Computers in Human Behavior Reports 5 (2022) 100161

Contents lists available at ScienceDirect

Computers in Human Behavior Reports

journal homepage: www.sciencedirect.com/journal/computers-in-human-behavior-reports

Complex, but in a good way? How to represent encryption to non-experts through text and visuals – Evidence from expert co-creation and a vignette experiment

Verena Distler [a,*], Tamara Gutfleisch [a,c], Carine Lallemand [a,b], Gabriele Lenzini [a], Vincent Koenig [a]

[a] University of Luxembourg, Esch-sur-Alzette, Luxembourg
[b] Eindhoven University of Technology, Eindhoven, Netherlands
[c] Mannheim Centre for European Social Research, University of Mannheim, Mannheim, Germany

ARTICLE INFO

Keywords:
Human-computer interaction
Encryption
User experience
Perceived security
Understanding
Online banking
E-voting

ABSTRACT

An ongoing discussion in the field of usable privacy and security debates whether security mechanisms should be visible to end-users during interactions with technology, or hidden away. This paper addresses this question using a mixed-methods approach, focusing on encryption as a mechanism for confidentiality during data transmission on a smartphone application. In study 1, we conducted a qualitative co-creation study with security and Human-Computer Interaction (HCI) experts (N = 9) to create appropriate textual and visual representations of the security mechanism encryption in data transmission. We investigated this question in two contexts: online banking and e-voting. In study 2, we put these ideas to the test by presenting these visual and textual representations to non-expert users in an online vignette experiment (N = 2180). We found a statistically significant and positive effect of the textual representation of encryption on perceived security and understanding, but not on user experience (UX). More complex text describing encryption resulted in higher perceived security and more accurate understanding. The visual representation of encryption had no statistically significant effect on perceived security, UX or understanding. Our study contributes to the larger discussion regarding visible instances of security and their impact on user perceptions.

1. Introduction

Streamlining people's interactions with technology might help improve usability but can lead to some unintended secondary effects in the context of security and privacy. In the quest to make interactions more "user-friendly", security mechanisms have often been hidden away from users under the rationale that they can introduce barriers to action, while Human-Computer Interaction (HCI) designers attempt to remove such barriers (Dourish et al., 2004). Accordingly, automated approaches of security that remove security decisions from the user's hands have emerged (Edwards et al., 2008).

But when users do not need to interact with security, they likely also do not need to understand security processes. This lack of understanding can lead to security issues (Adams & Sasse, 1999). Authors thus reasoned that security technologies should be highly visible and available for inspection (Adams & Sasse, 1999), with some explaining that only by making security-related actions and their consequences more visible, users are able to form accurate mental models about the security of an interaction (Spero & Biddle, 2020). Some authors also argued that security can even act as an enabling factor and a significant part of positive user experience (Pagter & Petersen, 2007).

To investigate these questions, in the present paper we focus on the security mechanism encryption, applied to provide confidentiality during data transmission on a smartphone application. To understand user perceptions, we investigate three concepts. First, we are interested in perceived security, which we define as how secure or insecure an experience felt to the research participant (see section 5.1.1). Second, we research user experience, which we define and discuss in section 2.3., and measure using the UEQ-S measurement (Schrepp et al., 2017). Third, we investigate the understanding of the security mechanism

* Corresponding author. University of Luxembourg, Human-Computer Interaction Research Group, 11 Porte des Sciences, Esch-sur-Alzette, Luxembourg.
E-mail addresses: Verena.distler@uni.lu (V. Distler), tamara.gutfleisch@mzes.uni-mannheim.de (T. Gutfleisch), Carine.lallemand@uni.lu (C. Lallemand), gabriele.lenzini@uni.lu (G. Lenzini), Vincent.koenig@uni.lu (V. Koenig).

https://doi.org/10.1016/j.chbr.2021.100161
Received 7 July 2021; Received in revised form 13 October 2021; Accepted 8 December 2021
Available online 25 December 2021

encryption based on a set of exploratory questions designed by security experts for non-expert users (described in section 5.1.1.). In the following, we will frequently refer to "understanding" as a shorter form of "understanding of the security mechanism encryption" for better readability.

This paper is organized as follows. In section 2, we introduce the background for our research, including research on visual representations of security mechanisms, use contexts of our research, and work on measuring subjective experiences. In section 3, we explain the research objectives. We then describe the iterative co-creation of representations of encryption with experts in section 4 (study 1), and the vignette experiment with non-experts in section 5 (study 2). Section 6 discusses the results of our work, before concluding in section 7.

2. Background

The question of how to represent security mechanisms visually remains a challenge. In the following, we use the term "visible instances of security" to describe any visible representation of a security mechanism to the user of a technology. A visible instance of security can encompass both visual and textual indicators (e.g., an image with some text).

Icons have long been used in graphical user interfaces to convey information (Blattner et al., 1989) and have the potential of being universally understood, even though sometimes, a range of different meanings can be attributed to a single icon (Rogers, 1989). As early as in 1999, Wiedenbeck (1999) evaluated the learnability of an application using buttons with text labels, icons, or a combination of both, measuring both the success of novice users learning how to use the application and measuring users' attitudes toward the application. Performance was best when using text labels only, or when combining icons with text labels; performance using the icon-only interface was much poorer. Ease of use was perceived as better for the icon-label interface, and perceived usefulness was higher for the icon-only and icon-label interfaces. This study seems to show that a combination of textual and visual representation may be most suited to convey information.

In the following, we will describe some of the research on visual representations of security mechanisms, and how they relate to user perceptions.

2.1. Research on visual representations of security mechanisms and perceptions

One instance of an encryption protocol that is familiar to many is HTTPS. It is used for implementing confidential communications towards interlocutors whose identity is certified as trusted.[1] Various studies explored how to visualize the presence of an HTTPS connection (resp. the lack thereof) to inform users whether they are transmitting sensitive data e.g., credit card numbers securely (resp. or insecurely) or to someone trusted (resp. or untrusted).[2]

Schechter et al. (2007) evaluated different connection security indicators and warnings, finding that participants failed to recognize the absence of a HTTPS indicator. Even when a warning page was displayed, suggesting that it may be unwise to visit an untrusted website whose certificate is invalid or expired, potentially suggesting that the website is not what it claims to be or that its identity was certified a long time ago and might have changed, many participants still took the risky action of visiting the website. The authors confirm prior findings that users seem to ignore HTTPS indicators and warnings. Felt et al. (2015) later designed visual indicators for the presence/absence of HTTPS secure connections, with the goal of improving understanding of these indicators, as well as adherence to the secure behavior, which they defined as not visiting the untrusted website. The authors were not able to improve understanding of the security warning, but improved adherence through opinionated design. Later, Felt et al. (2016) also designed new indicators for the presence or absence of HTTPS secure connections for browsers, and evaluated their effects on users. The authors indicate HTTPS in green with a padlock and the text "secure", HTTP in grey with a circle icon (resembling an "information" symbol) and the text "not secure", and invalid HTTPS in red with a triangle with an exclamation mark, and the text "not secure". The selection was implemented by Google Chrome. In a 2021 blog post, Chrome researchers highlighted previous research showing that the lock icon was often associated with a website being trustworthy, when really only the connection is secure (Panditrao et al., 2021). Due to this misalignment between people's interpretation of the icon and the actual security property it intended to indicate, the researchers planned to run experiments with removing or replacing the lock icon. The results are not publicly available at present.

Similar situations as with HTTPS arise with encrypted email. Also here "security" stands for several meanings such as "confidentiality", "sender/receiver identity authentication" and peculiarly to emailing and messaging,"integrity of a message", and "end-to-end encryption". Once more, this multifaceted role of the term "security" has given rise to several misunderstandings, while being a source of great confusion among users. The technical difficulty to make the whole encryption mechanism working as intended, which often requires users to perform additional actions such as creating and managing encryption keys on top of writing and sending messages, did not help the cause of securing the email, and encryption is still rarely used by laypersons. That said, even the goal of informing users of the presence of a mechanism to ensure the confidentiality of their messages through encryption has not been easy. As early as in Whitten and Tygar's (1999) seminal paper on the usability of PGP 5.0, usability issues made it difficult for non-expert users to make use of encrypted emails. In their study, most novices were unable to successfully encrypt their emails in a 90 minutes time period. Later work confirmed that usability issues, in addition to social factors (e.g., being viewed as paranoid for encrypting emails) play a role in the adoption of encryption (Gaw et al., 2006). Ruoti et al. (2013) evaluated a webmail system that uses security overlays with existing mailing services like Gmail. Their version of the tool was mostly invisible with automatic key management and encryption. Their participants were mostly able to use the system without any training, but the security aspects were so invisible that some mistakenly sent out unencrypted messages, and were concerned about trusting the tool. The authors then conducted a study with a prototype that used manual encryption, which enabled participants to avoid mistakes and led to more trust in the system. Lausch et al. (2017) reviewed security indicators in the context of secure emails and found that adding images of postcards, closed envelopes, and a torn envelope may warrant further work since they offered a relatively consistent interpretation. The authors also highlighted that the security indicators for encrypted email in different applications are mostly padlocks, but a variety of indicators exist for encrypted email (as well as signed and unsigned email), making it complicated for users to understand their meaning. They did not study text in association with the icons.

More recently, secure communication has often expanded to also include end-to-end encrypted messaging, for example for applications such as Signal or Whatsapp. Fahl et al. (2012) studied the usability and perceived security when encrypting Facebook messages, comparing combinations of manual and automatic encryption and key management. The authors found the highest usability in the versions of their prototype that included no display of security, where encryption was completely automated, or a combination of manual encryption and

[1] HTTPS also has the goal of authenticating the identity of the server for the reason that "secure" messages should be confidential but also sent to the intended recipient and not, despite confidentially, to an imposter.

[2] The potential issues are (1) their data will be sent in clear and can be read if the protocol is HTTP, or, (2) if the protocol is HTTPS, will be sent encrypted but to a recipient who may be who it claims to be (e.g., Amazon), but the certificate is invalid, or the recipient is not who it claims to be.

automatic key management. Researchers have also often focused on authentication-related interactions, which users can have difficulties understanding or performing (Vaziripour et al., 2017), sometimes noting that inconsistent interface design and technical wordings can make it difficult to use the tools securely and as intended by the designers (Abu-Salma et al., 2017). A recent study by Fassl et al. (2021) explored a user-centered design process to improve useable authentication ceremonies. Instead of incrementally improving existing ceremonies, they employed a user-centered process to design new ceremonies from scratch in collaborative design workshops, followed by a security evaluation to narrow the design space, an iterative storyboard prototyping approach to improve usability, and an online evaluation. This user-centered approach took into account the social aspects of authentication ceremonies. While their approach did not result in better UX or usability, participants gained an improved understanding of security implications of authentication ceremonies.

A study on textual descriptions of encryption during data transmission in multiple contexts found that the verbs "secure" and "encrypt" were perceived as relatively secure, but the study did not combine the text samples with images or icons (Distler, Lallemand, et al., 2020). While privacy icons do not serve to represent an underlying security mechanism per se, some of the insights from studies on how to represent privacy concepts are relevant to our study. A study on the design of privacy icons (Cranor, 2021) demonstrated the importance of placing link text next to the icons for participants to understand what it meant. It is also important to consider different user groups as the usability of icons also differs between different age groups, with older adults needing more time to select icons, but giving the same number of correct responses in a navigation task (Dosso & Chevalier, 2021).

