# Towards Understanding ATM Security – A Field Study of Real World ATM Use

Alexander De Luca[1], Marc Langheinrich[2], Heinrich Hussmann[1]
[1]Media Informatics Group, University of Munich, Amalienstr. 17, 80333 Munich, Germany
{alexander.de.luca, hussmann}@ifi.lmu.de
[2]Faculty of Informatics, University of Lugano, Via G. Buffi 13, 6904 Lugano, Switzerland
langheinrich@acm.org

## ABSTRACT

With the increase of automated teller machine (ATM) frauds, new authentication mechanisms are developed to overcome security problems of personal identification numbers (PIN). Those mechanisms are usually judged on speed, security, and memorability in comparison with traditional PIN entry systems. It remains unclear, however, what appropriate values for PIN-based ATM authentication actually are. We conducted a field study and two smaller follow-up studies on real-world ATM use, in order to provide both a better understanding of PIN-based ATM authentication, and on how alternative authentication methods can be compared and evaluated. Our results show that there is a big influence of contextual factors on security and performance in PIN-based ATM use. Such factors include distractions, physical hindrance, trust relationships, and memorability. From these findings, we draw several implications for the design of alternative ATM authentication systems, such as resilience to distraction and social compatibility.

## Categories and Subject Descriptors

H.1.2 [**Models and Principles**]: User/Machine Systems – Human Factors; K.6.5 [**Management of Computing and Information Systems**]: Security and Protection – Authentication

## General Terms

Experimentation, Security, Human Factors

## Keywords

ATM, security, authentication, design implications, field study, lessons learned

## 1. INTRODUCTION

New authentication systems are mostly created with the goal to be "better" than PIN or password (e.g. [4, 9]). "Better" usually refers to being more memorable, more secure, or both. Security is certainly the most important aspect when designing authentication systems for public settings (e.g. ATMs), yet memorability directly affects security as well, as hard to memorize secrets get written down and thus overall security suffers [1].

The standard approach to verify the appropriateness of a new ATM authentication system is to compare it to PIN entry in controlled laboratory experiments. However, such a laboratory experiment can never mirror completely the real situation when using an ATM. The role of the authentication process with respect to the entire interaction at an ATM remains unclear, since the actual process of ATM authentication outside of laboratory settings has not been sufficiently examined yet. For example, overall interaction speed is a very important aspect of public authentication, and it has been argued that alternative authentication mechanisms should thus also be judged by this factor (e.g., [4, 17]). PIN entry typically is faster than proposed alternatives, yet without knowing the "big picture" of an entire ATM interaction, it is difficult to assess the significance of this faster speed.

Previous research [13], based on semi-structured interviews, helped to identify basic factors that influence the decision to use an ATM, like privacy, social density, and time pressure. Nevertheless, the actual use of ATMs was not explored. Consequently, we decided to perform a number of field observations involving ATM use, in order to explore how people actually interacted with ATMs. As it has been previously shown in the domain of public display interactions [10, 15], field studies have the potential to uncover important facts and practices that otherwise cannot be asserted. The main focus of our observations was on the ATM authentication process, i.e., how people enter their PIN, whether and how people protect their PIN entry from skimming attacks, and what contextual factors affect security and secure behavior.

After analyzing the first field study, two additional follow-up studies were conducted: A second field observation with the focus on obtaining more detailed interaction times, and an additional set of interviews in public spaces in order to ground some of our findings.

This paper presents the results of the two field observations and the interviews, and derives a number of implications for the design and the evaluation of authentication mechanisms for ATMs. For example, our observations indicate that contextual factors have a high influence on security and usability of PIN authentication. A large number of observed interactions (11%) featured one or more distractions during ATM use (e.g., phone calls, discussion with friends, or handling shopping bags). Maybe not surprisingly, we also found that a majority of users (65%) did not take any precautions against PIN skimming attacks (such as shielding


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

PIN entry). Based on our findings, we offer a discussion of lessons learned for performing field studies on the use of privacy sensitive technology.

## 2. METHODOLOGY

The field observations were performed in six different locations in two central European cities, Munich (Germany) and Delft (the Netherlands). We chose ATMs that were available 24 hours a day, seven days a week, and which were located outside. This allowed for unobtrusively observing actual ATM interactions (see below for a description of the observation method).

The data for the primary field observation was collected over a period of nearly two months. Each ATM was at least visited four times, with at least one observation session on a Sunday and at least one session during "rush hour" (i.e., mid-mornings, noon, or early evenings). This was to ensure that the data collected was as broad as possible and did not, e.g., only include off-peak times, which could have biased the results. Rush hours and off-peak times were identified in pre-observations. Depending on the location (e.g. close to a supermarket) these times differed not only between cities, but also between locations within the cities. For instance, the rush hour close to a supermarket was between 5pm to 7pm while the rush hour at an ATM in a pedestrian area with shops and restaurants was during lunch time (around 1pm).

We also made sure to observe a variety of ATMs from different banks (six banks in total), since terminal software can significantly differ from one bank to another. At each ATM, 60 users were observed, resulting in an overall data set of 360 users, which were collected during 44 observation sessions. 199 of the observed users were male, 161 female.

All observations were performed and recorded by the one and the same researcher. This was necessary to keep the data comparable, since different people might apply different standards during the observation, deliberately or not. Even though multiple observers might have reduced the risk of accidentally missing data, we opted for this solution since we considered consistency more important than efficiency (speed of collecting the data).

