the need to spontaneously connect two devices over a wireless link has become prominent. Apart from exchanging business cards and appointments, spontanous wireless links can be used to send files to Bluetooth-enabled printers and to make electronic payments in busses, train stations, and coffee shops. To authenticate spontaneous wireless device communication, several secure device pairing protocols have been proposed that allow device authentication in the absence of a centralized security infrastructure. With no wires to verify actual connections, users cannot be sure what device they connected their wireless link to. The basic approach of spontaneous pairing protocols is thus the use of "out-of-band" channels, i.e., a secondary information channel that can be used to verify the authenticity of the primary wireless link. An example for such an out-of-band channel is the popular Bluetooth pairing method of displaying a 6-8 digit number on one device and having the user enter it on the other [5]. Here, the user's eyes and fingers act as a secondary communication channel between the two devices. Consequently, the usability of such methods is of crucial importance, as complex mechanisms might raise the probability of human error, prompt users to choose a lower security level, or lead them to abandon security altogether.

Existing usability studies that tried to compare the pairing methods proposed so far [10, 12, 13, 14, 15] mostly focussed on covering a high number of protocols and protocol variants. Consequently, prior work rarely investigated the use of such methods in real-life situations, but instead used a single generic setting (e.g., "pair these two devices") to determine the best method overall – regardless of the purpose of connecting those devices and the physical and social situation. Furthermore, prior studies predominantly recruited male study participants – mostly university students and researchers, often with technical backgrounds. Last but not least, the higher number of trial methods in these studies implied a significant cognitive load for participants, resulting in some 30-50 individual pairing tasks that each test subject had to go through in a single study [10, 15].

We decided instead to conduct a more explorative study, in order to determine the usability of proposed pairing methods *in specific situations*, and to elicit the needs and the un-

Symposium on Usable Privacy and Security (SOUPS) 2010, July 14–16, 2010, Redmond, WA, USA.

derlying mental models of users with respect to their security considerations in device pairing scenarios. We explicitly recruited a much more balanced set of test subjects, with users with diverse non-technical backgrounds. To limit the cognitive load of participants, we restricted the study to four carefully chosen device pairing protocols, aiming to span a wide range of channels (visual, audio, tactile), and degrees of user involvement (from completely passive to fully active). After having learned those four methods, participants were asked to choose among them in the context of three distinct pairing tasks, each one with a different real-world situation as a motivation. Last but not least, while previous studies only investigated device authentication, we also incorporated device *identification* into the pairing process, as choosing the device to connect to is often the most frequent and time consuming part of the process.

Our results show that device pairing methods are more than just means of connecting two devices: devices and methods used represent *people*, may make owners seem more professional (e.g., in a newly established business relationship), provide a playful moment between friends, or even act as an "ice breaker" when meeting someone new. The proper pairing method can reassure device owners that they handled a payment transaction well, or that they acted responsibly with customer data. Even more, methods evoke strong emotions: they are "annoying", drive users "crazy", and even make users "fall in love" with them.

2. RELATED WORK

The last few years saw a number of comprehensive studies that evaluated many of the proposed device pairing methods [10, 12, 13, 14, 15].[1] Our work differs from these studies in three important points: (1) we explored user preferences not in terms of pure pairing speed but by investigating particular situations and their corresponding social factors; (2) we reduced mental load on participants by testing only four representative pairing methods; (3) we recruited participants with diverse, non-technical backgrounds, and aimed for a more balanced gender composition.

Early comparative usability studies such as Suomalainen et al. [22], Valkonen et al. [24], and Uzun et al. [23] involved only simple methods based on string/number entry or comparison. The main emphasis was on measuring completion time and determining the error rate of methods. Qualitative data was not gathered, and the task given to participants was a generic pairing task that did not model any real-world situations.

In one of the most comprehensive studies, Kumar et al. [13] tested 14 variations of 8 basic methods, resulting in almost 50 individual test cases that each participant had to perform. Participants were mostly "technology-savvy" university students, with 70% male participants. While the authors argued that "if highly-motivated and technology-savvy young people do not react well to a given method, the same method will perform a lot worse with average users," our results suggest that non-technical participants do like newer methods, which performed less well in their study. Perhaps non-technical users are more excited about "what technology can do" or perhaps methods shunned by the technology-savvy fit better into their mental models of how security is provided.

In "Serial hook-ups", by Kobsa et al. [12], participants were told to imagine that they had just bought a new phone and needed to pair it to the old one. Study participants had to try 11 diverse methods, based on video, audio channel, button presses, and manual comparison. The authors proposed three "best" methods, based on the availability of displays: PIN-comparison or image-comparison for devices with a display, and (automatic and semi-automatic) audio-based comparison for devices without a display [12]. The study does not give insight into *why* users thought that a particular method would be more secure than another.

Kainda et al. [10] tested 14 methods, but placed a stronger emphasis on the trade-off between usability of a method and its susceptability to security failures. While users also preferred numeric comparison methods for their usability, the authors point out that numeric PIN entry, which requires the user to enter a number displayed on the screen into the partner device, is much less prone to accidentially confirming non-matching numbers and thus should be preferred, even if it ranks lower. The study did provide participants with a scenario – making an electronic payment to another device – yet it did not explore how this influenced the participant's choice of method. According to the usability rating used, using the phone's built-in camera to take a photo of a barcode displayed on the other device was classified as unusable. It is unclear whether this was simply a result of the low reliability of the employed 2D barcode decoder. Our results suffer from a similar bias, but probably to a lesser extent. Irrespective of the occasional unreliability of our 1D decoder, in our study, the barcode-based method was considered more secure than other methods and was thus relatively popular in payment scenarios.

A technical report by Kumar et al. [15] specifically explores scenarios involving two users. Their results show that people are unwilling to hand over their phone to strangers. This work confirms our belief that pairing methods must be explored in more realistic social settings.

All of the studies discussed above have focused only on *authenticating* the connection. They do not consider the additional step of *device identification*, i.e., pairing in the presence of other (potentially pairable) devices. Having to make a choice between several available devices significantly affects the pairing process – both in terms of time spent and the perceived security of the process. Our study, therefore, incorporates both identification and authentication methods.

Rashind and Quigley [18] compared five methods for device identification: shaking or bumping devices, simultaneous button pressing, "stitching", and touching both devices at the same time. Even though the focus of their study had not been on security, users raised privacy concerns and worried about the risks of undesired intrusions. The authors used storyboards to show to the participants different usage scenarios and found that both the purpose of pairing and the social context were important to users when choosing a method. This is very much in line with our own findings, though Rashind and Quigley did not explore the actual impact of these factors, nor users' perception of the security level of a method.

[1]Note that due to space constraints, we will not summarize the background of device pairing research here. Instead, we refer the interested reader to the excellent summaries in Kainda et al. [10] and Kobsa et al. [12].

3. METHODOLOGY

Designing proper usability studies that ensure a fair and comprehensive comparison of device pairing methods is a challenging task. First, the designer has to consider a large number of methods that have been proposed, the situations in which they apply and the type of devices they were intended for. The mental load on the participants should be considered; researchers have to carefully set the number of methods and options such that the user can learn and evaluate them appropriately. For this reason, we restricted the number of test methods to four methods that span a wide range of auxiliary channels and interaction models. Second, there are currently no consistent implementations for all these methods. Software development frameworks for mobile devices are still far behind those for desktop systems, and many methods require special libraries that are not robust or freely available (e.g., barcode decoder). We developed mock-ups of the four chosen methods to ensure consistency and reliability. Finally, the nature of wireless communications makes device pairing techniques intrinsically different from standard Internet security solutions and therefore hard to grasp even for technical people. We therefore trained the users by placing the protocols in adequate real-life situations, but kept them simple enough for users to understand. Furthermore, we paid special attention to ensure that users did not receive more training than they are likely to be given in real life.