In this article, we will focus on the security mechanism "encryption", a mechanism that is ubiquitously used to ensure secure digital communications, yet mostly invisible in the user interface. We will focus on encryption as a mechanism used mostly for confidentiality during data transmission on a smartphone application. We will address the question of how to display this security mechanism in two contexts, e-voting, and online banking when optimizing the experience for perceived security, UX, and understanding. We will now describe examples where visible representations of encryption were empirically assessed with end-users in these two contexts, and then situate our study conceptually.

2.2. Security mechanisms in specific use contexts

We will now introduce some previous work on visible instances of security mechanisms in the contexts of e-voting and online banking, as well as factors that were found to influence security perceptions in previous work.

2.2.1. E-voting

E-voting is a high-stakes use context where encryption is used to ensure vote confidentiality, together with other cryptographic mechanisms that are often employed to ensure a trustworthy electronic election process, for instance, to help users and authorities verify that votes are not lost, tampered with, or selectively discarded, and that the vote counting has not been compromised. Elections have a complex work and information flow, and it is hard for citizens to have a detailed picture of the whole process, with or without the use of security mechanisms which, of course, complicate the picture. In the following, we will focus on e-voting using a smartphone application; technology-supported voting at the polling station is out of scope for our purposes.

Existing e-voting applications can be hard to use and not always perceived as secure by the users. While vote verification is considered a cornerstone of secure elections, it also often leads to usability issues (Acemyan et al., 2014).

Remote e-voting is already used in some countries, for instance Estonia (Alvarez et al., 2009; Vassil et al., 2016) and Switzerland (Petitpas et al., 2020). A study compared the usability of multiple e-voting schemes and demonstrated that insufficient usability led to a considerable proportion of participants unable to cast a vote across voting systems, and many were unable to verify whether their vote had been taken into account. Overall satisfaction was low (Acemyan et al., 2014). A coercion-resistant e-voting with transparent verifiability protocol is "Selene" (Ryan et al., 2016). Selene allows voters to verify their vote using a tracking number to find their vote in clear on a bulletin board, providing a simple approach to vote verification. Distler et al. (2019) describe the design of an e-voting application based on the existing e-voting protocol Selene in two versions, one of which displays more security-related information to users. The version with "more information" ("version D") included a visual of encryption, whereas the other version displayed no encryption-related information. In addition, version D also included more explanation about the vote verification process. Their results suggest that the version displaying more information may perform better overall in terms of UX and psychological need fulfilment, even though they caution to interpret these results carefully since results were statistically non-significant, potentially due to a relatively small sample size for intergroup comparisons, suggesting that more work is needed. Marky et al. (2018) evaluated the usability of different implementations of the Benaloh challenge for cast-*as*-intended vote verification, comparing three approaches. Based on their results, the authors recommend using the mobile approach for deployment during elections, and using the automatic approach for those who do not own a smartphone or similar device.

2.2.2. Online banking

In most European countries, e-voting is not routinely used for major political elections. Online banking is a more common and more frequent interaction than voting for many people. Online banking can take place on a computer, using the browser, on a smartphone or other mobile devices, often using either a mobile application or the browser, in addition to various options for two-factor authentication that are frequently used (e.g., a second smartphone application, codes to be received via SMS, or a separate hardware token). This combination of options for online banking can make it difficult for users to accurately assess the presence or absence of security mechanisms during the interaction.

Online banking is perceived as security-critical by users (Distler, Lenzini, et al., 2020) and previous studies have found that perceived security and trust had a positive impact on the acceptance of online banking (Damghanian et al., 2016). Perceived risk had no significant effect on the acceptance of online banking, but on trust in online banking. Authors have argued that banks should take better steps to persuade their customers about the security and usefulness of their online banking systems (Özlen & Djedovic, 2017). Khan et al. (2017) investigated the acceptance of online banking in and found that perceived security, as well as performance expectancy, facilitating conditions, habit, and privacy value were important antecedents of behavioral intentions. A study in the context of financial technology (Lim et al., 2019) found that perceived security and knowledge have an effect on users' confirmation (the extent to which the users' expectation of a service are fulfilled) and the perceived usefulness of a mobile fintech payment services, but perceived security did not directly impact on user satisfaction and continual intention to use.

The studies above have in common that researchers evaluate subjective experiences, and frequently, attempt to design for a user-friendly interaction that users understand and perceive as secure. How to measure such subjective perceptions is a challenge that can in part be addressed through the concept of user experience.

2.3. Measuring subjective experience through user experience

The evaluation of people's interactions with technology is a challenge that was traditionally addressed by the field of usability, but the concept has shifted to the broader concept of user experience. Usability

focuses on the users' ability to achieve their goals effectively, efficiently and to their overall satisfaction (International Organization for Standardization, 2018). Authors have argued that a certain level of usability is necessary as a basis for a positive experience (Hassenzahl et al., 2013), but will not necessarily lead to a positive experience on its own. In addition to users being able to achieve their tasks, user experience also takes into account the non-instrumental qualities that many experiences fulfil (Hassenzahl, 2001). These non-instrumental qualities refer to functions an interaction fulfils that are not directly goal oriented, but instead could fulfil psychological needs such as feeling connected to others (relatedness) or self-actualization (Sheldon et al., 2001). Adopting user experience as a frame of reference can help obtain a broader understanding of how users perceive an interaction. An efficient way to measure UX are standardized scales such as the Attrakdiff (Hassenzahl et al., 2003) or the User Experience Questionnaire UEQ and its shorter versions, the UEQ-S (Laugwitz et al., 2008; Schrepp et al., 2017).

In addition to UX, we also measure perceived security and understanding of the security mechanism encryption, as described in section 5.1.1. Conceptually, we see security perceptions as related to the psychological need for security (Sheldon et al., 2001). The measurement of understanding of an interaction, or, in our case, of the security mechanism encryption, is more difficult to situate within the framework of UX, but understanding is often highly relevant in useable privacy and security (UPS) contexts where misunderstandings can lead to security issues.

2.4. Summary

There is a growing body of research that calls for more visible and transparent communication of security mechanisms to end-users. Dourish et al. (2004) argued that security technologies should be visible to users, to provide people with the means to understand the security implications of the current configuration of technologies they are using. This visibility should be expressed not as mathematically-grounded concepts of cryptography, but in terms that are adapted to the users' activities and needs at the time. Rather than making information about security mechanisms available when the user requests it, it should be available as a part of every activity in the system (Dourish & Redmiles, 2002), and authors argue that displaying security mechanisms more clearly could help improve users' mental models and understanding of the security state of their interaction (Spero & Biddle, 2020). Pagter and Petersen (2007) suggested that security mechanisms could in fact become a significant part of positive experiences by providing a perception of security. Indicators for the presence of the security mechanism encryption have been tested in contexts such as connection security indicators, encrypted email, encrypted messaging applications and e-voting, frequently finding that people's understanding was inaccurate and not always inducing perceived security when warranted (Acemyan et al., 2014; Distler et al., 2019). Perceived security was also an important factor for the acceptance of online banking (Damghanian et al., 2016; Özlen & Djedovic, 2017). Going beyond perceived security, people's understanding of the security mechanisms in place is also an important aspect to consider, to ensure that their understanding is as accurate as possible and avoid erroneous mental models. Finally, people's user experience, as a broader measure of people's overall impressions of the interaction, is a promising concept to provide additional information with regards to the subjective experience of security-critical interactions.

Despite existing user studies of various encryption technologies, the HCI and UPS communities mostly lack concrete guidelines on how to communicate many of these security mechanisms to end-users. Current practices also sometimes lead to misunderstandings of the security provided by a system. In particular, previous work does not describe causal relationships between specific textual and visual representations of the security mechanism encryption on perceived security, UX and understanding. In addition, existing work mostly focuses on one specific use context and implications on how to display the security mechanism encryption are thus not necessarily transferable to other contexts.

We will now describe how we will contribute to closing these gaps.

3. Research objectives

The main aim of this research is to evaluate the effects of textual and visual representations of encryption on non-experts' perceived security, UX and understanding in two contexts (e-voting, online banking). Our research design involved two studies. The purpose of study 1 is to inform the design of our vignette experiment in study 2.

In our first study (section 4), we involved experts from the fields of security, privacy and HCI to develop ideas on how to communicate encryption to non-experts via textual and visual representation:

- RQ1: How do HCI and security experts suggest to display the security mechanism encryption to non-expert users using textual and visual representation in the context of e-voting and online banking?

The first study allowed us to obtain the visual and textual representations for our main study; in our second study (section 5), we tested the impact of these representations on non-experts' perceived security, UX and understanding in a vignette experiment, comparing the use contexts e-voting and online banking.

- RQ2: What is the effect of *visual* representation of encryption on perceived security, user experience and understanding of the security mechanism encryption?
- RQ3: What is the effect of *textual* representation (including the complexity of text) of encryption on perceived security, user experience and understanding of the security mechanism encryption?

We also address an additional methodological question. Since to the best of our knowledge no measurement of understanding of encryption exists in prior research, we explore how we might measure non-experts' understanding of the security mechanism encryption across both studies (sections 4 and 5). Based on experts' suggestions in study 1, we created a measurement for understanding of encryption in study 2 and included it in the vignette experiment. We further openly asked participants what they thought our exploratory understanding questions were intended to measure and analyze the answers to this question.

4. Study 1: iterative co-creation of representations of encryption with experts

To address our first research question, we conducted multiple co-creation activities with security and HCI experts to find out how we may best represent the security mechanism encryption to non-expert users and how we may measure understanding of encryption. The study was approved by our institution's ethics board, and experts provided informed consent. To define the visual and textual representations of encryption, we used an iterative design process, where experts were confronted with previous experts' ideas and opinions. Fig. 1 shows the four stages of our iterative design process.

4.1. Ideation with security experts

4.1.1. Participants

In this first phase, five security experts were recruited for an ideation session. The experts were three PhD Candidates, one Postdoctoral Researcher and one Full Professor. The PhD Candidates had between 0 and 5 years of experience in the field, the Postdoctoral Researcher between 5 and 10 years, and the Full Professor more than 10 years of experience. The experts participated in sessions of 1–1.5 h each and were compensated with 40€ for their time. The participants were not part of the author team and recruited through the personal network of

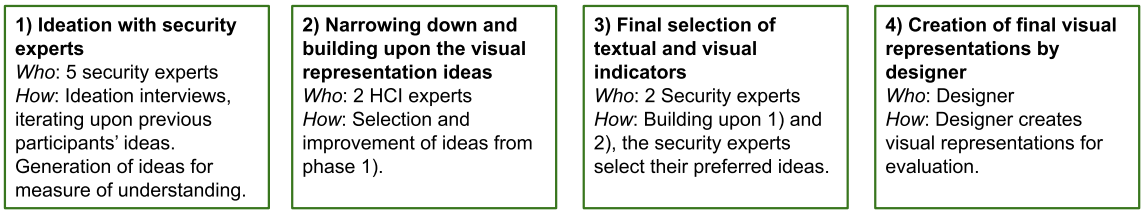


Fig. 1. Summary of methodology study 1.

the first author. Four of the ideation sessions took place in the user lab, and one remote. We pre-tested the protocol; it worked as intended and we only made minor changes to the protocol, allowing us to include the pilot participant into the final set of security expert participants.

4.1.2. Procedure

The facilitator guided the experts through a number of questions and tasks. First, the experts were invited to explore ways of describing encryption to non-expert users while explaining their thought process. Next, we asked the experts to generate ideas for questions they might ask non-experts to measure whether they had understood encryption being used during data transmission. We asked the experts to explain their thought process for the questions they came up with.