In order to remain unobtrusive during observations, we chose ATMs that were visible from public outdoor seating areas, i.e., street cafés and restaurants that had tables in appropriate locations outside. A large number of the outdoor ATMs that we could find were actually close to such spots. Thus, finding appropriate locations was not an issue. Considering these precautions, it is very unlikely that the observer did arouse suspicion amongst ATM users. Additionally, the observation sessions were kept rather short to minimize this risk.

### 2.1 Ethical and Legal Considerations

In order to ensure the privacy of the study subjects, we chose all of our observations spots in such a way that the hands of the subject could be seen but the keypad itself was not visible. Also, we positioned ourselves at a distance where the ATM screen could not be read. Most importantly, all observations are based on written data by the observer – no surveillance technology of any kind was used, i.e., neither videos nor photos were made.

We instead used a written checklist in order to ensure that no important information was missing. This list was based on procedures identified during an informal pre-study. The checklist included the following information:

- location
- gender
- time of day
- interaction time
- queue length behind user
- security measures
- start of interaction
- repeated PIN entry (yes or no)
- comments

In the first field study, *interaction time* was simply measured with a standard commercial stop watch. The beginning of the measurement was the moment of inserting the bank card, the time was stopped when the user took the withdrawn money (all our observed interactions resulted in a money withdrawal). We later performed a more detailed analysis of interaction times in a follow-up study (see section 2.4 below). The entry *security measures* featured a number of checkboxes for marking procedures that had been identified in the pre-study, such as "hiding entry with other hand" or "checking people standing close to the ATM". Finally, situational information that could not be narrowed down to a set of actions was written down in the *comments* section of the checklist (e.g., "with company", "on the phone" or "shopping bags").

To ensure untainted data, observations were only added to the data set if all of the above points could be collected with 100% confidence by the observer. The reasons for failed observations were mainly cars or other people that suddenly blocked the view to the ATM or the user. Roughly one third of all observations were thus discarded. There were some rare instances of interesting behavior (e.g. a user leaving the ATM after a failed authentication attempt) that lead to failed observations – these were also not added to the data set, but instead written down as additional comments in case they would help to gain further insights.

In the countries where we conducted the studies, no ethical review boards are in place for this kind of research. However, legal issues have to be considered. For instance, German privacy regulations state that without the explicit consent from the subjects, data can only be collected and stored anonymously.[1] However, once data has been rendered anonymous, it can then be used freely for scientific purposes. Since none of our subjects can be identified by any means (no videos and photos were taken), our data collection is truly anonymous. Furthermore, as the study was conducted in public spaces without the use of AV-equipment, our local legal counsel informed us that no consent from any institution (e.g., banks or city administration) was required. In connection with the previously mentioned measures to protect the subjects' privacy (e.g., not being able to see the actual PIN entered), we thus did not identify any legal or ethical issues with this study.

During the observation sessions, no frauds or safety issues came up. However, if this would have occurred, the observer would have of course abandoned the experiment and provided help/support as needed.

[1]Exceptions do exist of course, e.g., for law enforcement or the protection of private property.

## 2.2 Methodology Limitations

Since ATM interaction is a sensitive and private task, it was very important for us not to disturb the users' privacy. Therefore, we decided *not* to engage them in interviews after the observation. Consequently, some of our findings are necessarily based on (speculative) reasoning about the observed behavior, rather than on actual user feedback. Especially inferences on the use of security, the influence of company, and queuing strategies were not verified with those users exhibiting these behaviors. To fill these gaps, we performed additional interviews in public spaces with a focus on these aspects (cf. section 2.3 below).

When analyzing the observational data from our first study – and especially the comments – it became apparent that the time measured from entering the ATM card to the moment of money withdrawal was not entirely sufficient. Many users blocked the ATM for a significantly longer amount of time before and after the actual cash withdrawal, which we called *preparation phase* and *cleanup phase*, respectively. These phases include simple tasks like getting the ATM card from the wallet or putting down shopping bags. Based on our experiences from the first study, we reckoned that this overhead might in some cases be around 50% to 100% to the "interaction times" that we measured. To clarify this issue, we performed a second set of observations (cf. section 2.4) with a focus on input times.

## 2.3 Follow-Up: Public Interviews

To get a better understanding on users' security considerations, the influence of company, and users' queuing behavior, we conducted a number of public interviews some time after our initial field study. Interviews took place over a period of one day in the city center of Lugano (in the Italian speaking part of Switzerland). As we did not want to interview people who we had previously observed withdrawing money (cf. section 2.1 above), we do not think that the change of location for these interviews affects our findings. Also note that these interviews did not attempt to achieve statistical significance – we merely wanted to gain some insight into "people's thinking" with respect to ATM usage. While there might clearly be cultural differences between ATM users in Munich, Delft, and Lugano, we expect to be able to uncover the same basic set of attitudes in each of these locations (though we do not have evidence for this assumption).

Overall, 25 full interviews were conducted. That is, 25 participants answered all questions. Additionally, two interview partners did not use ATMs and thus were not asked any additional questions. The average age of the survey participant was 36 years. The youngest was 19 and the (two) oldest 64 years old. One participant did not agree to share his birth year. 16 participants were male, nine were female.

Two interviewers performed the interviews together. They were fluent in English, German, and Italian, and thus were able to cover a large range of possible interview partners. While we did not record nationality, all interviewees were in fact fluent in at least on of those three languages. People were semi-randomly picked. "Semi" refers to the fact that the interviewers tried to get people from as many different age groups as possible. Firstly, people were asked whether they would be available for a short interview. They were told that the interview was for a research project of the local university and that no private data of them would be collected. Approximately 30% of the approached people did not agree to participate in the interview.