In this study, we explore which security levels users prefer in given situations, when they are willing to use security, and how much time and effort they want to spend on pairing. We therefore designed each of the four methods – *Select the device* with PIN entry, *Take a picture*, *Listen up*, and *Push the button* – to run under three security levels: *Not secure*, *Secure*, and *Very secure*. The *Not secure* level is equivalent to running only device discovery or device identification, without authenticating the device or securing the communication. The *Secure* and *Very secure* levels imply both identification and authentication and differ in the amount of information transmitted over the auxiliary channel. Each level builds on the previous one, takes slightly longer time to complete, and typically requires an increased user effort.

3.1 Selected Methods

The four methods offer different automation degrees, by involving the user to varying extents in the connection process, and span a wide range of channels (visual, audio, tactile), and degrees of user involvement (from completely passive to very active). In all considered methods two communication channels are used: a primary and an auxiliary (out-of-band) communication channel. Here, the auxiliary channel is an authentic (typically low-throughput) channel that allows the exchange of Short Authenticated Strings between the devices. The methods primarily differ in the way they implement auxiliary channels. Their security depends on the size of the authentication string [7, 8, 16]; it has been shown [25] that, given appropriate protocol constructs, the use of short (20 bit) strings is sufficient to provide strong security guarantees. In the following we describe each of the methods and how they implement different security levels.

Select the Device is based on the Bluetooth Simple Device Pairing protocol [5], and entails device selection from a list and a PIN entry; this method is a de facto standard for device pairing today. If the user selects an incorrect device, he will connect to an unintended party; if he types in the incorrect PIN, the connection will fail. For the *Not secure* level, the user's device searches during four seconds for available devices and displays them in a list. The user chooses from the list the name of the device to which he or she wants to connect. For the level *Secure* the user's device additionally displays a 6 digit PIN (equivalent to 20 bits of data) and for *Very secure* a 9 digit PIN. The user types the PIN into the partner device to which he or she wants to pair. This method differs from the others because switching to a secure level involves adding a new kind of interaction (typing in the PIN vs. selecting from a list). For the other three methods the interaction type remains the same, but the completion time and number of (repetitive) tasks the user has to perform increase with the security levels.

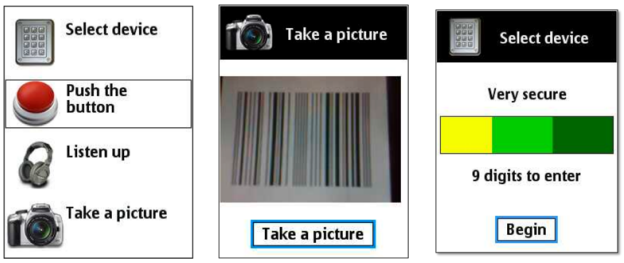
Take a Picture is based on Seeing-is-Believing by McCune et al. [17], namely using the phone camera to take a picture of a barcode displayed by the partner device. Assuming the user does not accidently take a picture of another barcode and that the barcode is successfully recognized, connecting to an unintended party is not possible. Identifying devices through barcode pictures is a well established procedure used even in systems for physical access control [4]. For the *Not secure* level the barcode contains the 48 byte Bluetooth MAC address of the device. For *Secure* the user must take a picture of an additional barcode displayed by the device and for *Very secure* the user takes three barcode pictures. The additional barcodes encode the authentication string. The security guarantees of this method vary depending on the encoding capacity of the barcode. Depending on display and camera capabilities, the length of the authentication string that could be transmitted through a single barcode picture, i.e. in the *Not secure* level of this method, could be equivalent to that transmitted in the *Very secure* level of other methods, e.g., the 9 digit PIN of *Select the Device*. However, we wanted to nevertheless introduce a differentiation between the three levels of *Take a picture* in order to quantify users' security needs as well as time and effort they are willing to spend to achieve different security levels.

Listen up is based on Loud-and-Clear [9] and the newer HAPADEP [21], and uses the audio channel for data transmission. It has the highest degree of automation among all the methods considered in our study, and places very little strain on the user. For *Not secure* the partner device plays a 3 seconds melody, which encodes its Bluetooth MAC. The user's device records, decodes and extracts the MAC, and establishes the connection. It is hard to estimate how many bits could be encoded in a 3 seconds audio transmission, but even in the very likely case in which the entire MAC address does not fit, transmitting the first or last 12 bytes and then matching these against the devices discovered or supplementing it with wireless messages would still provide a reliable enough implementation. For *Secure* and *Very secure* the melody lasts 6 and 9 seconds. The additional seconds are used to transfer authentication strings.

Push the Button is inspired by Button Enabled Device Authentication by Soriente [20]. The user's device makes short vibrations. With every vibration, the user pushes a button on the other device. The *Not secure* level requires 3 button presses, and with each of them, messages are broadcast, either by both devices or by one. Received packets

	Not Secure	Secure	Very Secure	Units
Select Device	0	6	9	digits of PIN
Take Picture	1	2	3	barcode pics
Listen Up	3	6	9	sec. of melody
Push Button	3	6	9	button presses

Table 1: Each of the four methods has three security levels, which correspond to different completion times and degrees of user involvement.



(a) Choose method (b) Take a picture (c) Security level

Figure 1: Application screenshots. (a) choosing a method, (b) taking a picture of the 1D barcode displayed by the other device (c) choosing the *Very secure* level for the *Select the device* method, which entails entering a 9 digit PIN.

are matched against the time intervals at which the button was pressed. If several connections are being established at the same time and in the same place, interference might occur. Following a similar concept, the Bump application [1] requires users to bump their phones in order to exchange phone numbers. A central server matches the bumps and facilitates the exchange. The apparent matching reliability of this application, which matches packets at a global rather than local scale is encouraging for our design. Currently, Bluetooth does not support message broadcast, but it is reasonable to assume that in the near future wireless spontaneous communications will be broadcast enabled (e.g., through the upcoming Wi-Fi Direct standard [26]). An alternative would be a WLAN infrastructure to which both devices are connected. Similar interaction concepts were proposed in SyncTab [19], Network-in-a-box [3], and WiFi Setup, and are available in several products on the market. For the *Secure* level the user must perform 6 button presses and for *Very secure*, 9. The additional presses are used for transmitting the authentication string, which could provide 9 and 18 bits of entropy (if we assume that the interval between two presses can be used to transmit 3 bits, as in the original paper [20]). This method requires increased user attention and is time consuming due to the low information entropy of the channel. Table 1 summarizes the options for security levels of each method.

We implemented mock-ups of the four chosen protocols in Python for Symbian S60. Instead of a 2D barcode we used an 1D barcode and the BaToo decoding library [2], because of the higher reliability of the 1D barcode decoder compared to existent 2D decoding libraries. To provide a realistic user experience, we preferred to use an actual barcode decoder instead of a pure mock-up. For Listen up we used an audio file sample from the original HAPADEP implementation. For Push the button we allowed a 500ms user reaction time (the time the user has to push the button on the other device once the first one vibrates). This is higher than the 300ms proposed by the original authors and is meant to minimize failure rates. The two devices used in the study were a Nokia N95 (the user's device) and a Nokia N96 phone (the partner device). Figure 1(a) shows the application screen for choosing one of the four methods, 1(b) for taking a picture of the 1D barcode, and 1(c) for using the *Very secure* level on the *Select the device* method.

3.2 Tasks

During the study, participants were given three hypothetical situations and asked which method (and which security level) they would choose. The moderator read the task description from the study script. In the following, we present the task descriptions the participants received.

Task 1: Print a document. *Imagine that you work for a consulting company. You are at the airport and will soon board the plane. You will fly to London to visit your client. You have saved your client's confidential financial report on your mobile phone. In the waiting area there is a printer. Connect your mobile phone to the printer, so you can send the financial report wirelessly to the printer. Pretend the display of the other device represents the printer's display.*

The goal of this task was to see how security-critical users perceived the document to be and how big they perceived the security threat to be. Subsequently, we asked participants questions to understand which criteria they used in their choice and whether the nature of the environment influenced this.