Then, the experts were presented with three visual representations that are, or could be, used to represent security concepts (hand-drawn images of padlock, shield, database with a key, see Fig. 2). We used hand-drawn images to alleviate any concerns about not being able to draw "well enough" in the next stage of the procedure. They were asked to rank these in terms of how well they represented encryption and to discuss critically what they liked and disliked about the visual representations.

Finally, the experts were asked to generate at least three ideas of visual representations of encryption, going beyond the common ones we asked them to critique. While the first expert participant was only presented with the three initial visual representations, the second expert was asked to critique the same three visual representations, plus the visual representations expert 1 had come up with. Expert 3 critiqued the initial visual representations, plus the visual representations experts 1 and 2 had come up with, and so on. Thereby, the experts built upon the ideas of previous experts, resulting in a rich collection of ideas for visual representations of encryption. Before presenting the previous experts' ideas, we redrew them so that they could not recognize a colleague's handwriting and would not hesitate to critique their ideas.

4.1.3. Results

The experts mainly critiqued that the padlock and the shield seemed too unspecific to indicate a security mechanism such as encryption, and associated the third icon with a database rather than any specific security mechanism. They also critiqued and built upon the previous' experts' ideas as shown in Table 3 (appendix), which demonstrates the evolution of the visual indicators through the various stages of study 1.

Fig. 2. Hand-drawn images of padlock, shield and database with key that experts were asked to critique.

The experts also suggested a variety of explanations of encryption, as well as ideas on questions that they might ask a non-expert to evaluate their understanding of encryption as a security mechanism. Using the pool of potential questions intended to measure understanding of encryption, we selected a set of questions (mainly excluding any repetitive questions) and presented these to a security expert for feedback and improvement. We then presented this improved set of questions to another security expert, who also suggested improvements. This iterative process led to six questions intended to evaluate whether participants' mental model corresponds to an interaction secured by encryption.

Overall, the experts suggested focusing on questions that evaluated the most important implications of encryption, as they would not expect non-experts to be able to explain how encryption worked specifically. We use these suggestions for how to evaluate understanding of the security mechanism encryption in study 2 (see section 5.1.1).

4.2. Narrowing down and building upon the visual representation ideas with HCI experts

In this step, our objective was to narrow down and improve the large number of ideas of visual representation generated by the security experts in Phase 1.

4.2.1. Participants

We recruited two HCl experts from the personal network of the authors (not part of the author team) to take part in a 1.5 h conference call. One of the experts had between 0 and 5 years of experience, the other between 5 and 10 years of experience in HCI. Their main expertise did not lie in the field of useable privacy and security. The experts were compensated with 40€ for their time.

4.2.2. Procedure

In the meeting, the experts were first presented with all the visual representations, and asked to individually think about which ones were most promising for use in a smartphone application (not in a tutorial context, but presented briefly as part of a smartphone interaction). They could also choose to modify visual representations they thought were promising but could be improved on certain aspects. After the individual task, both experts were asked to converge their opinions in a shared document and discuss which visual representations to keep, remove or modify. This phase yielded a set of visual representations that were deemed suitable for smartphone interactions and a set of recommendations on how to change visual representations to be more user-friendly. Using the HCI experts' suggestions, the first author created modified versions and presented them to the HCl experts for feedback the day after the initial call.

4.2.3. Results

Overall, the experts mainly opted to exclude representations that seemed too complex for being viewed only briefly in the context of a smartphone application, as these seemed more appropriate for tutorial-style interactions. They also asked to standardize the way certain

components were visualized (e.g., by always using the same visual representation to represent a polling station or bank).

4.3. Final selection of textual and visual indicators

4.3.1. Participants

In this phase, we recruited two security experts (one postdoc and one PhD researcher) who had not participated in the previous stages. One of the experts had between 0 and 5 years, the other between 5 and 10 years of experience in the field of security respectively. Participation took 20–30 min and was asynchronous. The experts were compensated with 20€ for their time.

4.3.2. Procedure

We created a shared worksheet to be filled out by our security experts participants. The participants were first presented with all the expert ideas on how to describe encryption to non-experts and asked to highlight their favorite ideas out of the options. Then, we asked them to build upon these ideas to create a textual description of encryption with a low level of detail; as well as a description with a high level of technical complexity (yet still accessible for non-experts). Then, we presented the visual representations that stemmed from phase two and asked the expert to select their four favorite visuals and explain their selection.

4.3.3. Results

As a result, we obtained six favorite visual representations of encryption. The experts also built upon all previous ideas to create explanations of encryption with varying levels of detail. We (the authors) selected and combined their preferred descriptions; one used only the term encryption ("Encrypting your data"), and two with higher levels of complexity: "Encrypting your data. Encrypting your data ensures that only your intended recipients can read your data." and "Encrypting your data using a digital key. Others require this key to read your data, and we made sure that only your intended recipients know it."

4.4. Creation of final visual representations by designer

We used the online platform Fiverr.com to find a designer to create the visual representations. We used the same color scheme for all visual representations for consistency and went through one additional iteration with the designer to simplify and standardize the visual representations, while closely representing the experts' ideas. Table 3 (appendix) shows how the expert visual indicators evolved through the stages of this study.

5. Study 2: Vignette experiment with non-experts

In the second study, our objective was to evaluate how well the experts' ideas communicated encryption to non-expert users, addressing research questions 2 and 3. Anyone who has not received formal training or work experience in information security or cryptography is considered a non-expert for the purpose of our study.

We tested all the combinations of textual and visual representation brought forward by the experts in a vignette experiment. Vignette experiments combine the advantages of survey and experimental research (Auspurg & Hinz, 2015). Respondents are typically presented with descriptions of hypothetical scenarios, which are experimentally manipulated by the researcher. The method is extensively applied to study normative judgements and behavioral intentions (Wallander, 2009). The experimental design allows achieving high internal validity because the variation in the observed outcome variables can be solely attributed to the experimental manipulation of vignette characteristics. Moreover, the vignettes are assigned randomly to respondents, thus the effect of vignette characteristics on outcome variables should be independent from respondent characteristics. Using a vignette experiment allows us to provide causal evidence regarding the relationship between visual and textual representations of encryption and our three dependent variables (perceived security, UX, and understanding).

5.1. Research design

To test the effect of the visual and textual representations of encryption on our outcomes of interest, we conducted an online vignette experiment in February 2021 which was approved by our institution's ethics board. Our experiment considers two contexts where security concerns are highly relevant: e-voting and online banking. We investigated the impact of these combinations on people's perceived security, UX and understanding of the security mechanism encryption used mainly for confidentiality during data transmission.

Participants were randomly assigned to either the e-voting version of our survey, or the online banking version (split-half experiment, see Fig. 3). Within each context, after providing informed consent, the participants were shown a series of images of smartphone screens aimed at helping them envision being in the specific scenario (i.e., having to make a bank transfer or voting for a candidate).

Our vignettes exhibited the encryption part of the data transmission process in each context. They were integrated as one image of a smartphone screen into the series of smartphone screens. The vignettes varied experimentally in the values of two dimensions: the visual and textual presentation of encryption. Both dimensions were based on the expert productions in study 1. Table 1 provides an overview of the values of textual presentation of encryption and Table 2 provides the same information regarding the visual presentation of encryption. "No text" in Table 1 means that in this condition no textual representation was displayed. Instead, the visual was presented on its own unless they were assigned to the vignette that combined no text and no visual (i.e., where neither a text nor visual was shown). This condition thus represents a common case in current smartphone applications, where no visual indicators of encryption are shown to users. The respondents who were assigned to this vignette were shown the series of smartphone screens without the vignette. Instead of displaying the vignette, the confirmation screen was directly shown to participants. We used this condition as our control condition. Fig. 4 shows an example vignette with the combination Text ID 4 und Visual ID 2. The experimental design resulted in 24 (4×6) vignettes, representing all possible combinations of visual and textual representations of encryption. We employed a between-subjects design. Each participant was exposed to one randomly assigned vignette only. Such an approach decreases the risk of the respondents detecting the objective of the experiment and avoids, for example, learning effects. In each context, this vignette was followed by an image of a smartphone screen confirming the success of the interaction.

After participants had looked at all of the images of smartphone screens, we then asked them to rate (1) the perceived security, (2) the UX of the simulated interaction (3), and their understanding of the security mechanism encryption used mainly for confidentiality during data transmission. We provide the full questionnaire as Supplemental Material.

5.1.1. Measurements

5.1.1.1. Perceived security. We measured perceived security ("How secure or insecure did this experience feel to you?") on a scale of 1 (not secure at all) to 10 (very secure).

5.1.1.2. User experience. We evaluated UX with the 8-items short version of the UEQ, the UEQ-S (Schrepp et al., 2017) which measures UX in two dimensions: pragmatic quality and hedonic quality. Each dimension is measured with a 7-point semantic differential scale with four items. Pragmatic quality of experience is measured with the differentials obstructive/supportive, complicated/easy,

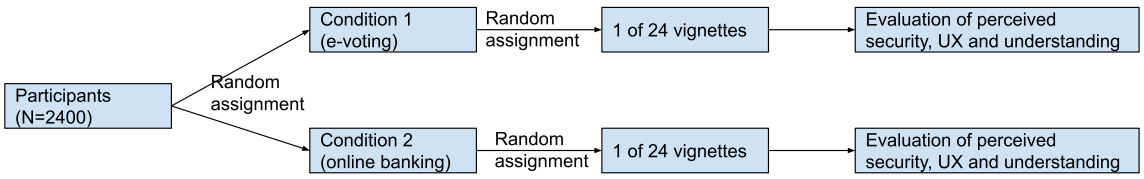


Fig. 3. Overview of the study design (not shown: demographic questions).

Table 1
Values of the experimental variable: textual representations of encryption.

Text ID 1	*No text*
Text ID 2 (technical term "encryption")	Encrypting your data.
Text ID 3 (lower complexity)	Encrypting your data. Encrypting your data ensures that only your intended recipients can read your data.
Text ID 4 (higher complexity)	Encrypting your data using a digital key. Others require this key to read your data, and we made sure that only your intended recipients know it.

inefficient/efficient, confusing/clear. Hedonic quality is measured with the differentials boring/exciting, not interesting/interesting, conventional/inventive, usual/leading edge. For our analysis, we generated two mean value indices representing the two dimensions.

5.1.1.3. Understanding of the security mechanism of encryption. Finally, we used an exploratory measure of understanding of the security mechanism encryption resulting from study 1. Participants rated how much they agreed or disagreed with the single items on a 5-point scale for the following question items. Since it is an exploratory measure, respondents were given the option "not sure":

- The connection is protected so that hackers cannot steal the data I'm sending.
- I am using a secure communication channel.
- Even if someone steals the data that I am sending, they won't be able to see what it means.
- Nobody can impersonate me unless they know my digital key.
- Nobody can see what I am sending without holding my digital key.
- My actions on the application are not revealed by someone listening in on the channel.

We reversed the scale such that higher values on the 5-point scale meant more agreement or, in other words, that participants thought that the interaction was secured by encryption. We generated a mean value index based on the six items. Observations for the option "not sure" were counted as missing values. For respondents with missing values, the mean understanding was calculated based on the items for which valid information was available. However, we applied weights that assign higher values to respondents who rated all six items (i.e., either agreed or disagreed to all six items) when constructing the index. For instance, a participant who answered all six items without selecting "not sure" would be given the full weight of 1, a participant who answered "not sure" on three out of six questions would be given a weight of 0.5. Respondents with missing values (i.e., not sure) on all six items were assigned the weight of 0 and thus were excluded from the analysis (N = 99). We used this weighted mean value index to measure understanding of encryption in our analysis. This approach allows us to use as much of the available information as possible without unnecessarily reducing the sample size.