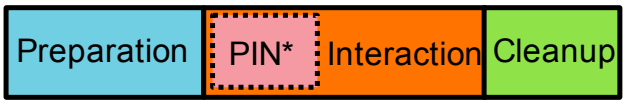


**Figure 1: The different phases measured for the in-depth time study. **PIN was not measured.***

The first question was about whether the interviewee actually used ATMs or not. Out of 27 interviewees, only two stated that they did not use ATMs at all. One person explained that "*I don't trust those machines, so I don't use them.*" For participants who said that they used ATMs, we continued with the following questions:

- Approximately how many times per week do you use ATMs?
- Do you worry that someone might steal your PIN when using an ATM?
- How do you protect your PIN entry?
- If you are in company, would you still protect your PIN? (If no, why not?)
- If there is a queue at the ATM, would you wait in line? On what does your decision depend?
- What is your alternative to queuing at an ATM?

Participants were told that they could answer those questions freely. We did not interrupt them as long as they felt like talking. During that period, the interviewers took notes to record the answers – again, no recording devices other than pen and paper were used. After each interview, participants were given a small reward (a piece of chocolate) for answering the questions. The final question asked whether they were willing to provide us with their birth year. All but one participant gave us this information.

## 2.4 Follow-Up: In-Depth Time Measures

To get a better understanding of the time overhead that is spent at ATMs besides our previously measured "interaction time", we conducted a follow-up field study in Munich. In contrast to the first study, we used a custom program installed on an Android-based smartphone to easily record the individual interaction phases. Meant as a supportive study to gain insight on the influence of preparation and cleanup on the overall interaction time, this study featured a significantly smaller amount of only 24 observations. The in-depth measurements were performed on two ATMs that were also used in the primary field study. At each ATM, twelve data sets were collected in four observation sessions.

Three different times were measured (see figure 1):

- *Preparation*: time from blocking the ATM to the beginning of the actual interaction (i.e., the previous begin of our "interaction time", when the user entered the card)
- *Interaction*: time from card entry until cash/receipt withdrawal (previously called "interaction time")
- *Cleanup*: the additional time the ATM was blocked by the customer after the last withdrawal

Splitting an interaction up into several consecutive steps can help to identify usability factors and to uncover different

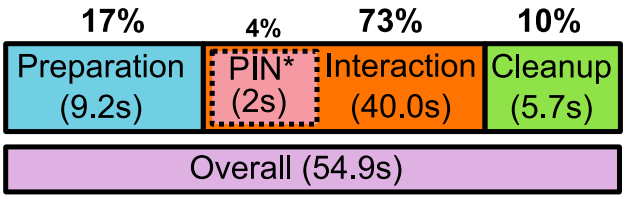


**Figure 2: The different phases including their average times. **PIN is a subset of interaction and is based on related work.***

effects that might have stayed hidden otherwise. This has for instance been done by Bauer et al. [2] when they analyzed the usability of the Grey authentication system.

Apart from the use of a smartphone to acquire more in-depth recordings of times, the same methodology and ethical rules were applied for this study as they were for the initial observations. Thus, due to the private nature of the observations, we could not record even more detailed breakdowns of the interaction time, in particular the time spent on entering the PIN. This would have required us to observe the actual ATM screen, which we tried to avoid for ethical reasons. Based on previous work, however, we know that PIN entry is very fast and usually takes around two seconds only (e.g. [7, 5]). We tacitly assumed similar timings for our observations.

## 3. FINDINGS

This section presents findings based on the two field studies and the field interviews, grouped along five main properties: overall interaction time, user distractions, input errors, queuing behavior, and employed security measures.

### 3.1 Interaction Time

In the main field study, an interaction session took on average 45.9 seconds (SD: 15.1s). The fastest user was finished in only 19.9 seconds while the longest took 125.3 seconds. Sessions were typically measured from the moment the user inserted the card until the cash or the receipt (if any) was taken. As we pointed out above, our observation positions did not allow us to isolate authentication times (i.e., PIN entry) in these measurements – taking PIN entry measurements from prior work [7, 5] (2 seconds) these would thus be less than 10% of the total average interaction time that we observed.

A detailed analysis of the data revealed that factors like queues and the use of security measures did not significantly influence interaction time. For instance people hiding their PIN entry (mean: 45.9s) did not take significantly longer than users that did not perform such security measures (mean: 44.4s).

However, during our observations we noticed that the actual interaction with the ATM was only part of the time that a single user would block the machine. Significant overhead came from "preparation" and "cleanup" actions taking place before and after actual ATM use, respectively. These actions included: arranging shopping bags; finding the bank card; putting the withdrawn money into the wallet; arranging personal items (e.g., putting away wallet); and finishing a phone call or a conversation with a friend.

These times were measured in our follow-up study described in section 2.4 above. Our in-depth measurement later showed that preparation and cleanup actions would take around 27% of the time that the ATM was blocked (17% preparation, 10% cleanup). In one extreme case, preparation and cleanup made up even 66% – for a user that arrived with a dog and a child in a pram. Before he could use the ATM he had to make sure they were safe (e.g. blocking the weels of the pram), which he later had to undo again during the cleanup phase. In the "best" case, they took only 16% of the time that the ATM was blocked. This was a user that performed a strategy that we could observe four times during our observations: he had his cash card already prepared when he approached the ATM, rendering the "preparation" time practically zero.

The average time for preparation (9.2s) was higher than for cleanup (5.7s). The different phases, their average times, and percentages are depicted in figure 2. The most important thing to notice here is that standard PIN authentication takes only a fraction of the overall time that a user is in front of an ATM.