Task 2: Make a payment. *In London, you will also visit a good friend. Before you board the plane, you want to buy him a bottle of whisky in the duty-free shop. You hear the announcement that your flight's boarding has just begun. Connect your mobile phone to the payment terminal to pay for the bottle.*

In this task we tried to evaluate whether users perceive a higher security threat when paying compared to printing in the previous task, the effect of the time pressure on their choice, and whether they are generally more concerned about protecting private than business data.

Task 3: Send electronic business cards. *You are now in London, at your client's site. At a conference, you meet the CEO of another company, who is interested in doing business with you. You want to exchange electronic business cards with him. Use your mobile phone and connect wirelessly to his phone.*

The goal of this task was to investigate whether users differentiate between different sensitivity levels of data and type of environment, and whether the business nature of the setting influences their choice.

Table 2 summarizes the three tasks in terms of the devices participants had to connect, the data that was to be transmitted, the place where the connection was hypothetically performed, and the amount of time pressure participants were theoretically facing.

3.3 Session Structure

Sessions lasted between 50 and 110 minutes with an average of 70 minutes and a standard deviation of 15 minutes. They involved one participant at a time and were run by one

	Devices to connect	Data to send	Place	Time pressure
Task 1	Phone with printer	Confidential financial report	Airport lounge	Some
Task 2	Phone with payment terminal	Credit card information	Duty-free shop	High
Task 3	Phone with CEO's phone	Business card	Business event	Low/None

Table 2: The three chosen tasks simulate diverse real-world situations.

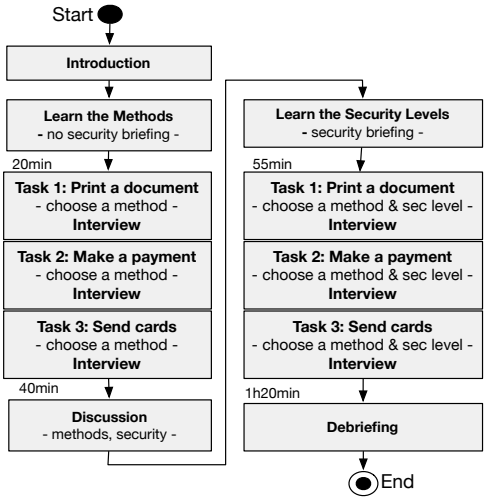


Figure 2: In the first part of the study we did not mention security. Participants learned the four methods as in the *Not secure* variant and performed the tasks. In the second part they had to choose both a method and a security level for the tasks. Typical time taken to reach the point of the study is presented above the rectangles.

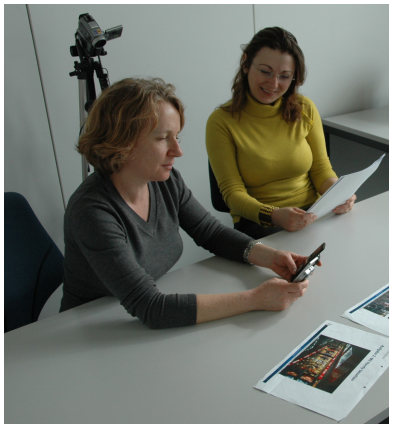


Figure 3: Study session: The moderator (on the right) reads instructions and task description from the script. The participants (on the left) pairs the two devices. A video camera is recording the session.

moderator. Figure 2 shows the outline of a session. In the "Introduction" phase, users were asked to fill in the background questionnaire. To ensure no bias was being created, in each session we used a script to introduce the purpose of the study and explain the methods. We recorded each session using a video camera placed behind the participant (see Figure 3). We took an additional audio recording with a laptop.

To motivate the study, we told participants that the methods they would learn could, for instance, be used to send a friend some pictures while sitting together in a restaurant. There could easily be dozens of mobile phones in the restaurant, so the role of the methods is to ensure that the pictures will not arrive at a neighboring table by mistake. If users knew that the purpose of the study was security related, their behavior might have changed. Therefore, during the first stage of the study (the left side of the Figure 2), we did not mention security. During the "Learn the Methods" phase participants were introduced to the *Not secure* variant of each method. The name of the level and the existence of different security levels were not mentioned. To avoid bias, we introduced the methods in pseudo-random order, overall covering all 24 possible permutations.

To teach the methods, the moderator read step by step detailed instructions and waited for the participant to execute each one before moving on. This process simulates the user buying a new phone with usage instructions for the methods or, probably more realistically, running the methods once in the shop under the salesperson's guidance. Results from Kainda et al. [11] suggest that users rarely read instructions when using a new system, and wait until they they cannot get something done. A minimum amount of explanation was given on how the methods work. For example, for *Take a picture* and *Listen up* participants were told that the barcode and the melody contain messages which their phone decodes. For *Push the button*, we said the two devices synchronize each other through the button presses. If the participant failed to execute a method, the moderator would start over again until the participant felt comfortable with the method. Our pilot studies showed that it is important for participants to successfully execute each method on their own until they succeed, otherwise they will avoid it throughout the study and consider it too hard. Finally, participants were asked to run all the methods again by themselves. The learning phase took between 15 and 40 minutes.

The moderator then read the task description and asked the participant to choose the method he or she would use in real life to establish the connection. As each task was presented, the participant was shown a picture of the potential situation (i.e. an airport lounge, a duty-free shop, and business people talking at an event). After each task, a small semi-structured interview followed. If peculiar answers or inconsistencies emerged, further questions were asked to explore the answers. Special care was given to ensure that participants understood the methods (within the limits of the script information) and that peculiar or incorrect beliefs emerged from users' general perceptions and security mental models, not from lack of clarity on the tasks and methods. If, during the task phase, the moderator realized that the

participant had not properly understood a method, she went back to the learning phase and explained the method again. However, to avoid bias, no further details beyond what the script contained were provided, even if the participant asked for more. At the end of all three tasks, the participant was asked to speak freely about his general impression of the methods and was asked whether he worried about security and how it influenced his choice.

The second part of the study, depicted on the right side of the Figure 2, was security oriented. In the "Learn the Security Levels" phase, the participant was told that in the way he or she had used the methods until then, anybody with the proper tools could listen in on their communication, and read and possibly modify the data transmitted. For each method, following step by step instructions, the participant learned the three security levels. Screens similar to Figure 1(c) allowed the user to select the desired security level. At the end of this phase, participants were asked to run on their own all methods with the *Secure* level and then again with the *Very secure* level.

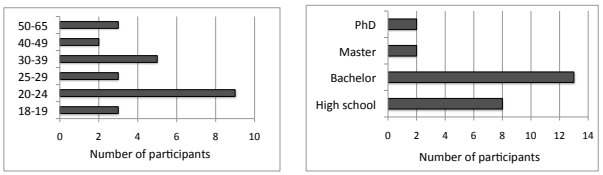
Finally, the participant was asked to think about each of the three tasks again and decide which method and which security level he or she would use. To refresh his or her memory, shorter tasks descriptions were re-read. As in the first part of the study, at the end of each task a semi-structured interview was carried out to understand the participant's choices and preferences. A larger set of questions explored the perceived threat level and how different factors, such as the data being sent and the social setting, would influence the choice. The session ended with a free discussion.

3.4 Participants

We recruited 25 participants – 10 male and 15 female – through an online job advertisement website hosted by ETH Zurich but regularly visited by people not affiliated with the university. None of the participants had studied computer science nor had taken security nor advanced computer courses. Professions and study areas varied widely, with no more than two participants from the same field, and included secretaries, housewives, veterinary medicine students, a cook, economists, a sales person, lawyers, psychology students, a journalist, etc. Figure 4 summarizes the demographics. All participants reported owning a mobile phone. Seventeen participants reported that their mobile phone was Bluetooth or WLAN capable, fourteen said that they do not use Bluetooth nor WLAN regularly on their device, seven use it once every few months or once per month, three weekly and one several times a day. Several participants said they had never used such *"advanced"* phones before and many said they *"don't know much about technology."* The study was conducted in German in our offices in Zurich and involved one participant at a time. We gave a monetary reward of 20 Swiss Francs (aprox. $17) to each participant.