To further assess the quality of our measurement of understanding, we asked participants what they thought the "understanding" questions meant in an open-ended question ("In your own words, what was the question above about?").

5.1.2. Recruitment and participants

We invited a sample of 2400 participants from Prolific who were based in the UK. Prolific allows researchers to recruit potential participants according to specific selection criteria. Participants are notified through the recruitment platform once they are eligible to take part in a research study. In terms of recruitment criteria, we did not specify constraints regarding gender, education or other factors. The included participants were notified automatically and redirected to the survey. The sample was non-representative. Note, however, that a representative sample is not necessary to achieve internal validity with experimental data. We recruited 2400 participants with the objective of obtaining 50 answers per vignette in both contexts (2*50*24 = 2400), which the literature suggests as the rule of thumb to obtain enough statistical power (Auspurg & Hinz, 2015). We excluded anyone who had participated in pre-tests of the study. The data collection period took one day in early 2021.

In total, 2457 respondents started the survey and 2417 completed the survey (i.e., answered all questions). We excluded respondents who had previously worked or studied in a field related to cybersecurity from further analysis. We also excluded respondents who did not pass the attention check questions. For any participants who filled out the survey twice (presumably by saving the link to our survey), we excluded their second participation from our analysis and kept the first time they participated. Finally, we excluded respondents who had not answered all the relevant questions from further analysis (i.e., who dropped out before answering the questions related to our dependent variables). Our analytic sample included 2180 participants, of which 1087 were randomly assigned to the context of e-voting, and 1093 were randomly assigned to online banking.

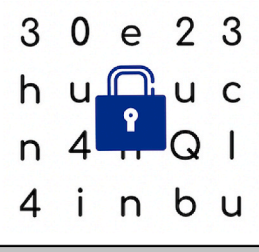
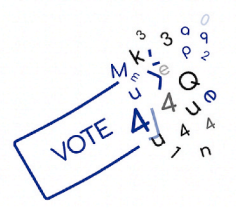
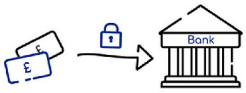
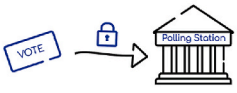
The participants were 68.8% women, 30.7% men, the remainder being non-binary and a gender that was not listed (0.5%). Participants were 38 years old on average (SD = 12.5). Around 55% of the respondents have a university degree (Bachelor or higher).

5.1.3. Experimental data

Since we employed a between-subjects design, our data comprises 2180 vignette ratings from 2180 respondents. A Chi-Square Test of Independence between vignettes and context revealed a non-significant result, suggesting that the split-half experiment worked. On average, each vignette was evaluated 45 times (e-voting: 45 times; banking: 46 times). Tables 4 and 5 (appendix) show that all bivariate correlations between the values of our two vignette variables are close to zero ($r < 0.1$) and not statistically significant, ensuring efficient estimation. Similarly, all correlations between the values of our vignette variables and key observed respondent characteristics (education, age) were close to zero ($r < 0.1$) and not statistically significant, indicating that the randomization worked. The only exception is respondent gender, of which single values correlated significantly with one value of visual representation, but these correlations were also close to zero (see Tables 4 and 5). We performed robustness checks to test the influence of respondent characteristics on our findings (see section 6).

In both contexts, respondents used the whole answer scale for the dependent variables and the distribution of ratings was left-skewed for perceived security, UX (pragmatic quality), and the weighted index for understanding, thus tending towards more positive values on the respective scales (see Figs. 9–12, and 15-16 in the appendix). Hedonic quality of UX was symmetrically distributed in both contexts (Figs. 13

Table 2
Values of the experimental variable: visual representations of encryption.

	Visual ID 1	Visual ID 2	Visual ID 3
Online banking	*No visual representation*		
e-voting	*No visual representation*		
ID	**Visual ID 4**	**Visual ID 5**	**Visual ID 6**
Online banking			
e-voting			

and 14 in the appendix).

5.1.4. Data analysis

To analyze our experimental data, we conducted Ordinary Least Squares (OLS) regressions using robust standard errors to account for heteroskedasticity. We estimated separate models for our dependent variables (UX pragmatic, UX hedonic, perceived security, and understanding) and the two contexts. We first estimated the overall effect of textual and visual representation of encryption on each of our dependent variables. We then estimated the effects of the single values of textual and visual representation for each dependent variable in a second model. If we found statistically significant effects, we tested whether the effects of textual and visual representation varied between contexts in a third set of models. These models were conducted based on the full sample and including an interaction term between the variable indicating the context and the variables indicating textual and visual representation. We performed several analyses to assess the robustness of our findings (see our discussion in section 6). We provide these additional analyses as supplemental material.

Regarding the qualitative analysis, we categorized all qualitative answers about what participants thought the understanding questions were aimed at. Once the initial codebook was created, we conducted a test session with 8 HCI experts, who applied the codes to a subset of 400 answers. They commented on any codes they thought were unclear and suggested improvements, which we used to update the codebook. Using the updated codebook, we then conducted a double-coding session with an HCI expert who double-coded the answers from 250 participants (11% of answers). Since there was a large number of codes and potential combinations, the probability of agreement by chance was low. We thus used a simple measure of percentage agreement. We defined agreement between coders as the exact same combination of codes. For the questions assessing qualitative answers to the understanding question, the two coders achieved an agreement of 86%.

We provide the syntax files used for analysis and the data (with potentially harmful meta data removed), as well as our annotated analysis, as supplementary material.

Fig. 4. Sample vignette, combination of Text ID 4 and Visual ID 2.

5.2. Results

5.2.1. Bivariate correlations between dependent variables

We found a statistically significant and positive correlation between UX and perceived security in both the context of e-voting (pragmatic, r = 0.40; hedonic, r = 0.36; 5% significance level) and online banking (pragmatic, r = 0.46; hedonic, r = 0.29; 5% significance level). Thus, higher values on UX mean higher values on perceived security in both contexts.

There is a statistically significant and positive correlation between perceived security and understanding in both contexts (e-voting, r = 0.56; banking, r = 0.40; 5% significance level), meaning that the better the understanding, the higher the perceived security.

Overall, the size of the correlation is moderate suggesting that our three dependent variables capture distinct dimensions of the interaction.

5.2.2. Experimental evidence

5.2.2.1. Perceived security. Table 6 (appendix) shows the results of OLS regressions predicting perceived security regarding the overall effects of text and visual representation. In both contexts, we observed a positive and statistically significant overall effect of text representation (compared to no text) on perceived security. The effect was highly significant ($p < .001$) in the banking context and significant at the 5%-level in the e-voting context. When looking at the single values of text representation (see Table 7), "lower complexity" and "higher complexity" showed statistically significant and positive effects in the banking context (both $p < .001$). Fig. 5 shows the results graphically. Lower complexity text increased perceived security by almost one scale point (0.74), similarly, high complexity text increased perceived security by 0.70 compared to no text. However, the differences between the two effects was not statistically significant. Although we observed the same pattern in the e-voting context, the effect sizes were slightly smaller than in the banking context. Moreover, we only found a statistically significant effect of high complexity (5%-level). The values of the visual representation of encryption showed relatively small effects on perceived security, which were not statistically significant in both contexts.

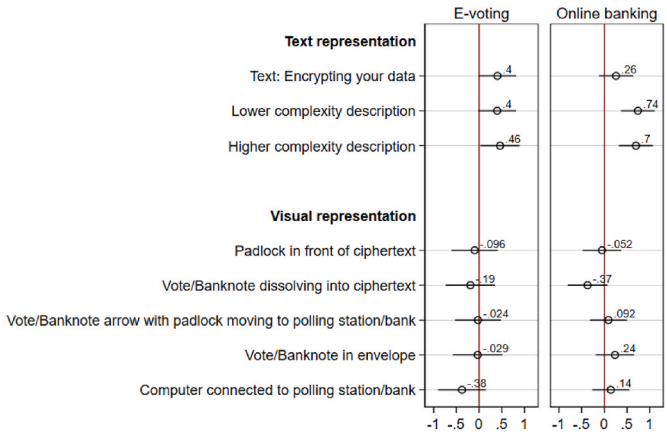


Fig. 5. Coefficient plot: single effects of vignette values on perceived security. N = 1087 in e-voting, N = 1093 in online banking.

Table 8 (appendix) shows the results of a regression model including an interaction effect between context and textual representation. The interaction effect suggested slightly more positive effects of lower and higher complexity in the banking contexts compared to the e-voting contexts, but was not statistically significant. Thus, our results do not suggest substantial differences in the effects of text presentation between the two contexts. Since we found no substantial and significant effects of visual presentation in any of the two contexts, we did not estimate a model including an interaction of visuals and context.

In summary, we found evidence that textual representation of encryption increases perceived security while visual representation has no effect.

5.2.2.2. User experience (UX)

5.2.2.2.1. Pragmatic quality of user experience (UX-PQ). In both contexts, the overall effect of text and visual presentation were close to zero and not statistically significant (see Table 9 in the appendix). We observed similar results regarding the effects of the various versions of text and visuals on UX-PQ (see Table 10). The effects were relatively small and not statistically significant (see also Fig. 6, which shows the results graphically). Some of the effects of the versions of visual presentation showed a negative sign, suggesting a decrease in UX-PQ. In both contexts, none of the observed effects were statistically significant. The only exception is the padlock in front of ciphertext (visual ID 2),

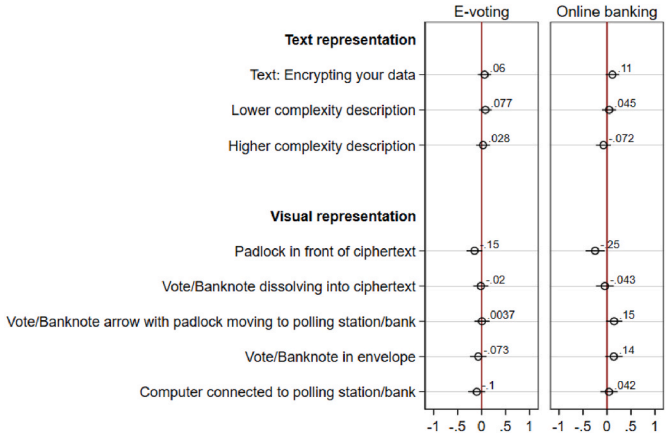


Fig. 6. Coefficient plot: single effects of vignette values on UX-PQ. N = 1087 in e-voting, N = 1093 in online banking.

which had a statistically significant negative yet small effect on UX-PQ in the context of banking ($p < .01$).

5.2.2.2.2. Hedonic quality of user experience (UX-HQ). Regarding the overall effects of text and visual presentation, we observed relatively small and non-significant effects in both contexts (see Table 11). Similarly, we observed relatively small and close-to-zero effects of the versions of text and visual presentation of encryption on UX-HQ in both contexts (see Table 12 in the appendix). Some of those showed negative signs, however, the effects were not statistically significant in most cases. We found a statistically significant and positive effect of visual representation ID 4 (vote/banknote arrow with padlock moving to polling station/bank) on UX-HQ in the context of voting ($p < .05$). We provide the coefficient plot for the versions of the text and visuals in Fig. 7.

Overall, we found little evidence suggesting that textual and visual representations impact UX, with two exceptions: padlock in front of ciphertext regarding UX-PQ in the context of banking, and vote/banknote arrow with padlock moving to polling station/bank regarding UX-HQ in the context of voting.