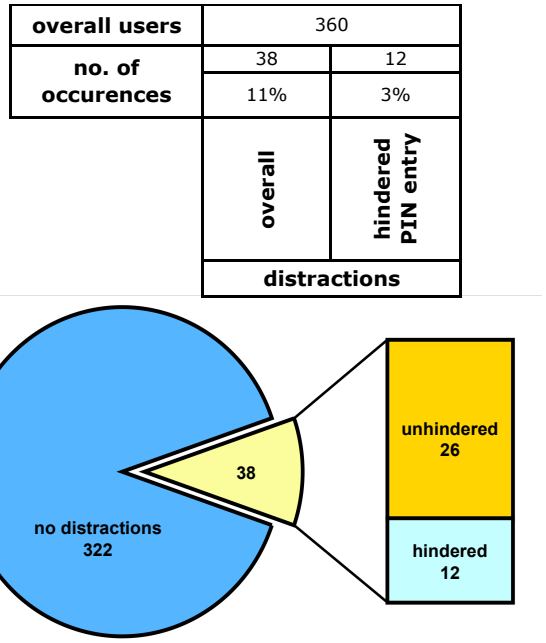


**Figure 3: 11% of users were distracted during PIN entry, for 3% this even hindered PIN entry and led to errors.**

### 3.2 Distractions

Our initial observations revealed a number of factors that distracted the user during the actual ATM interaction, i.e., they either interrupted the interaction with the ATM or slowed down the preparation or cleanup phase. One of the most common distractions was a friend or partner that spoke to the user during the interaction. Other factors were for instance shopping bags or prams that partially required the continuous attention of the user. Overall, 38 people (11%) were distracted by various factors during their ATM use (see figure 3).

In an extreme case, a user came to the ATM with a dog and his child in a pram. Before he could even think about

starting the interaction with the ATM, he had to take care of both, effectively blocking the ATM in the process. Also, during the interaction the child repeatedly required attention, resulting in a loss of focus on the actual task.

## 3.3 Input Errors

ATMs give users three tries to authenticate to the system. In case the user fails to do so, the bank card will typically be confiscated by the machine. While the distance to the ATMs did allow the observer to see only the general interaction with the keypad, but not the actual PIN input, we distinguished errors from successful input in the following way: All ATMs in this study used screen keys for providing access to their different services (e.g., account balance, withdrawal). To activate any ATM functionality after a successful authentication, the user *had* to use one of the screen keys. That is, the hand had to be moved away from the keypad to the screen. Going directly back to keypad input without touching any of the screen keys thus meant that the user was forced to correct the PIN. In some of the cases, users even removed the card after an error occurred and restarted the authentication process all over.

Out of the 360 users we observed in the initial field study, only six failed to authenticate correctly at the first attempt. These six users subsequently spent more time ensuring that they would "get it right" on their second attempt. The average time for an interaction that included a failed authentication session was 103.1 seconds – more than twice the average time of a session without a failed authentication. However, due to the small amount of errors, this difference is not statistically significant. The low error rate correlates with standard PIN entry error rates from laboratory studies (e.g. [7, 16, 17]).

We observed one user who first applied security measures but failed to authenticate correctly: shielding her PIN entry with the other hand meant that she could not see which buttons she was pressing. After her first attempt failed, she gave up on shielding her PIN entry and then was able to enter the PIN correctly.

## 3.4 Queuing Behavior

If an alternative authentication method takes longer than PIN entry, one might expect this to have an effect on accumulated waiting times. If authentication took, say, twice as long, would queues in front of ATMs get much longer? During our observations, we were thus interested in actual queuing behavior: how long do ATM queues typically get, and how do people deal with long queues, both while waiting and while withdrawing?

Big queues almost never occurred during our observation sessions[2]. In 251 of the 360 sessions, no one was queuing behind the user. Queues with a length of one appeared 88 times; queues with a length of two 19 times. We only observed two instances when the queue had three or more people: one time three people where queuing, once we saw four people in line. At a length of two, we saw people approaching the ATM but when they realized there was already a queue they seemed to change their mind and turned to go away.

To get a better understanding of this behavior, and to understand reasons for and against queuing, we included two corresponding questions in our follow-up questionnair study. When asked if they would queue in front of an ATM, three of the 25 interviewed participants stated that they would never queue. Four users said that they always queue, no matter how long the queue. The remaining 18 participants stated that it would depend on the circumstances. When analyzing the interview logs, we identified four such influencing factors: Urgency, queue length, the availability of an alternative, and the perceived safety of the queue. We will briefly describe these factors in turn below. Note that Little et al. [13] also identified *time pressure* as an important factor toward ATM use. However, this was only mentioned by one of the 25 participants. We assume that our way of phrasing our question in our interview did play a role in this notable absence from the list of factors.

### 3.4.1 Urgency

11 out of 25 interviewees were only willing to queue if they urgently needed cash.

### 3.4.2 Queue length

Six participants explicitly mentioned an acceptable queue length. None of them said they would accept a queue bigger than three. One user said that he would only queue if it was urgent, and only if the queue length would be two at maximum.

### 3.4.3 Alternatives

The most important factor for our participants when deciding on queuing was the availability of alternatives – not only the alternative of having another ATM close-by, but also other means.

14 participants stated that they would only go to another ATM if a) the alternative ATM would not apply charges, and b) if it would be located close-by. Two participant mentioned that they would always queue due to the lack of alternatives: both were with banks that had very few ATMs in town from which they could withdraw money without being charged. Four users stated that a queue would make them skip cash withdrawal altogether, given that they were on their way to shop at a place that supported paying by card (e.g., a local supermarket).