3.5 Data Analysis

We transcribed all audio recordings into English. For each question in the interviews, we tried to identify trends and place answers in a few big categories. Most questions explored the preferred method and/or security level in a given situation. On a first pass, we categorized answers first based on the chosen method and second based on the reason for choosing the method. Next, we observed patterns across different questions and tasks. The perceived security and usability of the methods emerged in different places throughout the session. Finally, higher-level conclusions such as mental models, perceived security, the need for control, and the role of social context, emerged through associations and combinations of all of the above.



(a) Age (b) Education

Figure 4: (a) shows that many participants were young, but higher age groups were also represented; (b) presents the education levels equivalent to US degrees. 15 participants were female and 10 male.

4. RESULTS

In this section we present participants' choices of methods, perceived security and mental models, and we draw conclusions on influencing factors. We refer to participant 1 as P1, participant 2 as P2, and so on. We start by presenting high-level takeaways, discuss method preference for each task and decision factors, then present perceived security, mental models and the role of social factors. The results of our study are purely qualitative. We do report the number of participants who fall into a given category, but we do not imply statistical significance.

Non-technical users do like newer methods: Previous studies mostly recruited participants with technical backgrounds and concluded that users prefer simple methods like number comparison instead of the newer ones, inferring that newer methods will perform even worse with non-technical users [13]. In our study, while the most popular method was indeed *Select the device*, on average half of the people preferred another method in any given task. If designed and explained well, non-technical users do embrace non-standard methods: *"Take a picture is cool"* (P8); The methods are *"reliable, fast, uncomplicated"* (P12); *"interesting and exemplary"* (P20); *"really cool, especially Listen up, that is really great"* (P25). Given a more reliable barcode decoder and a lower number of pictures to be taken for the *Very secure* level, the *Take a picture* method is very likely to receive an even higher user acceptance.

Different users prefer different methods: We found no single method that fits all users. In terms of personal preference, opinions differ widely. Some users said *Push the button* is *"funny"* and *"cool,"* while others said it is *"silly," "annoying"* and *"cumbersome."* The most controversial methods were by far *Take a picture* and *Listen up*. Some participants excluded *Listen up* because of the sound while others thought the method is very practical and the sound would not bother them or people around. While P16 said about *Take a picture*: *"It would drive me crazy if somebody would want to do that to my phone"*, P6 has *"fallen in love with it."*

Same user prefers different methods in different situations: Three participants explicitly stated that *"in different situations different methods are applicable."* Figure 5

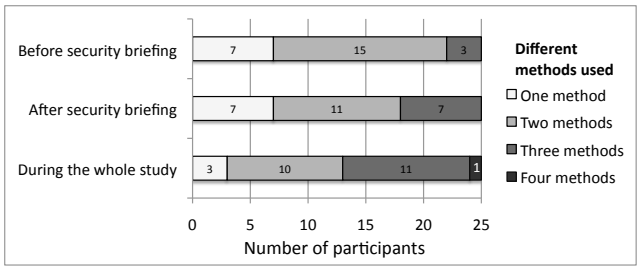
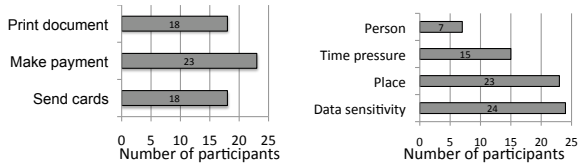


Figure 5: Participants preferred different methods in different situations. For example, only seven people chose the same method for all tasks before the security briefing.



(a) Same method again (b) Main decision factors

Figure 6: (a) For each task, users said they would use the same method again if encountering the same situation. This suggests that user's choice, although fine-grained, is not aleatory. (b) Users chose methods based on the sensitivity of data, the place where the connection is established, the time pressure they are in, and the person handling the other device.

shows the number of participants that used different methods for the three tasks in the study. For example, only seven people used the same method in all three tasks before the security briefing. In terms of security levels, five users chose the same security level for all tasks, all of whom used *Very secure*. No user chose three different security levels, which might be an indication that users are more willing to vary the method used than the security guarantee.

Same user prefers the same method in the same situation: Although very diverse and fine-grained, participants' choices for methods were not aleatory. For each task, we asked participants questions of the kind "would you use another method, if, for example, you had to print another document?" Figure 6(a) shows the answers for each task. Almost unanimously the answers were *"no, if it works, I would always use this one"* or *"once good always good."* The few answers of the kind *"yes, would use another method"* were almost always followed by a condition: *"if a less sensitive document [were to be printed]"* or *"if not in a hurry."* Only three participants said they would use an alternative method for paying, seven for printing and seven for exchanging business cards.

To show users' diverse preferences for the methods and how many factors play a role, we give the following policy example. Figure 7 depicts some of the decision rules mentioned by more users.

Example policy: P6 (male, 19 years old) chose *Select the device* for printing: *"I can see the list"* and *Take a picture* for paying because it gave him a double assurance: *"By taking a picture, I actively recognize whether this is the device I want."* P6 thinks *Listen up* is less secure than *Take a picture* because *"my device could make a mistake"* and *"I would send my payment data to somebody else. [...] If another person's device rings I cannot walk over, take his phone and say 'Sorry, I have to delete my data from your device.' He would say 'Are you crazy? What are you doing with my phone, somebody was calling me, what do you want?' I find this kind of risky, even when I hope that the mobile phone always chooses the right device."* But for exchanging business cards P6 would, nevertheless, use *Listen up*: *"Now we are at a meeting. If the CEO is there, he can see whether he received something from me. [...] It would not be so bad if somebody else received my business card because that is not something extremely personal. In this case the connection needn't be double-verified."* After the security briefing, P6 said he would use a lower security level if he were printing his own tax document because it is not so sensitive. In the office or at home, he feels *"generally safer"* because he is alone, and therefore would choose *"one security level less."* To pay he would still use *Take a picture*, with *Secure*, and to exchange business cards *Listen up*, also *Secure*.

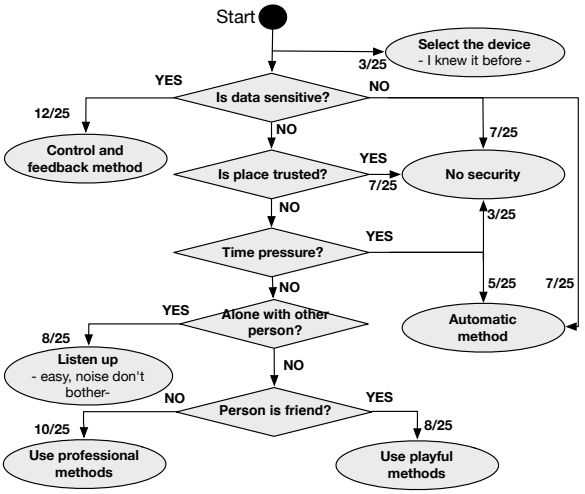


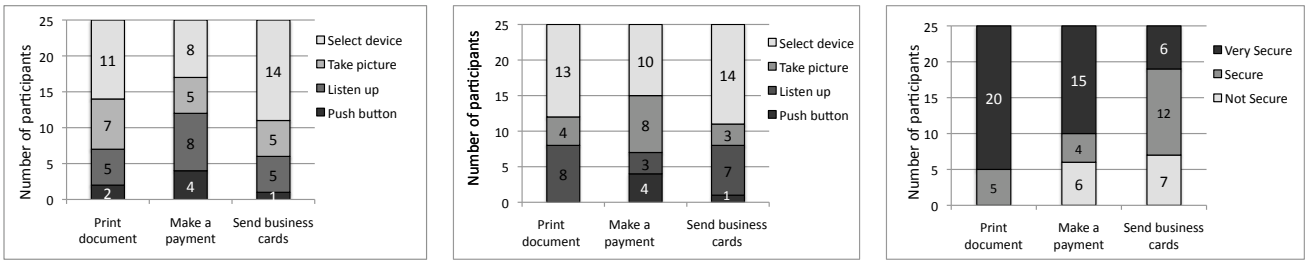
Figure 7: Participants showed fine-grained decision process based when choosing a method. For example, 7 out of the 25 users said that in a trusted place they would not use security.