5.2.2.3. Understanding. Table 13 (appendix) shows the results regarding understanding of encryption. The textual representation of encryption had a statistically significant and positive overall effect on understanding of encryption in both contexts (e-voting: $p < .001$; online banking: $p < .001$). We found no statistically significant overall effect of visual representation and the effect was close to zero in both contexts. When looking at the single values of text (see Table 14 for the full model and Fig. 8 for the graphical presentation of results), Text version 3 (highest complexity) has the strongest positive effect (similar to our results regarding perceived security) in both contexts. In the context of e-voting, the difference between the effect of higher complexity and lower complexity as well as the simplest version "encrypting your data" vs. no text was statistically significant ($p < .05$ and $p < .001$, respectively). In the context of online banking, only the difference between higher complexity and the simplest text version as well as between lower complexity and the simplest version was statistically significant. The difference between the effects of higher and lower complexity was not statistically significant. The visual representation of encryption had no statistically significant effect on understanding in both contexts. All effects were relatively small.

Similar to our results regarding perceived security, the interaction terms between context and text was not statistically significant and rather small in both contexts, suggesting that the effect of text on understanding does not vary in a relevant way between the two contexts (Table 15).

In summary, we found evidence that a textual representation of the security mechanism encryption increases the understanding of encryption. In both contexts, more complex textual representations had the greatest influence, although we found no relevant differences between high complexity and low complexity, at least in online banking.

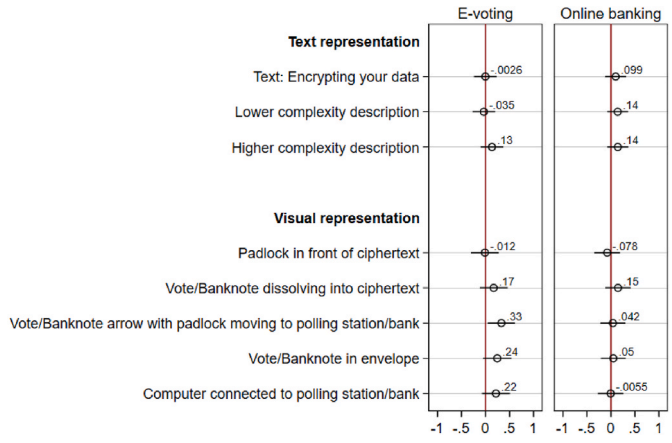


Fig. 7. Coefficient plot: single effects of vignette values on UX-HQ. N = 1087 in e-voting, N = 1093 in online banking.

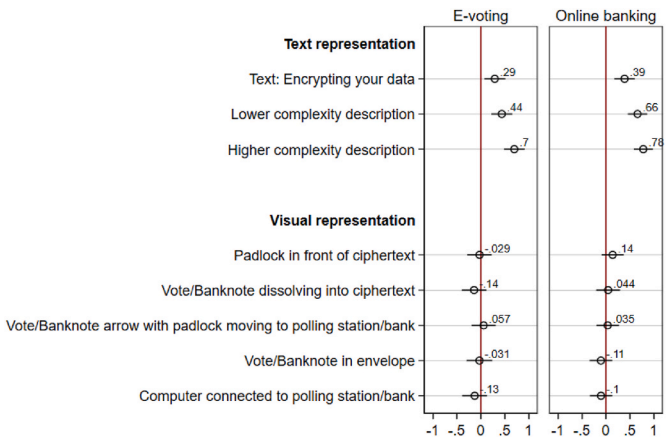


Fig. 8. Coefficient plot: single effects of vignette values on understanding. N = 1087 in e-voting, N = 1093 in online banking.

To synthesize, our results show that UX-PQ and UX-HQ are positively correlated with perceived security, as is our measure of understanding encryption. The textual representation of encryption had a statistically significant and positive overall effect on both perceived security and understanding of encryption in both contexts, with more complex versions of the text having a greater influence. The visual representation of encryption had no substantial or statistically significant effect on any of our dependent variables.

5.2.3. Results of qualitative analyses

We will now describe the qualitative results.

Since the "understanding" questions were exploratory, we asked participants what they thought these questions were about (commonly referred to as "face validity"). 65% of answers mentioned that they were about security in general, followed by encryption (19%) and hacking (14%), as well as authentication (3%), impersonation (3%) and fraud (2%).

Most of these concepts are closely related to encryption during data transfer, which can, for instance, provide confidentiality and protection from fraud, hacking or impersonation to a certain degree. The only concept which one can argue is not necessarily related to encryption during transmission is authentication which participants frequently related to login details. In the qualitative answers, we could see that participants who mentioned authentication seemed to mix up the "digital key" mentioned in the description of encryption with a password. We also explored the terms participants used to qualify these concepts. Participants mostly thought that the question aimed at exploring their feelings, knowledge, thoughts, understanding and perceptions.

Overall, these qualitative results show that our participants thought that our understanding questions measured concepts closely related to what we intended to measure, albeit they often expressed this more generally as a notion of overall security.

6. Discussion

In this section, we reflect on the three research questions of this paper. First, we discuss our results regarding the expert co-creation study (RQ1). Second, we discuss the results from our vignette experiment (RQ2 and RQ3). Next, we discuss the exploratory measure of understanding brought forward by our experts, and reflect on its usefulness. We also discuss the limitations of the present work and

suggest directions for future research.

6.1. Generating ideas for visual and textual representation of encryption using a multidisciplinary panel of experts

The results from study 1 answered RQ1 and provided us with the elements needed to inform study 2. Expert insights are used in a variety of studies in useable privacy and security, for instance in order to compare their behaviors with non-expert behaviors (Busse et al., 2019; Ion et al., 2015), their security perceptions compared to non-experts (Gallagher et al., 2017), or to compare the security concerns experts had as compared to non-experts (Murillo et al., 2018). In this article, we used a different approach and did not compare expert and non-expert behaviors, perceptions or understanding. Instead, we recruited a mix of security and HCI experts and asked them to generate ideas in an iterative co-creation process. We found this approach helpful, in particular by asking the experts to build upon the earlier experts' ideas, which encouraged them to go beyond the ideas that first came to their mind. This approach differs from a recent study using co-design methodologies with non-expert users (Fassl et al., 2021). The authors highlight that the initial framing of the security threat and task heavily influenced participants' ideas for solutions. This difficulty when using co-design methods for displays of technical security with non-experts led us to avoid using the non-experts to come up with ideas of how to display encryption as we would have needed to explain what encryption is first. However, explaining encryption to non-experts is a non-trivial task and there is limited work on how to best do this. This paper makes a contribution to this gap. Both our approach and co-design methodologies with non-experts seem suitable to put the user at the centre in the design of user-centered displays of security, but our iterative approach of combining expert knowledge from multiple domains (study 1), followed by a an evaluation with non-experts (study 2) might be more suitable for technical topics where co-design with non-experts is initially difficult since they are not familiar with the subject matter and empirical guidance on how to create a common frame of reference is lacking.

6.2. Putting expert ideas to the test: experimental results of the effects of visual and textual representations on dependent variables

The results from study 1 informed study 2, which addressed RQ2 and RQ3 in a vignette experiment. We found that the visuals had no statistically significant effect on any of our dependent variables, while the version of the text had a statistically significant effect on perceived security and understanding. We might have expected the visual representations to be "intuitive" ways of displaying the security mechanism of encryption to users, not requiring any reading and able to convey a "great deal of information concisely" (Blattner et al., 1989, p. 12). However, our study does not confirm such assumptions that visuals are necessarily more intuitive ways of displaying information. Previous work frequently measured the effects of icons and text in terms of observed measures such as task completion times or error rates. For instance, Huang et al. (2019) compared two experimental groups of older adults, one of which interacted with an ATM interface that only used text, and one of which used an interface that combined icons and text. Task completion (measured in terms of use of the help button and number of steps required to complete a task) was better for the participants in the group that saw both icons and text, although effect size remained relatively small. Similarly, Majrashi (2020) found that combining text with icons in a smartphone menu led to faster task completion times and fewer mistakes. Both studies did not measure any subjective indicators of experience, as was done in this study. In other studies that included self-report measures of experience, the combination of visual representation with text labels, as well as text-only led to better learnability and ease of use (Wiedenbeck, 1999). In work on the visuals representing privacy choices, it was also necessary to add a text to the icon for research participants to understand their meaning (Cranor, 2021). Note that, based on these studies, one might have expected the visual indicators to have a positive effect in our study when combined with textual indicators, but this was not the case – even when combined with textual indicators, the visuals had no statistically significant effects. Our results were however in line with research that found that textual indicators have a positive effect on user perceptions (Cranor, 2021; Wiedenbeck, 1999).

We can hypothesize on the reasons why there was no significant effect in our study. All of our visuals were novel to users since they were based on the expert iterations in study 1. This novelty might require participants to engage in greater mental efforts to process the visuals which have no previously assigned meaning. Indeed, previous work has found familiarity to be a relevant factor for the guessability of physical safety warning signs, for instance (Chan & Ng, 2010) and is generally considered relevant for the speed and accuracy with which icons and objects can be identified (McDougall et al., 2016). Wogalter et al. (2006) also describe the different symbol-to-concept relationships, from representational symbols that directly or closely relate to the represented concept, to more abstract or arbitrary symbols, with a more distant relationship to the concept. A sign with a crossed out match would be an example of a representational symbol, directly displaying the meaning of "do not light a match". In our case, such direct representation was not possible as encryption does not have an equivalent, well-known real-world concept that could be visualized to represent encryption. The digital processes represented by the visuals are not always familiar to non-experts, for instance the concept of data transmission in visual IDs 4 and 6. Also note that many of these studies investigate performance measures such as number of errors participants make or completion time. Self-report measures with a focus on variables such as UX and perceived security are comparatively rare, we cannot exclude that the tested visual representations might have an effect on measured variables that were out of scope of this study.

In our study, more complex text had a positive effect on understanding of the security mechanism encryption and at least in the online banking sector also on perceived security. Considering that more complex text introduces friction to the interaction by introducing additional information and an additional step compared to our control condition, our work lends some empirical support to work arguing that introducing some friction into experiences may create more mindful experiences (Cox et al., 2016). Recent work also made an argument for "security-enhancing friction", friction that encourages users to behave more securely (Distler, Lenzini, et al., 2020). The friction introduced through the descriptions of encryption can be seen as friction that helped improve the understanding of encryption, which is in itself a positive result for the security of our users. Of course, our work does not allow us to make statements about behaviors.

6.3. The challenges of creating an exploratory measure of understanding of encryption

In our studies, we created and used an exploratory measurement of understanding for encryption. We asked experts which questions they

might ask non-experts to evaluate whether they had understood that encryption was being used, upon which we iterated twice with other security experts. We then used these question items in study 2 as an exploratory measure of understanding. Our qualitative analysis shows that these questions were mostly perceived by non-experts as measuring security in general or encryption. The answers suggesting that the items measured security in general did not provide any details about the security mechanism providing the security, but they seemed to understand the general implication of providing protection to some degree. One possibility for these results is that participants lack the necessary vocabulary to associate our six items with encryption and therefore associate these items with the more familiar term security. However, it might also be the case that the six items capture security perceptions in addition to understanding. Also, the answer option "not sure" was used relatively frequently, although no question item seemed to stand out in terms of difficulty to provide an answer (approximately 20% of participants for each question item). These ratings could indicate that the participants did not understand the question, or they might have understood the question, but were not sure about its answer. For these respondents, we might have over- or underestimated understanding of encryption. As a robustness test for our weighted mean value index of understanding, we generated an index excluding all observations for "not sure" and re-estimated our models using this index as a dependent variable. These analyses did not reveal substantial changes in our findings. We provide this additional analysis as supplemental material.