### 3.4.4 Perceived safety

One participant had a different view on ATM queuing than the rest of the interviewees. Instead of considering it a time burden to queue, she instead considered the safety aspects of the queue. Depending on the type of people in line, she stated that she would not queue "*if there are strange people nearby*".

## 3.5 Observable Security Measures

During the main observations, we found that only 124 out of 360 users (around 35%) made observable efforts to secure their PIN entry (57 female, 67 male). A summary of secured and unsecured input is depicted in figure 4.

The most common security measure was hiding the PIN entry with the second hand or the wallet (120 out of 124). Many ATM interfaces propose this method when prompting for PIN entry (see figure 5). Four out of the six ATMs in our study displayed such a hint. Interestingly, users at such ATMs were not more likely to protect their PIN entry.

The remaining four users that applied security measures

[2]In the following, we count only people in line to use the ATM, not the people accompanying them.

| | secured | | | unsecured | | |
|---|---|---|---|---|---|---|
| overall users | 124 | | | 236 | | |
| no. of occurrences | 120 | 1* | 4 | 12 | 21 | 203 |
| | hiding input | checking for manipulations | checking surrounding | hindered | watched by company | no reason |

**Figure 4: Number of users that did or did not apply observable security measures. *One user applied two different security measures.**

did not hide the PIN entry, but instead checked their surrounding and verified that no one was standing nearby. One user additionally checked the ATM intensively for manipulations. To do so, he employed behavior as commonly proposed in the media and displayed on many cash machines. This mainly included grabbing and shaking the card slot and keypad to look for loose parts.

With 236 out of 360, almost two thirds of the observed users did not observably secure their input in any obvious way. This number increases when considering the users that only weakly secured their PIN entry. For instance, 15 users shielded their input only toward the screen, but left their PIN entry visible from the sides.

In the interviews, we wanted to get a better understanding why users would not protect their PIN entry. Therefore, we firstly asked them whether they are worried about someone stealing their PIN while using an ATM. 14 users, i.e. more than 50%, were not afraid of the risk of PIN theft. One of them even mentioned that "*the bank puts up cameras, so I am safe*".

Surprisingly, 19 out of 25 participants (including some that said that they were not worried about their PIN being stolen) stated that they would actually take security precautions, with 11 of these mentioning that they would always hide their input. This is a much higher percentage than we found in our primary field study, where barely a third secured their input. While part of this discrepancy could be attributed to "white lies" during the interview, a closer look at our interview logs revealed a more nuanced explanation: Several of the mechanisms people said they employed to secure their PIN entry were difficult – if not impossible – to detect during our observations. Consequently, the percentage of people securing their PIN entry could have been much higher than 34%.

For instance, three participants mentioned that they would hide their PIN entry with their body, blocking the view for onlookers. This is a rather large number considering that there was only a sample of 25 persons. However, during the field study, there was no situation in which a user efficiently blocked the view with his or her body. In all cases, our view to the keypad remained unblocked. Another three said they usually tried to choose an ATM inside a building, or that they would always choose the same ATM as a security measure. Six participants mentioned that they would check the surrounding while they were approaching the ATM. If


**Figure 5: Examples of how ATMs visualize to their users that they should apply security measures. Top: instructions to hide the PIN entry. Bottom: a visualization of how the card slot should look like.**

there was no one in sight, they would not hide their input. Since queues were rather seldom during our field studies, some users might not have hidden their input due to that reason. Finally, one user said that he would always do the input very quickly so no one could see it.

Interestingly, the majority of participants in the interview did not consider the danger of hardware based attacks, such as video recording and fake keypads. That is, many of the described measures – like fast input or hiding the input with the body – are rendered useless by those attacks. Therefore, a user might feel secure (e.g. when there is no one around) when she actually is not secure at all.

From both our observations and the interviews, we can infer that many users do not protect their input (203 during the observations) – or do so rather ineffectively. However, the reasons can be manifold. Apart from the obvious lack of interest, or a lack of threat awareness, we found three instances in which other factors hindered PIN security: physical hindrance, memorability, and trust display.

### 3.5.1 Physical Hindrance

Securing PIN entry against cameras and shoulder surfers typically requires a second hand to shield the keypad. We observed several instances where users simply did not have a free hand to spare to protect their input. For instance, they were holding shopping bags that they did not want to (or were unable to) put down. Other users were holding their mobile phone, having calls or even holding children in their arms. Overall, twelve instances of hindered, unsecured


**Figure 6: A staged example of a user that cannot hide the PIN entry due to physical hindrance.**

PIN entry were observed (see figure 4). An example of this (staged by the authors) is depicted in figure 6.

### 3.5.2 *Memorability*

Even though a four digit PIN is a rather short token to memorize, the increasing number of cards and services that depend on different PINs can make it difficult to remember them, prompting research into more memorable authentication methods (e.g. [14]). While during the 360 observations we only observed four sessions in which users forgot their PIN, these four cases vividly document how badly PIN entry fails when it does. Even though the first two cases were observed at two different ATMs on two different days, both users reacted in exactly the same way: after their first failed input attempt, both pulled out a notebook or piece of paper from their purses (in which they also kept their ATM card!) and consulted it for their PIN. After checking their notes in this way, both users could authenticate successfully. The third and fourth case showed similar behavior. Instead of having the PIN written down, however, those two users checked their iPods for their PINs.