In the following section, we summarize the results and impressions of the 25 participants and outline some of the main factors influencing their choices.

4.1 Preferences and Decision Factors

Previous studies tried to identify the preferred method and rate easiness of use. Our results show that users do not always use the easiest or fastest method, nor the one they like best. For example P11 said *"Push the button annoys me,"* but he would use it for printing a sensitive document *"even if I don't like it,"* because the method seems secure and it gives him a sense of control: *"There I have a direct influence on the devices, I synchronize them myself."*

Before security briefing: Figure 8(a) displays the participants' choice for methods in the first part of the study. To print, eleven people chose *Select the device*, seven *Take a picture*, five *Listen up* and two *Push the button. Listen up*



(a) Method choice before security briefing (b) Method choice after security briefing (c) Security levels choice

Figure 8: (a) displays method choice for the three tasks in the first part of the study, before security was mentioned. We then introduced the security levels. Participants had to perform the same three tasks again, choosing (b)the preferred method and (c) the preferred security level.

became more popular for paying because it was perceived as fast (users were in a hurry to catch the flight) and *Push the button* as well because, even if tedious, it was perceived as secure, which is very important when dealing with money. Only six people did not mention security as a selection criteria during the first two tasks, all of which chose *Select the device* every time. *Take a picture*, *Push the button*, and occasionally *Select the device* were regarded as secure because users felt in control. All participants who chose *Take a picture* or *Push the button* for printing or paying said they did so for security reasons. People who chose *Listen up* said they did so because it is fast and/or easy. Reasons for choosing *Select the device* were more diverse: easy, fast, somewhat secure, *"I knew it before,"* or *"I am certain it works."* For exchanging business cards, users once again tended more to *Select the device* (14 participants), which generally was considered professional and most adequate in business settings.

After security briefing: Figure 8(b) displays the preferred method and Figure 8(c) the chosen security levels, in the second part of the study. For printing a document, twenty people used *Very secure* and five *Secure*. For paying, fifteen people used *Very secure*, four *Secure* and six *Not secure*. When making a payment, the reason for lowering the security level was mostly the hypothetical time pressure in the task, while for exchanging business cards it was the low sensitivity of data. When keeping security high in task 3, some users said they wanted to seem responsible in front of the CEO, would like to maintain security by default or worried that there is always a risk (see Section 4.4).

For each task, we asked participants to sort five criteria used to choose a method in their order of importance. According to the average ratings, *security* ranked first for printing, followed by *ease of use*, *speed*, *professional look-and-feel*, and finally by *fun*. For paying, speed became the second factor, while for exchanging business cards speed and ease of use were both ranked first.

	Not secure	Secure	Very secure
Select the device	14s	29s	29s
Take a picture	24s	28s	34s
Listen up	9s	15s	17s
Push the button	20s	22s	36s

Table 3: Completion times for participant 6 in seconds, for each combination of methods and security levels.

The varying differences in completion times for security levels for the four methods was a reason for switching to another method. Table 3 depicts the completion time for P6. For example, *Listen up*, *Very secure*, took 17 seconds, only 8 seconds more than *Not secure*. *Push the button*, however, took 16 seconds more for *Very secure*, compared to *Not secure*. The least number of people, six out of twenty-five (compared to twelve and thirteen for the first two tasks), changed their chosen method for exchanging business cards after the introduction of the security levels, which might indicate the weight of social factors in this situation.

There was a tension between users' tendency to choose a default method and security level, and their tendency to adapt to various data protection requirements. Interestingly enough, these tendencies were at odds even for the same participant. After having said *"Why are there three security levels? I would always use the highest one,"* P24 nevertheless said he would use the *Secure* level (i.e., only the second highest level) for printing his own tax document: *"My tax data is not so secret. I have an average salary."*

Overall, users varied both the security level and the method used depending on a wide range of factors: the sensitivity of data being transmitted, the place where the transaction was made, the time pressure, the person operating the other device, the social setting, people present, noise level, and perceived security threat. Figure 6(b) depicts the main decisive criteria and the number of participants using them.

Data sensitivity: An overwhelming twenty-four out of the twenty-five people used sensitivity of data as criteria in their choices, e.g. when exchanging business cards or printing their own tax document. Surprisingly though, people did not only change the security levels based on the sensitivity of data, but also (and maybe more or equally often) the method used. P16 said: *"If it is about money the method has to be extremely secure, and it can also be more tedious. It is a completely different situation than before [when printing], where it can be easy, or when you have nothing to lose."* P15 would use *Listen up* to print her own tax document: *"in the worst case, it is not so bad if it goes somewhere else. [...] But for the financial report of my client, I would not want to risk that. An error could occur; I could believe that the sound comes from this device but it would not be so."*

Place: Figure 9 shows the number of people for whom place was a decisive factor. Summed up, twenty-three people used place as a selection criteria in at least one situation. For example, when printing at home instead of in the air-

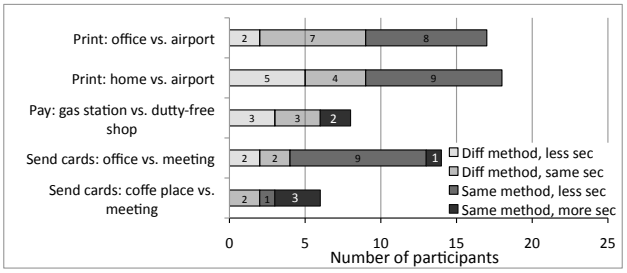


Figure 9: The place where the connection is being established is a decisive factor in choosing the preferred method.

port, nine participants would use the same method but less security. To pay at night in a gas station, when no other customers are waiting in line, eight participants would make a different choice than in the airport. If they would be exchanging business cards with the CEO in the office, in the first part of the study, nine participants said they would use a different method, eight of which opted for *Listen up*. For six participants, a coffee place requires a different choice than the conference.

Time pressure: Fifteen participants mentioned time pressure as a decisive factor. Six people said they would have used a higher security level, had they not been in a hurry to board the plane, four of which had used the *Not secure* level in the task. P20 felt under time pressure when printing in the airport and said he would use a higher security level to print in the office because he would have more time. Unsurprisingly, when in a hurry, nine participants selected a faster method and/or a lower security level. Furthermore, under stress, four participants preferred less attention demanding methods. P17 said: *"if you are under stress you are careless."* P5 worried that *"because of the rush I could not take pictures so well"* and P14 said that she *"could make a mistake when typing in the number."*

Person: Seven participants said they would choose a different method to exchange addresses with a friend than when exchanging business cards with the CEO.

In the following section, we give more insight into participants' reasoning and mental models related to security, as well as factors that increase the perceived security threat in specific places and situations.

4.2 Perceived Security

In the first part of the study, nineteen people mentioned security as a choice criterion and sixteen said they had worried about it. (Even if they used security as criteria, some people said they didn't worry about security because they were reassured by the use of an adequate method.) However, what participants worried about was not cryptographic protocols nor malicious attackers. Instead, they worried about connecting to the wrong device by mistake and about how to avoid errors. P14 said about printing: *"If I chose the right device, then I am not concerned that somebody else would get the document, even if I use no security."*

Four participants believed *Select the device* provided high security assurance because they could *"see the name"* and then they could be sure nothing bad would happen. Seven people said *Select the device* is not secure, but only one person worried that somebody might try to impersonate the printer; the other five were concerned with accidently choosing the wrong device, having more devices with the same name in the room, or that in real-life they would not know the name of the device. During the study, the name of the partner device was displayed on its screen. *Select the device* was regarded more secure in the office than in a public place by five participants, not because of a lower risk from attackers, but because *"in my own office I would definitely recognize the devices"* (P18) or *"if I have set-up the printer myself, I then know exactly which one it is. Maybe I even used it several times. I don't really feel insecure"* (P25).