6.4. Limitations and future work

Our study has some limitations and open questions for future work remain.

6.4.1. Visual and textual representations

There are some limitations related to the visual representations we used. The visual indicators we evaluated were closely based on the HCI and security expert ideas and were not redesigned by an icon designer following guidelines for icon design. A previous study compared crowdsourced security indicators by non-experts with designer-drawn icons. In their evaluation, the crowdsourced indicators performed no worse, and sometimes better than the designer-drawn icons (Egelman et al., 2015), providing some support to our approach. However, future work could redesign the icons following icon design guidelines and evaluate the effectiveness. We also tested the vignettes in the particular context of a smartphone interaction, a context for which the visual representations may have included more details than is typical in such interactions. While the visuals did not have a significant effect in this context, we cannot exclude that they might have positive effects in, for example, a tutorial setting aimed at teaching non-experts about encryption, a potential avenue for future studies. We also did not test animated designs, which is an open question for future research. Future work could address the effect of familiarity with visuals on user perceptions, for instance using eye tracking to investigate how fast people are able to react to the visuals, and whether they react more efficiently to indicators that are commonly used.

Some limitations need to be acknowledged regarding the textual representations. Our study focused not only on textual representations of encryption in general but also the degree of complexity of textual representations. Complexity was defined based on technical concepts introduced in each version of text, but other characteristics of textual representations might have an effect on our outcomes of interest. For example, future studies could explore the impact of text length in addition to the mentioning of technical concepts. In our study, more complex text provided more details on the ongoing process, which made more complex text longer. Thus, we cannot clearly separate the effect of text length and technical terms. Also, the number of technical concepts in one text might additionally play a role, which could be assessed in future work.

Overall, a promising result of our study is that complex, carefully designed descriptions of encryption had a statistically significant effect on perceived security and understanding. We hope to see more work in the future on how to design text that describes technical security concepts to non-experts in a user-centered way.

6.4.2. Generalizability

A potential limitation of the present work concerns the generalizability of our results to real-world interactions with technology. Our participants were encouraged to pay close attention and might have paid less attention in a real-life context. Thus, we might have overestimated the effect of textual representations on our dependent variables in the vignette experiment. Nevertheless, as discussed, our results are in line with previous studies finding an effect of textual representations on perceptions and/or performance. It would be relevant for future studies to implement varying representations of security mechanisms in real-life use contexts, where participants might pay less attention to the details of a smartphone application, and compare the results to our outcomes. Also, the generalizability of our results is further limited to the textual and visual representations used in our design (including the general layout of our vignettes such as color), but other relevant combinations of text and visual representations might exist. These could be assessed in future research.

6.4.3. Measuring understanding of encryption

For our exploratory measure of understanding, we, as well as our experts, found it challenging to define what level of understanding of such a technical concept we could expect non-experts to have. A challenge of measuring understanding of encryption is to make sure that the wording of the questions stays sufficiently non-technical for non-expert users, but at the same time measures the intended concept. Given that understanding of encryption constitutes a relevant concept for many security-relevant interactions, future work should continue iterating upon our exploratory items. For example, although we had conducted qualitative pre-tests of the questionnaire, more extensive qualitative investigation of the "understanding" items should reveal the reasons for participants' frequent selection of "not sure" as an answer. Overall, we think that the items were a useful first step in measuring general understanding of encryption, but we acknowledge the exploratory nature of our measurement and that further research is needed to validate and further develop this study's measurement of understanding.

6.4.4. Theoretical concepts

Our study also leaves some open questions on a theoretical level. Indeed, typical models of UX (Hassenzahl, 2008; Mahlke, 2008) and instruments assessing UX (Hassenzahl et al., 2003; Laugwitz et al., 2008) do not include indicators for understanding of underlying processes or perceptions of security. While psychological need theories include the need for security as drivers of satisfying events (Sheldon et al., 2001), assessment is relatively broad and thus difficult to apply in the field of useable privacy and security. But of course, the field of useable privacy and security has long extended beyond the concept of usability and

includes a broad scope of research; for instance aiming to improve user understanding and perceptions of security (Abu-Salma et al., 2018; Distler et al., 2019; Spero & Biddle, 2020) or applying co-design methodologies for security processes (Fassl et al., 2021). In the future, it would be relevant to see work theorizing on the links between UX and useable privacy and security, reflecting on the extent to which the broad range of issues addressed by the field of useable privacy and security can be addressed under the umbrella of UX. The field would further profit from empirical work assessing the relationship between the concepts of understanding, user experience and understanding, strengthening our theoretical knowledge of user perceptions in the context of security-relevant interactions.

7. Conclusion

There is an ongoing debate whether security mechanisms should be visible or hidden away from users. User-centered design typically aims to let users complete their tasks as easily and quickly as possible (Krug, 2000), leading to many security mechanisms being hidden away from the user, who thus have no indication they are happening in the background. This lack of visibility can backfire when users lack understanding of security processes, potentially leading to security issues (Adams & Sasse, 1999) and leaving users unable to form accurate mental models of the security of a system (Spero & Biddle, 2020). Authors have thus argued that security should be highly visible and ready to be inspected by users (Adams & Sasse, 1999).

Our study brings empirical evidence to the ongoing discussion "should security mechanisms be visible or hidden away from users" by answering two main research objectives. First, we addressed the question of how HCI and security experts suggest displaying encryption to non-expert users using textual and visual representation, using an iterative co-creation process (see section 4). Second, we wanted to understand what the effects of the resulting visual and textual indicators are on perceived security, user experience and understanding, comparing two use contexts: e-voting and online banking. To this end, we conducted an online vignette experiment with non-expert users to test the effect of the representations on our outcomes of interest (see section 5).

In summary, the textual representation of encryption significantly increased both perceived security and understanding of encryption in both use contexts. More complex text describing encryption resulted in higher perceived security and more accurate understanding. Representing encryption through text thus seems to be a promising solution to improve understanding and improved security. Overall, we found little differences in our results between the two use contexts. We found no statistically significant or substantial effect of textual representations on UX. Finally, visual representations of encryption had no statistically significant effect on any of our dependent variables.

Overall, our study contributes to the larger discussion regarding visible instances (including text and visuals) of security and the impacts they may have on user perceptions. Our study supports the hypothesis that more visible instances of security support more accurate understanding (Spero & Biddle, 2020), but also, perceived security. We also attribute this effect to the extensive design phase of the tested vignettes with a multidisciplinary panel of experts; as well as pre-tests that enabled us to improve upon any expert suggestions that participants perceived as confusing. We therefore interpret our results as an encouragement to carefully design and pre-test technical descriptions for improved understanding and perceived security in a user-centered way.

While the vignette experiment is a frequently used methodology to measure normative judgements, attitudes, and behavioral intentions in sociology (Wallander, 2009), to the best of our knowledge, it has rarely been applied to evaluate interface designs in UPS contexts (Al-Natour et al., 2020) or other HCI contexts (Vance et al., 2015). Our work demonstrates that this method can be applied to empirically evaluate details of interface design. Its strength lies in the results that give insights into the causal relationship between visual and textual design choices and outcome indicators, free from confounding factors.

Our results demonstrate the relevance of measuring the effects of user interface elements such as visual and textual indicators on facets of experience such as perceived security, UX and understanding. We hope that future work will provide more empirical research-based guidance on how displays of technical security might look when optimizing these user-centered indicators going beyond UX alone and including security perceptions and understanding.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank our research participants for their time and insights.

This work was supported by the Fonds National de la Recherche (PRIDE15/10621687).

Appendix A. Supplementary materials

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chbr.2021.100161.

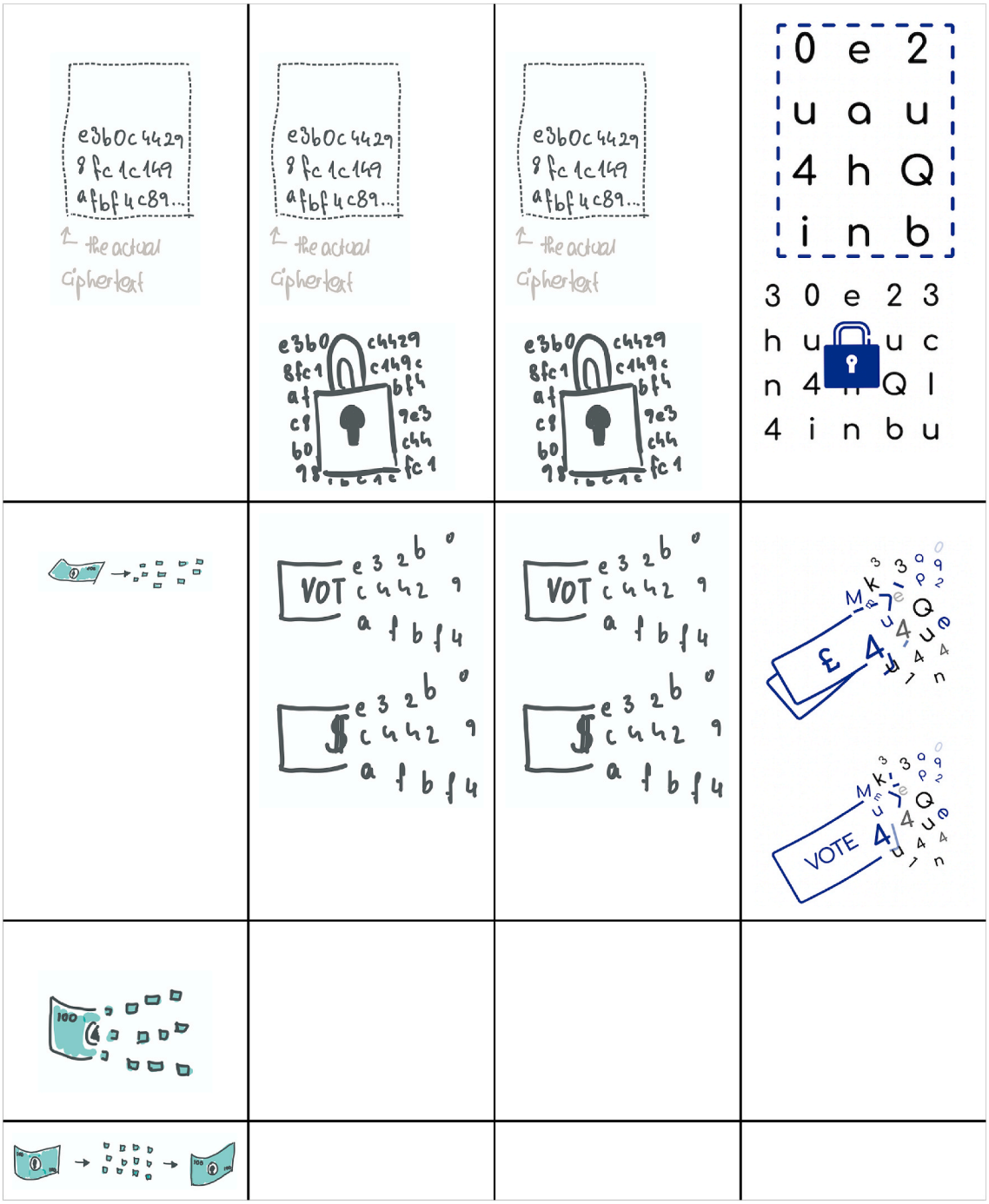
Appendices.