Writing down PIN numbers or passwords to remember them was already reported as a major problem of token based authentication systems (e.g. in [1]). Within the scope of authentication in public, the danger even increases since an attacker can even more easily get into possession of the token, which the user carries around.

### 3.5.3 *Trust Display*

In many cases, users were with friends, family members, or partners. Of the 60 users that were not alone at the ATM, we found 22 instances (37%) in which users performed their PIN entry in plain view of their company (see figure 7). "Plain view" not only refers to not actively hiding the input, but more often meant that from their position, the accompanying persons could easily gaze on the whole interaction. In one case, a father even dictated his PIN to his (young) son so that he could have the "fun" of entering it.

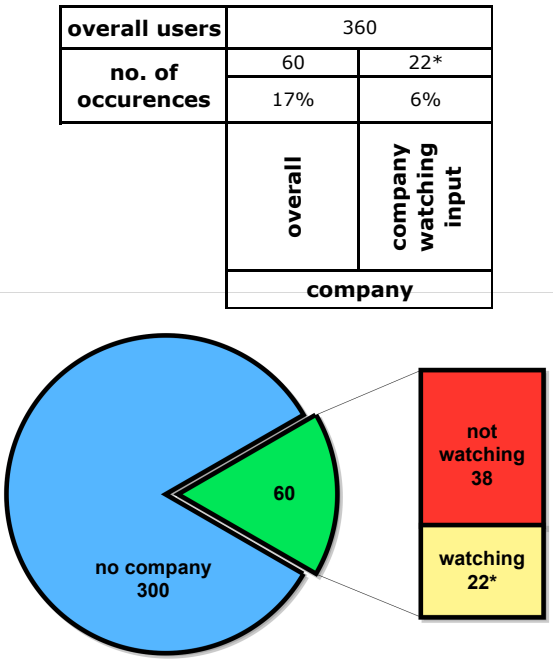


**Figure 7: 17% of users were in company, 6% let their companions watch their PIN entry. *Only one user that was watched by her companions applied security measures.**

Sharing (or at least not hiding) one's PIN in these situations might constitute a proof of confidence – or the other way around: hiding one's PIN might be constructed as a sign of mistrust toward the accompanying friends and family. The problem of social pressure and social factors has also been discussed by Kim et al. [12]. Social factors were one of their design criteria for their tabletop authentication system. To take the social pressure from the users, their systems are designed in a way that security is enforced. Our observations seem to support the importance of social factors on security.

To get a deeper understanding on this, the last block of questions in our follow-up interview study was "*whether users would protect their input if they are in company*". 13 Participants stated that they would still protect it while in company. One of these 13 mentioned that whenever he is around friends that used an ATM, he would look away since "*I don't want to put pressure on them*". The remaining twelve said that they would not protect their input while in company. However, only four of them were users that stated to hide their input with the other hand. Out of the participants that stated that they would not protect the input when friends were close, four stated that they would not protect it since they trusted their friends.

## 4. IMPLICATIONS

The insights we gained during our observation provide important feedback for the evaluation of authentication systems for ATMs. Therefore, in this section, implications for the design of authentication systems for public spaces are discussed, directly derived from our observations.

## 4.1 Authentication only a minor task

The numbers from our observations suggest that authentication only takes a marginal part of the whole interaction time with an ATM. With 46 seconds on average (or 54.9s when considering preparation and cleanup), more than 90% of ATM interaction is spent navigating menus and waiting for the withdrawn money (and optional receipts) to appear, etc. Distractions such as minding bags or talking to friends add further delay.

Being seen as a minor task that has to be done to be able to perform the actual task (e.g., withdraw money), it is questionable wether significantly slower authentication systems will be accepted by users. Considering an interaction time of 52.9 seconds, a system that takes, say around 12 seconds (e.g. [9]) adds an overhead of around 18% to the overall time.

The fact that we rarely observed longer queues (>2) during the observation, and that in our interviews we found that people based their decisions to queue or not on manifold factors, renders the "threat" of accumulated waiting times less significant. We can therefore support survey findings from [13] that people judge waiting time with respect to their time constraints and their need for cash. It seems that a queue length of two is a borderline that many people are only willing to cross if it is urgent and if their time constraints allow for it. However, increased authentication time can also have an influence on people waiting in the queue and would increase overall waiting times over accumulation.

Considering common authentication mechanisms from literature (e.g. [3, 7, 8, 9, 16, 19]), both waiting and overall interaction times can increase drastically if the authentication mechanism takes significantly more time. If, for instance, the interaction time for an authentication mechanism takes around 45 seconds (which is the average overall interaction time that was observed during the field study), the second user in the queue would have to wait twice as long as with PIN authentication. This indicates that time is a very important factor when creating an authentication system for public terminals, which can decide over acceptance or rejection of a system.

Within this work, we cannot provide an exact borderline on how long an authentication mechanism for ATMs should be. However, we argue that PIN authentication is only accepted by users since it is very easy and – maybe most importantly – extremely fast. Therefore, it is highly appropriate for ATMs, since the overall task is very short and PIN still only requires a small fraction of the overall time. A rule of thumb could be that an alternative authentication mechanism for ATMs should only require a fraction of the overall time (< 10%) that a user spends at the machine.

## 4.2 Security should not require an active user

There are several observations that support the notion that the security of an authentication mechanism should not rely on the way the user interacts with the ATM. In some cases, physical constraints (e.g. heavy shopping bags) did not allow the user to apply additional security precautions. Other examples had users try to hide their PIN entry, but using an angle that left the keypad in plain view for a "shoulder surfer". Much more often, however, was the case that users did not even try to hide their PIN input, either out of negligence or (potentially) as some sort of proof of confidence.