Device naming was confusing for many participants, even after the learning phase. Most of these users thought the name of the device was "Nokia N96" because the model "N96" was printed on the device, above the screen. P11 tried to infer the name: *"I thought the printer is in the Lounge. That's why I chose this Lounge printer."* For the paying task, there was an accidental misspelling in the name of the device users had to choose: the other phone's display said "Du**tt**y-free A" and the user's device showed "Du**t**y-free A" in the list. Only one participant out of the fifteen that went through this screen observed the name mismatch. This confirms that people are very likely to tolerate some sort of spelling mistakes during the identification phase. Feedback and verification measures are therefore of extreme importance. For three participants, typing in the PIN was valuable as double confirmation that they indeed chose the right device.

Five participants said that automatic methods are secure, because the user cannot make a mistake. Afraid that she might select the wrong device or type in the wrong number, P20 used *Listen up* for paying. When using *Take a picture*, P10 said: *"It seems the most secure to me, with this method I think an error is not possible."*

For some users, perceived security was more important than the predefined security levels. Unaware of the role of authentication strings, P20 believed that *Take a picture*, *Not secure*, is more secure than *Listen up*, *Very secure*: *"In my opinion it doesn't change much for one or two pictures. [...] I think it is secure enough, even with one picture."* We asked participants how concerned they are about somebody seeing their credit card data during the wireless transmission, on a Likert scale from 1 to 7, where 7 is very concerned. P20 rated *Take a picture*, *Not secure*, with 2 and *Listen up*, *Very secure*, with 7. About *Push the button*, P15 said: *"how does this increase security if I press three or six times? It's no extra step, no double confirmation. What I find good there is that something new happens, there is a new aspect."* P19 said about *Listen up*: *"From the security point of view, it doesn't matter to me if it is 3 or 6 seconds."* Understanding what makes people perceive a method as secure is of crucial importance in designing systems.

Seven participants explained that a method is secure if it provides control, feedback, and double assurance, and allows operating both devices. When dealing with sensitive data, users would go through the trouble of using a tedious method as long as it fulfills these requirements. Although considered tedious, hard to perform, and slow, *Push the button* was preferred by some participants in security relevant situations because it is interactive, it seemed very precise, and provided control over both devices. *Take a picture* was also regarded as secure, because of the double confirmation and control.

Control also means the ability to cancel the connection at

any point. P15 worried about Listen up and the lack of stop and cancel functionality of the prototype: *"How can I stop this if I hear it comes from another phone? Could I stop it?"* Other participants worried they could not distinguish from which device the sound would come from. P15 said: *"I don't know if it comes from the purse, from a phone, or from the payment terminal."*

Four participants said that if they put in an extra effort for security they feel at peace. P2 said: *"If I use security I have the feeling that I did something about it, so I am less worried."* Twelve participants said they would use more security than they consider necessary to be on the safe side or that they would have security enabled by default. P12 said for printing: *"Better a bit more secure than too little."* Similarly, although P17 believed the home is more secure than a public place, she still used the highest security: *"I would set the settings to that and then change it rarely. Simply because it is set to this."*

Seven participants said that, when dealing with sensitive data, they prefer methods where they have control and double feedback, but that for less sensitive data or under time pressure, automatic methods are better. P22 said: *"When you have to pay, it is better to have feedback but it can also be more tedious. It is a completely different situation than before, where it can be easy, or when you have nothing to lose."* Similarly, P16 said: *"For everything that is not sensitive data I would immediately use Listen up. [...] I think that you need to have interaction here, to have the feeling that it is secure."* It is not enough for designers to create the most secure and the easiest method, if it does not also instigate users trust. The designer might need to introduce, even artificially if necessary, explicit steps to provide for the feeling of security.

4.3 Mental Models

Even when properly accounting for perceived security and mapping this to the actual security guarantees of the protocols, designers need to be aware of user mental models, user requirements, and their implications. For example, four people said that when handling sensitive data in a public place the method should be discrete. P19 said: *"If it is really confidential, then other people don't need to be aware that I try to setup a connection."* With *Listen up*, *"people would wonder what I do with the music, what is that. [...] If it sings then people will perhaps look, throw a look at the financial report."* When printing sensitive data, three participants chose a method because it required physically interacting with the printer. P15 said: *"It is a confidential document, and therefore I have to take some precautions, that I am next to the printer when it gets printed."*

Our interviews revealed several mismatches between people's mental models and current system designs and operations. As credit card information gets used more and more for small payments in daily life, it is crucial to convey to users the importance of properly securing *every* transaction in mobile payment applications. Alarmingly enough, six out of the twenty-five participants, even very well educated and security-concerned people, said they would use less or no security when buying a pack of cigarettes than when buying a bottle of whisky because the price is lower. Although we did not specify whether credit card information or electronic cash gets transmitted wirelessly, subsequent questions referred to credit card information. P17 even said: *"even though I assume it is the same data that is being transfered, it is less money and one thinks it is not so bad."*

Keeping users alert about security in time is a challenge that security designers should keep in mind. If nothing bad ever happens, even security-aware users are very likely to lower their guard. P22 never hands out her credit card in a restaurant: *"That is the biggest mistake,"* but even though she thinks *Listen up* is not secure enough for paying, she admits that eventually she would no longer use *Select the device* which she considers secure: *"I would be weak and select* Listen up*, because I will have gotten to trust it."* To avoid such cases, security sensitive applications like mobile payments should enforce security by default, and not give users a choice to opt out.

Three participants thought devices are predestined to fulfill specific purposes and cannot act as other types of devices. For P7, *Select the device* is not appropriate when connecting to a phone because there might be more phones with the same name, but printers are less common devices, so then the method is good. Although one of the most diligent and security concerned participants, P15 chose not to enable security for payment, because she was convinced that she cannot transfer money from her phone to the phones of people around. This is a very dangerous assumption, given that sniffers and protocol implementations are possible. Relay attacks on in-shop credit card payments have been demonstrated [6].

Two participants worried that data can be stored and reused, but only if in a non-obfuscated format. P24 believed that the *Push the button* is more secure than *Take a picture*: *"When you push [...] there are several steps that one maybe cannot as easily trace back like with a picture that one can recall.* Push the button *is more secure because you cannot trace it back."* Even compared to *Select the device "it is more discrete, more hidden in the device."* P7 said: *"For the printer I would maybe worry that the data is saved somewhere and then it could be printed out again. For paying I worry less about this."* Countless incidents of in-store credit card cloning dismiss this assumption.

When connected to an unattended device, participants generally wanted interaction and control over both devices. However, if the partner device is operated by a person, this requirement diminished because the other person could act as a feedback provider, confirming that the data arrived at the right place. P15 said: *"If I can coordinate this with the second person, I am certain that no other person can take my data."* This can be a dangerous assumption, since the protocol is just as vulnerable to eavesdropping and man-in-the-middle attacks. Since the scenarios we explored covered only less sensitive data exchange with a human party, it would be interesting to see if users' concern increases if instead of the business cards sensitive data was transmitted.

When being alone or in a trusted place, participants generally felt safer. P23 said: *"If there are so many people here, you don't feel so protected anymore."* Also, when alone, the probability of connecting to an unintended party is lower. For printing, thirteen participants said there is less risk in the office. Ten participants would not use security if printing in the office instead of in the airport. Extending the concept to paying at night alone in a gas station could have undesired consequences in the presence of eavesdropping, unattended devices. P14 said that in the gas station he would not enable security for paying: *"I would be sure I pay to the intended*

party, because there is nobody else around."