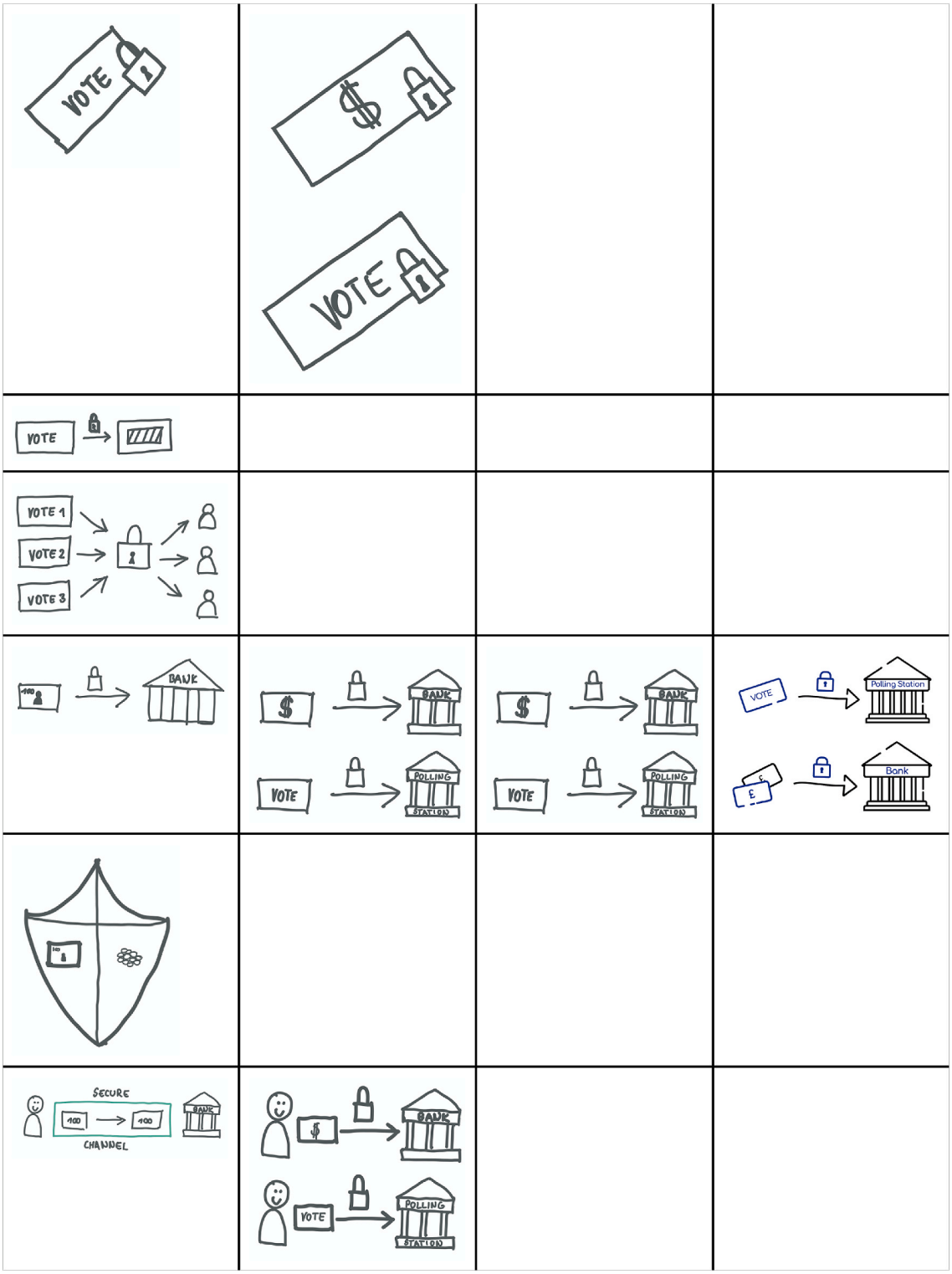
Expert co-creation – Additional Details

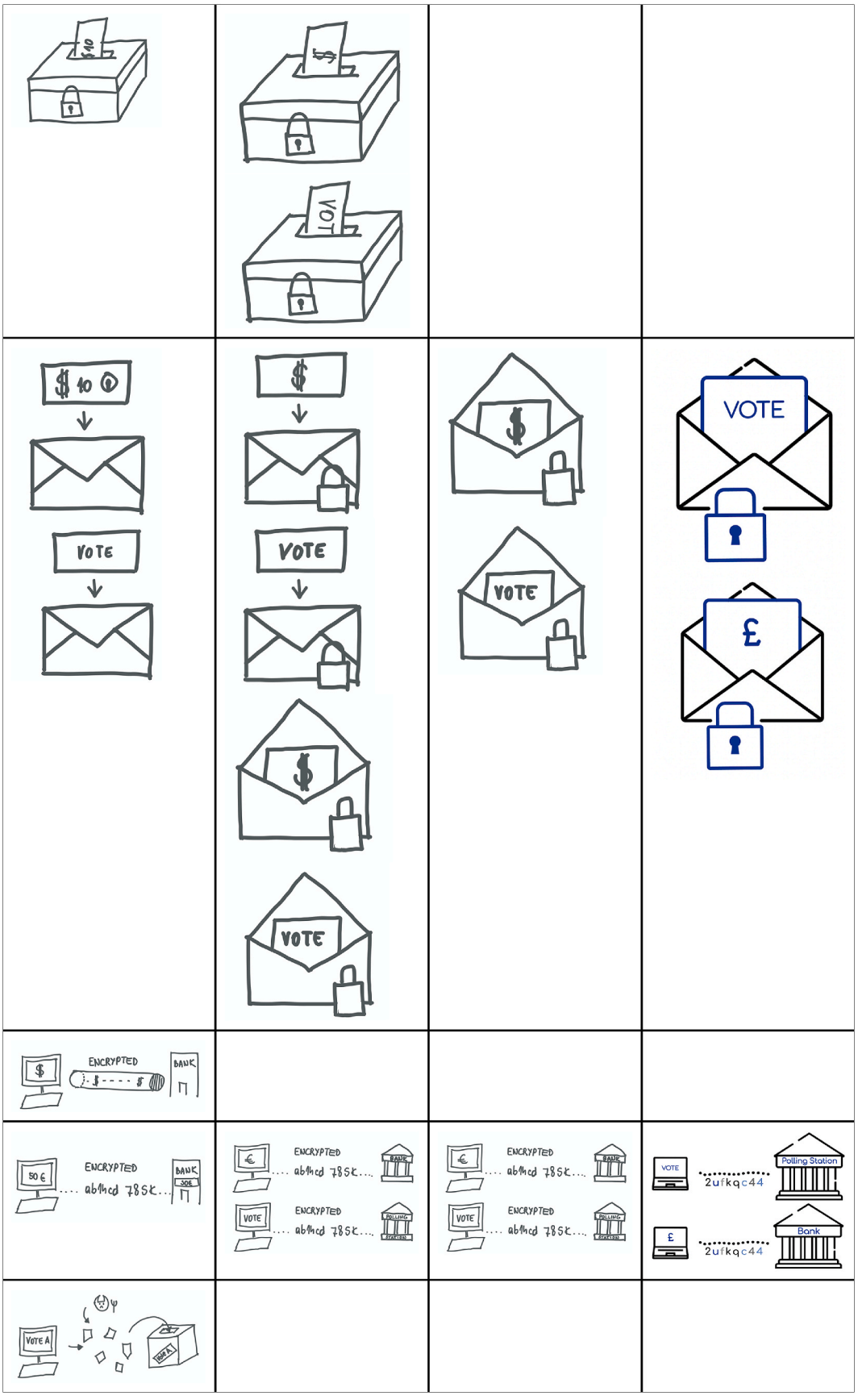
Table 3 summarizes the results from the expert ideation phase (1), the selection and improvement with HCI experts (2), the final selection by security experts (3), and the final visual representations that were created by a designer (4). The last column shows the visual representations that we used for the visual representations in the vignettes.

Table 3
Iterations of the visual representations of encryption

1) First ideation with security experts	2) Selection and improvement with HCI experts	3) Final selection of visual representations	4) Final visual representations by designer
			
			
			
			
			
			


e3b0c4429
8fc1c149
afbf4c89...
the actual ciphertext
0 e 2
u a u
4 h Q
i n b
3 0 e 2 3
h u u c
n 4 Q l
4 i n b u
VOTE
VOTE


VOTE
$
VOTE
VOTE
VOTE 1
VOTE 2
VOTE 3
BANK
$
BANK
VOTE
POLLING STATION
$
BANK
VOTE
POLLING STATION
VOTE
Polling Station
£
Bank
SECURE
100
100
BANK
CHANNEL
$
BANK
VOTE
POLLING STATION


$ to
VOTE
$
VOTE
VOTE
$
VOTE
VOTE
£
ENCRYPTED
BANK
50 €
ENCRYPTED
abfhcd 785t...
BANK
50€
ENCRYPTED
abfhcd 785t...
VOTE
ENCRYPTED
abfhcd 785t...
ENCRYPTED
abfhcd 785t...
VOTE
ENCRYPTED
abfhcd 785t...
VOTE
2ufkqc44
Polling Station
£
2ufkqc44
Bank
VOTE A

Correlations

Table 4
Bivariate Correlations of Vignette Values for Online Banking

	Text_1	Text_2	Text_3	Text_4	Visual_1	Visual_2	Visual_3	Visual_4	Visual_5	Visual_6	Male	Female	Non-binary	Gender not listed	Univ. education	Age
Text_1	1.00															
Text_2	−0.32*	1.00														
Text_3	−0.32*	−0.35*	1.00													
Text_4	−0.32*	−0.34*	−0.34*	1.00												
Visual_1	−0.07*	0.05	0.00	0.01	1.00											
Visual_2	0.02	0.04	−0.04	−0.03	−0.21*	1.00										
Visual_3	0.05	−0.06*	0.01	0.00	−0.20*	−0.19*	1.00									
Visual_4	0.05	−0.03	0.00	−0.02	−0.21*	−0.20*	−0.19*	1.00								
Visual_5	−0.04	−0.00	−0.02	0.06*	−0.21*	−0.19*	−0.19*	−0.19*	1.00							
Visual_6	−0.02	0.01	0.03	−0.03	−0.21*	−0.20*	−0.19*	−0.20*	−0.20*	1.00						
Male	−0.01	−0.02	0.04	−0.01	−0.01	−0.03	0.08*	−0.03	−0.02	0.02	1.00					
Female	0.01	0.02	−0.04	0.01	0.01	0.03	−0.09*	0.03	0.03	−0.01	−0.99*	1.00				
Non-binary	−0.00	−0.01	−0.01	0.02	0.00	0.01	0.05	0.01	−0.03	−0.03	−0.05	−0.10*	1.00			
Gender not listed	−0.02	−0.02	−0.02	0.05	−0.01	−0.01	0.07*	−0.01	−0.01	−0.01	−0.02	−0.04	−0.00	1.00		
Univ. education	0.00	−0.04	−0.00	0.04	−0.02	0.00	0.03	0.03	−0.04	−0.00	−0.05	0.05	−0.02	0.03	1.00	
Age	0.00	0.01	−0.01	−0.00	0.02	−0.01	0.03	0.00	0.02	−0.06*	0.08*	−0.07*	−0.03	−0.00	−0.04	1.00

Pairwise correlations between values of vignette variables (as dummies) and respondent gender (as dummies for each category), university education (Bachelor degree and higher as dummy), and age (continuous). Pearson's correlation coefficient. *p < .05.