Clearly, an alternative authentication mechanism needs to minimize the ability of the user to disclose the shared secret (e.g., the PIN) by accident or through negligence. For instance, Sasamoto et al. [17] created a system that does not disclose the authentication token by simple observation. Also Kim et al. [12] created their authentication systems in a way that makes it impossible for the user not to hide it.

In other work it has already been noted that security is seldomly a user's primary goal [11, 18] and that users are "bad" in protecting their authentication tokens [1]. These results support our claim for authentication mechanisms that have security built-in. However, this often comes at the cost of usability and has to be handled carefully.

## 4.3 Social compatibility

When designing an authentication mechanism that does not require an active user, the problem of social compatibility might – but does not necessarily have to – already been solved. Results from the field observations as well as results from the field interviews indicate that social factors can lead to insecure behavior. Therefore, authentication mechanisms should be compatible with social norms.

That is, to commit secure behavior, a user should not have to perform an action that might be misinterpreted as showing mistrust to a person accompanying her.

## 4.4 Memorability not majority problem, but still major one

Out of the 360 users, only four were not able to correctly recall their PIN at the first try. While it could thus be argued that memorability is not a problem for the large majority of users, this might be premature. Firstly, in the few cases where it was a problem, severe security problems resulted (e.g. PIN written down). Secondly, our results are most likely biased toward the most often used PIN. If we would have required people to recall PINs of membership cards or seldom used credit cards (which increasingly require a PIN as well), we might have gotten a very different picture.

Therefore, especially for authentication systems for public spaces, memorability deserves a lot of attention.

## 4.5 Authentication in highly distractive environments

As our observations showed, distractions can appear in manifold ways, and in particular in the form of ongoing social interactions (chat). ATM authentication mechanisms should therefore remain simple and work even without giving them their full attention. For instance, an imaginable authentication game that requires the user to follow a row of events might not be appropriate for an ATM.

# 5. LIMITATIONS OF THE RESULTS

Since the main observation took place in two central European cities, it has only limited validity with respect to other cultural areas (e.g., Asia) or in less urban settings.

The unobtrusive nature of the observations did not allow for in-depth findings on whether people check the hardware of an ATM (keypad or card slot) for manipulations. However, our general findings suggest that people only rarely use this security measure.

As for any study that involves direct contact to the participants, the field interviews might have been slightly biased since the participants might have wanted to "look good" or

"do it right". Therefore, the numbers on hidden input might be higher than they are in reality, which our field observations seem to confirm.

# 6. LESSONS LEARNED

In preparation for and while performing the observations discussed in this work, several lessons were learned. The presented lessons have proven especially helpful when dealing with sensitive and private data – as field observations on ATMs surely do. We argue that in this work we could show the value of observations in revealing important information about a study subject that could not have been revealed in laboratory studies. The lessons learned are meant to help any researcher that wants to conduct usable privacy and security observations in the public.

## 6.1 Pre-studies

As mentioned in the methodology section, we performed a set of pre-studies to figure out what data we could collect and how to best collect it. Pre-studies of this nature are especially helpful when an observer has to rely on written observations only. To be compliant to ethical and legal rules, in a scenario such as observing ATM use, no recordings of any kind should be made. Thus, a well defined and well prepared checklist can help significantly to ease the work of the observer. In this work, the pre-studies helped us to significantly improve the checklist used to collect the data during the actual observations. Therefore, pre-studies can be highly recommended to get an idea on which data an observer can and wants to measure.

## 6.2 Abide to strict rules

To guarantee validity and comparability of the gathered data, the observer should abide to strict rules. This also helps to avoid unethical behavior. During the field observations, we applied strict rules on when a data set was valid and thus could be added to the overall data. For instance, a rule stated that if the line of sight was blocked for any amount of time, the data would be discarded. While this led to a significant amount of observations that had to be discarded, it also helped to gather good and comparable data. A rule that was supposed to avoid unethical behavior was that the observer was positioned in a way that he could see when the keypad was touched but could not see the keypad itself.

## 6.3 Know the limitations

We are aware that there are many limitations when doing observational research, and so should anyone who attempts to perform this kind of research. For instance, observed behavior might be wrongly interpreted. Also, the results are most probably limited to the specific cultural area they have been collected in. Even with these limitations, however, they can lead to important insights.

## 6.4 Difference to laboratory studies

Not surprisingly, field study results can significantly differ from laboratory study results. For instance, during the work on MobilePIN [6], 89% of the participants stated that they would use measures to secure their PIN entry. In our observations, however, we could only observe 34% really doing so (though 79% claimed to do so in the interviews).

# 7. CONCLUSIONS

On the basis of a field study, an additional in-depth study, and a small set of street-interviews, we were able to identify several factors that are likely to influence the performance and security of authentication mechanisms for ATMs. Our observations revealed practices that suggest specific design decisions for ATM authentication systems, e.g., over 65% of users did not hide their PIN entry at all, or did so only weakly. This suggests that security for ATMs cannot rely on the user but needs security features which are "built in" into the authentication mechanism. That is, the security of a system should not rely on active secure behavior of a user.

The observations further helped to identify contextual factors that can have a great impact on the systems. Simple factors like prams, shopping bags, phone calls, etc., can be a reason for not applying security or for being slow. We also found that social factors (showing trust) can be a reason for bad security decisions.