Participants often refrained from using a method because they had not understood it or felt it didn't *"make sense."* We witnessed first-hand the value of explanation and education, which constitutes a big challenge in the real world when introducing new applications. Five people preferred a method in a given situation because it seemed *"appropriate,"* or it resembled something they knew, e.g., debit card payments or passwords. Some participants had even more surprising criteria: P17 chose *Take a picture* to exchange business cards *"because business cards are more visual. And it goes better with something optical."* If instead she was exchanging an mp3 file, she would have used *Listen up*, and to exchange a financial report with the CEO, she would use *Select the device.*

4.4 Social Factors

Our results show that designers should pay careful attention to ensure the methods comply with social conventions, otherwise users might compromise security for social compliance. For example, P16 said: *"Listen up would be more secure, but it draws more attention than it should."* Furthermore, lowering the security level in the office is not necessarily because of lower risk: *"It is more quiet, and if we are both there, it would feel awkward if it rings too long."* Social factors influence requirements for interaction models, ease of use, speed, and security and were cited by twenty users as a reason in their choices. Fourteen people said they would be embarrassed to use one of the methods in a social setting: ten with *Push the button*, three with *Listen up*, two *Take a picture*, and one if typing in the PIN for *Select the device.*

Participants said the method used is critical for building a good business relationship. Eight people said they would use a different method with a friend than with the CEO. All of them decided to change from *Select the device* to other methods if pairing with a friend's phone. P5 said: *"If it is somebody that I know then I would either use* Take a picture*, or put the phones together and use* Listen Up*, because I am closer to him, he is not such a big boss."* P7 said: *"When I know the other person well, I think any method would be appropriate. When it is somebody important or whom I do not know, I would take the most professional method: Select the device with typing in the PIN."*

P16 used *Push the button* to pay for the feeling of control but used *Listen up* with the CEO: *"The other methods would be too personal, if I now have to press around on his phone. It would be silly to have to tell him he has to press the button 3 times when my phone vibrates, or if I would have to push the button on his phone, because it is too personal, too close. I want to build a good relationship with him. If you don't know a person too well you don't want to go like a bull at a gate.* Take a picture *is just as inappropriate. [...] When it is about a business contact, I would like it to be the easiest for him, and for the situation: the method that could least go wrong."* On the other hand, *"if it is friends or acquaintances or my parents or whatever, then I don't care. They know me and I know them, so it doesn't matter which method I use."*

Depending on the social context, even the speed requirements of the method vary. When establishing a new personal contact, e.g., in a business relationship, the method should be faster than normal, easier and not disruptive. For exchanging business cards, P21 used a lower security level: *"It would not be the most pleasant when establishing personal contact to spend such time in this technique [Take a picture, Very secure]. So it should work relatively fast. Additionally I don't have to concentrate very much and I can nevertheless continue engaged in conversation with the discussion partner while I establish the connection."*

To make a good impression and protect the CEO's data, users sometimes seemed even to exaggerate the security requirements. P9 chose *Listen up* and *Secure* to pay, but *Select the device* and highest security to exchange business cards: *"Out of respect towards the CEO. I wouldn't want his data to arrive to somebody else but me."* P8 also used *Very secure*: *"It shows that you worry about the data security of somebody else, which could further strengthen the business relationship."* In fact, even though it was just about business cards, only seven people disabled security in task 3.

A funny anecdote was provided by P24, who would use *Take a picture* to exchange business cards. At the conference the *Secure* level is enough, because business cards are not so important, but in the office he would use *Very secure*:*"In the office the CEO is next to me and maybe he sees that I use the highest security. He probably expects that. At the conference there are also other people. He is more attentive when there are no other people around. And he sees that I choose highest security."* In the coffee place he would again use the middle security level: *"The CEO sits on the other side of the table. He doesn't necessarily see this."*

We asked participants to rate the sensitivity of the data contained on the business cards. P9 said there is a difference between his business cards and the CEO's and rated the CEO's with 5-6 on the Likert scale (7 is the highest) and his own with only 3-4. *"Maybe he has his private address written there, which nobody should have."* P13 would rate the CEO's cards as extremely sensitive, 7 on the Likert scale, if his private number would be on them. She used the highest security level for exchanging the cards: *"I hope that the CEO does not give everybody his business cards, but just to me."*

Participants used higher security when dealing with somebody else's data to make a good impression, but also out of a sense of responsibility towards other people's data. Eleven participants said that business or confidential data is more important than private data or were extremely concerned with protecting the CEO's business card. It would be interesting to conduct a cross-culture study, to explore whether people have different behavior in different countries.

Twelve people said they would use a lower security level and maybe even a "less secure" method to print their own tax document instead of the financial report. Fourteen participants said the tax report is less important: *"My tax document is mine, private, but what concerns the company does not belong to me. That I do for the company. So I have more responsibility"* (P14). P17 said *"I think it has less priority because it is something personal, and if it is a customer's, a business contract, you have to be twice as careful."* P16 also thinks *"tax document data is no longer so sensitive as the financial report, because the financial report concerns other people too, while the tax document just me. So it has more consequences, because I would damage other people too, if I were not secure."* P21 said: *"I am accountable in front of the CEO when I handle his confidential data. So I take the highest security."* For P23, losing money would be comparable to losing trust if she handled data irresponsibly. Eight participants said the financial report is more important even than the payment.

Several participants said that, in a more relaxed environment or among friends, the methods can provide a playful moment. P16 said: *" It depends if the other person knows it already, but for example, if he doesn't know these methods, I would have the demo effect with* Take a picture*: 'Look, it works!' "* P19 thinks that even with the CEO the method *"maybe plays a role to establish contact. If we have a bit of fun together, it will remain in his memory."* At the conference he would use *Select the device*, but in the office *Take a picture*, and even *Push the button* could be appropriate: *"Maybe it even has something that connects us, an ice breaker."* With a friend he would normally use *Take a picture* because *"it is more intimate"* and joked about how he would maybe even use *Push the button "to annoy somebody, like my grandmother, because she cannot do it."*

Finally, the right method is very dependent on the social situation. When printing a document in the airport, P16 thinks that *"it would be totally ridiculous if I wanted to take a picture [of the printer]. Even* Push the button *is a bit foolish. The PIN is professional."* P10 said: *"Noise is a criteria. In the meeting I cannot use Listen up."* P17 also wants a silent method: *"At a dinner you meet and talk to people; vibrations and sounds are not appropriate."* P9 agrees: *"When there are other people present, I think it is better to be discreet. If I am alone with the CEO then I would use Listen up, otherwise Select the device."* But in the airport *Listen up* is appropriate, and he also chooses the highest security: *"The airport is always noisy, so the music wouldn't bother."*

In the office seven users switched to *Listen up* when printing. P18 said *"If I'm alone the sound can't come from somewhere else."* However, only two people switched to *Listen up* in the office for exchanging business cards. This might be due to the higher weight of the social factors: choosing a method that would be appropriate for interacting with the CEO. Different decision factors have different priority for different people.

5. DISCUSSION

We conducted a laboratory user study with 25 participants to investigate the usability of device pairing methods in different real-life situations and the security needs users perceive. We tested four methods that span a wide range of auxiliary channels (visual, audio, tactile) and require different levels of user involvement (from very active to fully passive). Our results show that users do worry about security, but not in terms of malicious attackers or data encryption. It is not protection against man-in-the-middle attacks or eavesdroppers that makes a method secure in their perception. Instead, a method is perceived to be secure if it reassures users through double confirmation and control that everything went as planned and they indeed connected to the intended device.

Users prefer different methods in different situations. For example, when dealing with sensitive data, control and feedback are needed, and when handling less sensitive data or under time pressure, automatic methods are preferred. Furthermore, social factors greatly influence method requirements. For example, when connecting to a friend's phone the method can be playful, but with a newly met person in a business environment professionality is required. Similarly, in the office, at home, in a public place, or in a meeting, different methods and security levels are desired. We investigated factors influencing users' method choices in different real-life situations and detailed their perceived security and mental models.

Laboratory studies cannot predict with full confidence real-world behavior. In the three tasks, we tried to depict real-life situations as clearly as possible and hope that provided answers are consistent with real-world behavior. Strong correlation between reported previous security concerns during real-life situations and participants' security concerns during the study might be a reassuring fact. Future work should investigate user behavior in the wild, over a longer period of time. For example, an initial study could deploy device paring methods among the participants at a one-week conference and see which methods people use for exchanging business cards.

Method preferences were influenced to some extent by the reliability of the software and mock-ups used, a pitfall that any user study will encounter. For example, sometimes the barcode decoding library did not focus on the first try. This made some users believe that *Take a picture* is not as reliable as *Listen up*, which always worked due to the mock-up nature of the prototype. Nevertheless, *Take a picture* seems to have awakened the most enthusiasm in users. It would be interesting for future work to provide better understanding of to what extent even small unreliability in a method will make users avoid it.

Were we to conduct the study again, we would refrain from using the term "PIN" in the description of the *Select the device* method. Three users associated the number with the PIN of credit and debit cards, two of which were confused when, the number was displayed on the partner device's screen. One of these users indicated that for printing, unlike for paying, she would like to see the number displayed on the screen. It would be interesting to see if users maintain this association for payment tasks but not for others in the absence of the "PIN" naming.

Future work should also test whether users have a lesser requirement for control over the process if connecting to a device operated by a person (e.g., the CEO's phone) than to an unattended printer or payment terminal, regardless of the sensitivity of the data being exchanged. Furthermore, we would like to test our results and conclusions against other pairing methods and see if control and confirmation are still the major factor in user security perception or if methods such as distance bounding protocols intrinsically inspire more trust. One participant mentioned during the study that when putting the phones together he automatically feels safer.

6. GUIDELINES FOR DEVELOPERS

Creating a technically secure and highly usable method is not always sufficient to meet users' needs. The method should also comply with users' security perception and be appropriate for the specific social situation.

1. **Map perceived security to method guarantees:** Designers should create methods whose actual security guarantees are consistent with users' perceived security. To achieve this, it might be necessary to introduce redundant steps, controls, cancel buttons, and double confirmations.

2. **Include security by default:** We detected several mismatches between users' mental models and system designs, which prove the need to include security by

default when dealing with sensitive data, such as a customer entrusting a confidential financial report or a bank issuing a credit card. Also, our results show users' willingness to have security enabled by default.

3. **Support several methods:** Some users liked *Take a picture* very much and disliked *Listen up*, and others felt exactly the opposite. To account for diverse personal preferences, mobile devices should support a set of different pairing methods.

4. **Account for social factors:** No single method is adequate for all situations. Users are likely to bypass security before breaking social norms. Designers should provide appropriate methods for professional environments, public and private places, and interaction with friends or strangers. The user could, for instance, choose between several variants: meeting mode, quiet room mode, professional mode, play/fun mode, etc.

Acknowledgements

We would like to thank Lorrie Cranor for guidance in planning the pilot studies, Jonathan McCune for discussion on methods and security levels design, and Christina Pöpper for verifying the accuracy of interview translations. Last but not least, we thank Yves Geissbühler for the initial implementation of the Python software used in the study.

7. REFERENCES

[1] Bump. http://bu.mp/.

[2] R. Adelmann, M. Langheinrich, and C. Floerkemeier. Toolkit for bar code recognition and resolving on camera phones – jump starting the internet of things. In *GI Jahrestagung (2)*, pages 366–373, 2006.

[3] D. Balfanz, G. Durfee, R. E. Grinter, D. K. Smetters, and P. Stewart. Network-in-a-box: how to set up a secure wireless network in under a minute. In *SSYM'04: Proceedings of the 13th conference on USENIX Security Symposium*, pages 15–15, Berkeley, CA, USA, 2004. USENIX Association.

[4] L. Bauer, L. F. Cranor, M. K. Reiter, and K. Vaniea. Lessons learned from the deployment of a smartphone-based access-control system. In *SOUPS '07: Proceedings of the 3rd Symposium on Usable Privacy and Security*, pages 64–75, July 2007.

[5] Bluetooth SIG. Bluetooth Special Interest Group. Simple Pairing Whitepaper (Revision V10r00), 2006.

[6] S. Drimer and S. J. Murdoch. Keep your enemies close: distance bounding against smartcard relay attacks. In *SS'07: Proceedings of 16th USENIX Security Symposium on USENIX Security Symposium*, pages 1–16, Berkeley, CA, USA, 2007. USENIX Association.

[7] C. Gehrmann and K. Nyberg. Enhancements to Bluetooth baseband security. In *Proceedings of Nordsec 2001*, 2001.

[8] C. Gehrmann and K. Nyberg. Manual authentication for wireless devices. *RSA Cryptobytes*, 7:2004, 2004.

[9] M. T. Goodrich, M. Sirivianos, J. Solis, G. Tsudik, and E. Uzun. Loud and clear: Human-verifiable authentication based on audio. In *ICDCS '06: Proceedings of the 26th IEEE International Conference on Distributed Computing Systems*, page 10, Washington, DC, USA, 2006. IEEE Computer Society.

[10] R. Kainda, I. Flechais, and A. W. Roscoe. Usability and security of out-of-band channels in secure device pairing protocols. In *SOUPS*, 2009.

[11] R. Kainda, I. Flechais, and A. W. Roscoe. Two heads are better than one: Security and usability of device associations in group scenarios. In *SOUPS*, 2010.

[12] A. Kobsa, R. Sonawalla, G. Tsudik, E. Uzun, and Y. Wang. Serial hook-ups: a comparative usability study of secure device pairing methods. In *SOUPS*, 2009.

[13] A. Kumar, N. Saxena, G. Tsudik, and E. Uzun. Caveat emptor: A comparative study of secure device pairing methods. In *PerCom*, pages 1–10, 2009.

[14] A. Kumar, N. Saxena, G. Tsudik, and E. Uzun. A comparative study of secure device pairing methods. *Pervasive Mob. Comput.*, 5(6):734–749, 2009.

[15] A. Kumar, N. Saxena, and E. Uzun. Alice meets bob: A comparative usability study of wireless device pairing methods for a "two-user" setting. *CoRR*, abs/0907.4743, 2009.

[16] S. Laur and K. Nyberg. Efficient mutual data authentication using manually authenticated strings. In *CANS*, pages 90–107, 2006.

[17] J. M. McCune, A. Perrig, and M. K. Reiter. Seeing-is-believing: Using camera phones for human-verifiable authentication. In *Proc. IEEE Symp. on Security and Privacy*, pages 110–124, 2005.

[18] U. Rashid and A. J. Quigley. Interaction techniques for binding smartphones: A desirability evaluation. In *Human Centered Design – First International Conference, HCD 2009, Held as Part of HCI International 2009, San Diego, CA, USA, July 19-24, 2009 Proceedings*, pages 120–128, 2009.

[19] J. Rekimoto. Synctap: synchronous user operation for spontaneous network connection. *Personal Ubiquitous Comput.*, 8(2):126–134, 2004.

[20] C. Soriente, G. Tsudik, and E. Uzun. BEDA: Button-enabled device pairing. In *Proc. IWSSI 2007*, pages 443–449, September 2007.

[21] C. Soriente, G. Tsudik, and E. Uzun. HAPADEP: Human asisted pure audio device pairing. Cryptology ePrint Archive, Report 2007/093, March 2007.

[22] J. Suomalainen, J. Valkonen, and N. Asokan. Security associations in personal networks: A comparative analysis. In *Proc. ESAS 2007*, pages 43–57. Springer-Verlag, 2007.

[23] E. Uzun, K. Karvonen, and N. Asokan. Usability analysis of secure pairing methods. In *Proc. USEC 2007: Usable Security*, February 2007.

[24] J. Valkonen, A. Toivonen, and K. Karvonen. Usability testing for secure device pairing in home networks. In *Proc. IWSSI 2007*, pages 457–462, September 2007.

[25] M. Čagalj, S. Čapkun, and J.-P. Hubaux. Key agreement in peer-to-peer wireless networks. *Proceedings of the IEEE (Special Issue on Cryptography and Security)*, 94(2):467–478, Feb. 2006.

[26] Wi-fi. Wi-fi alliance announces groundbreaking specification to support direct wi-fi connections between devices. http://www.wi-fi.org/news_articles.php?f=media_news&news_id=909, October 14, 2009.