However, there are aspects of ATM authentication mechanisms that this study cannot answer, but which are nonetheless of great importance when creating respective authentication systems. Most likely deployment cost are one of the most decisive factors in this context: how much will it cost service providers to update all their ATMs (or public terminals) to a new system? Other factors could be, e.g., resistance to vandalism.

This work represents a first step in uncovering ATM use in the wild, hopefully helping to gain a broader insight on the real factors and constraints of ATM authentication. For future work, we would like to extend our observations to other forms of electronic payment (e.g., ticketing machines, supermarket checkout), where we expect slightly different circumstances leading to noticeable differences in use. For instance, we believe that in a supermarket setting, we might experience even more insecure behavior. Also, we would like to encourage other researchers to perform similar studies in different cultural and/or urban settings since we are highly interested in how these findings will apply there.

# 8. ACKNOWLEDGMENTS AND ACCESS

We would like to thank the many people that read the first version of this work and significantly helped to improve it. Furthermore, we would like to thank Marcello Scipioni for his highly appreciated help with conducting the interviews. If you are interested in getting access to the data sets used in this work, please contact the first author of this paper. We are more than happy to share them with any researcher interested in this topic.

# 9. REFERENCES

[1] A. Adams and M. A. Sasse. Users are not the enemy. *Commun. ACM*, 42(12):40–46, 1999.

[2] 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, New York, NY, USA, 2007. ACM.

[3] S. Chiasson, P. C. V. Oorschot, and R. Biddle. Graphical password authentication using cued click-points. In *12 th European Symposium On*

*Research In Computer Security (ESORICS), 2007*. Springer-Verlag, 2007.

[4] L. Coventry, A. De Angeli, and G. Johnson. Usability and biometric verification at the atm interface. In *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 153–160, New York, NY, USA, 2003. ACM.

[5] A. De Luca, M. Denzel, and H. Hussmann. Look into my eyes! can you guess my password? In *SOUPS '09: Proceedings of the 5th symposium on Usable privacy and security*. ACM, 2009.

[6] A. De Luca, B. Frauendienst, S. Boring, and H. Hussmann. My Phone is my Keypad: Privacy-Enhanced PIN-Entry on Public Terminals. In *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group (CHISIG) of the Human Factors and Ergonomics Society of Australia (HFESA), Melbourne, Australia, Nov. 2009*. ACM, Nov. 2009.

[7] A. De Luca, E. von Zezschwitz, and H. Hussmann. Vibrapass - secure authentication based on shared lies. In *27th ACM SIGCHI Conference on Human Factors in Computing Systems*. ACM, Apr. 2009.

[8] A. Forget, S. Chiasson, and R. Biddle. Shoulder-surfing resistance with eye-gaze entry in cued-recall graphical passwords. In *Proceedings of the 28th ACM International Conference on Human Factors in Computing Systems - CHI 2010, Atlanta, Georgia, USA, Apr. 2010*. ACM, Apr. 2010.

[9] E. Hayashi, R. Dhamija, N. Christin, and A. Perrig. Use your illusion: secure authentication usable anywhere. In *SOUPS '08: Proceedings of the 4th symposium on Usable privacy and security*, pages 35–45, New York, NY, USA, 2008. ACM.

[10] E. M. Huang, A. Koster, and J. Borchers. Overcoming assumptions and uncovering practices: When does the public really look at public displays?. In J. Indulska, D. J. Patterson, T. Rodden, and M. Ott, editors, *Pervasive*, volume 5013 of *Lecture Notes in Computer Science*, pages 228–243. Springer, 2008.

[11] C. Jackson, D. R. Simon, D. S. Tan, and A. Barth. An evaluation of extended validation and picture-in-picture phishing attacks. In *In Proceedings of Usable Security (USEC07)*, 2007.

[12] D. Kim, P. Dunphy, P. Briggs, J. Hook, J. Nicholson, and P. Olivier. Multi-touch authentication on tabletops. In *Proceedings of the 28th ACM International Conference on Human Factors in Computing Systems - CHI 2010, Atlanta, Georgia, USA, Apr. 2010*. ACM, Apr. 2010.

[13] L. Little. Attitudes towards technology use in public zones: the influence of external factors on atm use. In *CHI '03: CHI '03 extended abstracts on Human factors in computing systems*, pages 990–991, New York, NY, USA, 2003. ACM.

[14] W. Moncur and G. Leplâtre. Pictures at the atm: exploring the usability of multiple graphical passwords. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 887–894, New York, NY, USA, 2007. ACM.

[15] P. Peltonen, E. Kurvinen, A. Salovaara, G. Jacucci, T. Ilmonen, J. Evans, A. Oulasvirta, and P. Saarikko. It's mine, don't touch!: interactions at a large multi-touch display in a city centre. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 1285–1294, New York, NY, USA, 2008. ACM.

[16] V. Roth, K. Richter, and R. Freidinger. A pin-entry method resilient against shoulder surfing. In *CCS '04: Proceedings of the 11th ACM conference on Computer and communications security*, pages 236–245, New York, NY, USA, 2004. ACM.

[17] H. Sasamoto, N. Christin, and E. Hayashi. Undercover: authentication usable in front of prying eyes. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 183–192, New York, NY, USA, 2008. ACM.

[18] A. Whitten and J. D. Tygar. Why johnny can't encrypt. In *In Proceedings of the 8th USENIX Security Symposium*, 1999.

[19] S. Wiedenbeck, J. Waters, L. Sobrado, and J.-C. Birget. Design and evaluation of a shoulder-surfing resistant graphical password scheme. In *AVI '06: Proceedings of the working conference on Advanced visual interfaces*, pages 177–184, New York, NY, USA, 2006. ACM.