Table 5
Bivariate Correlations of Vignette Values for E-Voting

	Text_1	Text_2	Text_3	Text_4	Visual_1	Visual_2	Visual_3	Visual_4	Visual_5	Visual_6	Male	Female	Non-binary	Univ. education	Age
Text_1	1.00														
Text_2	−0.33*	1.00													
Text_3	−0.35*	−0.34*	1.00												
Text_4	−0.33*	−0.32*	−0.33*	1.00											
Visual_1	−0.04	0.02	−0.00	0.03	1.00										
Visual_2	0.06	−0.01	−0.02	−0.03	−0.20*	1.00									
Visual_3	−0.02	−0.02	0.04	0.00	−0.20*	−0.21*	1.00								
Visual_4	−0.00	0.06*	−0.03	−0.03	−0.21*	−0.22*	−0.22*	1.00							
Visual_5	0.01	−0.00	0.01	−0.01	−0.18*	−0.19*	−0.19*	−0.20*	1.00						
Visual_6	−0.00	−0.04	0.00	0.04	−0.19*	−0.20*	−0.20*	−0.21*	−0.18*	1.00					
Male	0.06	−0.02	0.01	−0.05	−0.05	−0.03	−0.01	0.08*	0.01	−0.00	1.00				
Female	−0.06	0.02	−0.00	0.05	0.05	0.03	0.00	−0.08*	−0.01	0.01	−0.99*	1.00			
Non-binary	−0.01	0.02	−0.01	−0.01	−0.03	−0.03	0.04	0.00	0.05	−0.03	−0.04	−0.10*	1.00		
Univ. education	−0.01	−0.00	0.05	−0.04	−0.03	−0.00	0.03	0.03	0.02	−0.04	−0.05	0.04	0.06	1.00	
Age	−0.02	0.08*	−0.05	0.00	0.02	−0.02	0.01	0.07*	−0.06	−0.02	0.09*	−0.08*	−0.03	−0.06*	1.00

Pairwise correlations between values of vignette variables (as dummies) and respondent gender (as dummies for each category), university education (Bachelor degree and higher as dummy), and age (continuous). Pearson's correlation coefficient. *$p < .05$.

Distribution of rankings

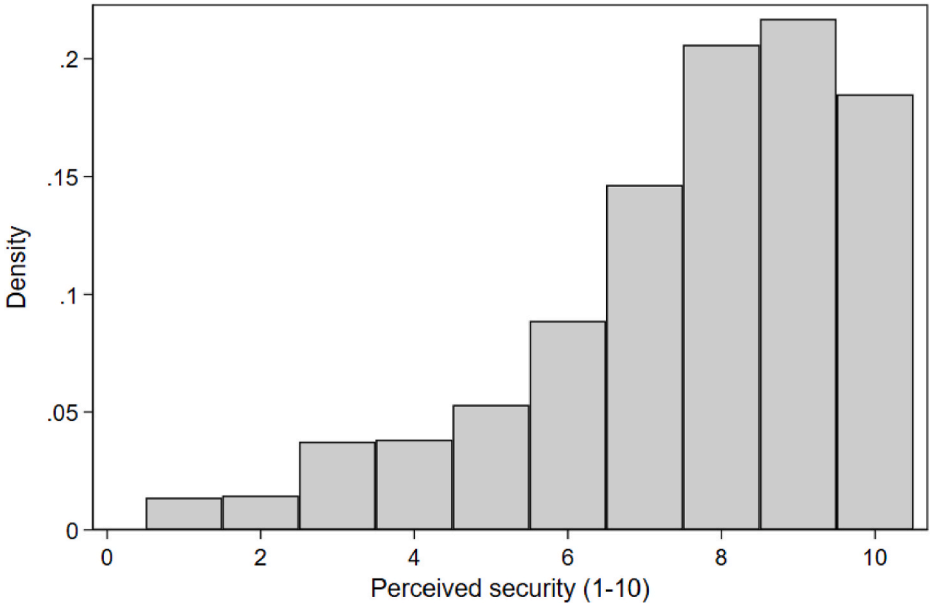


Fig. 9. Distribution of rankings of perceived security in online banking context. N = 1093.

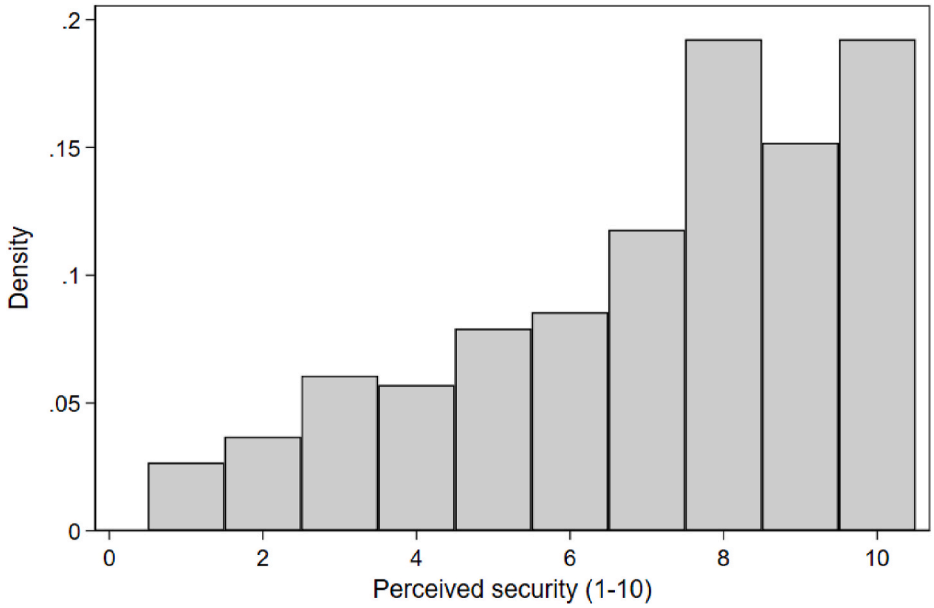


Fig. 10. Distribution of rankings of perceived security in e-voting context. N = 1087.

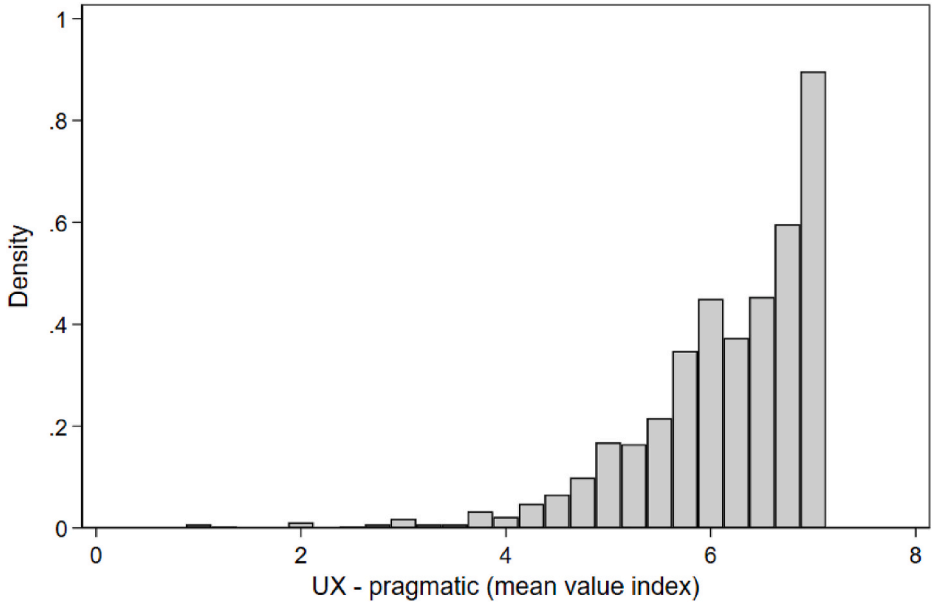


Fig. 11. Distribution of rankings of pragmatic quality of UX in online banking context. N = 1093.

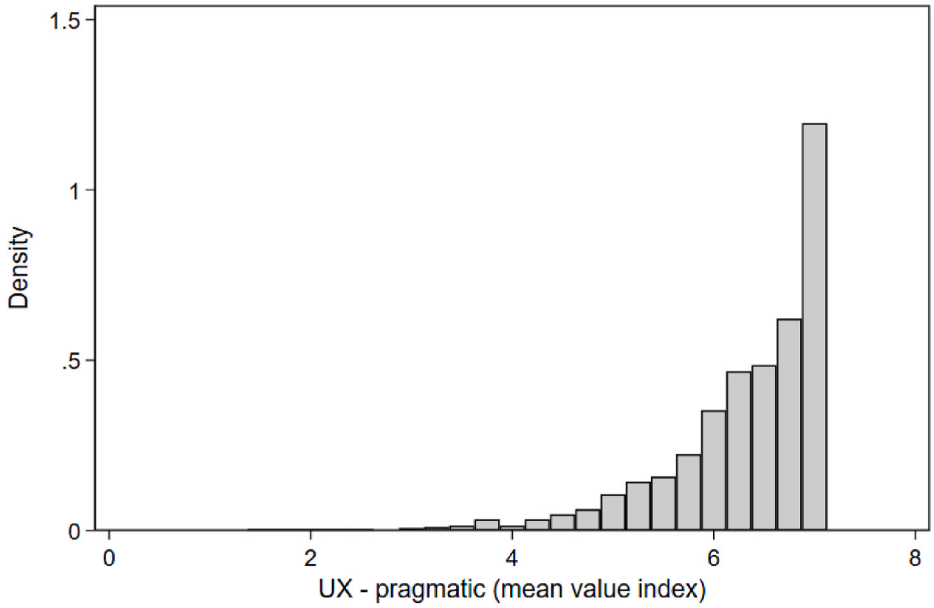


Fig. 12. Distribution of rankings of perceived pragmatic quality of UX in e-voting context. N = 1087.

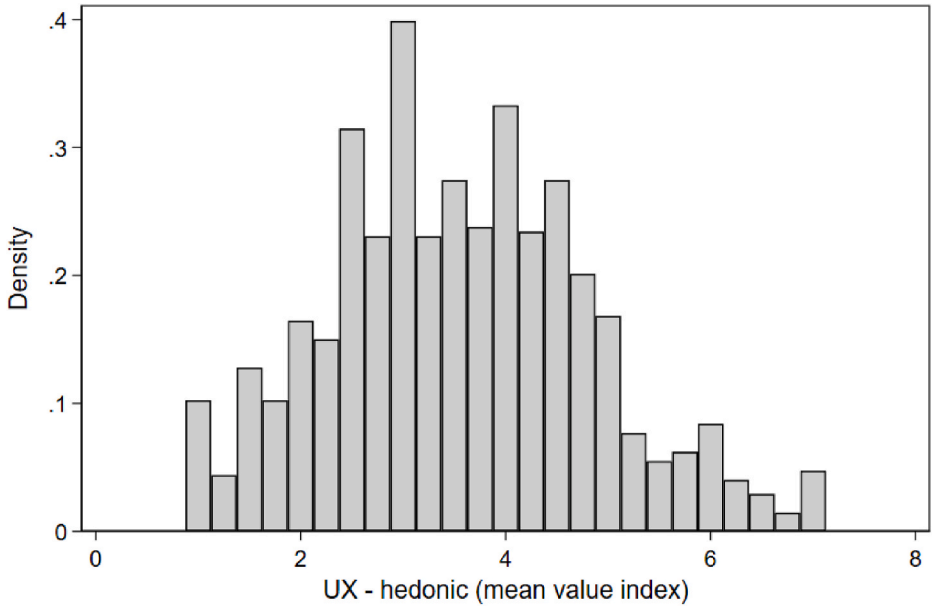


Fig. 13. Distribution of rankings of hedonic quality of UX in online banking context. N = 1093.

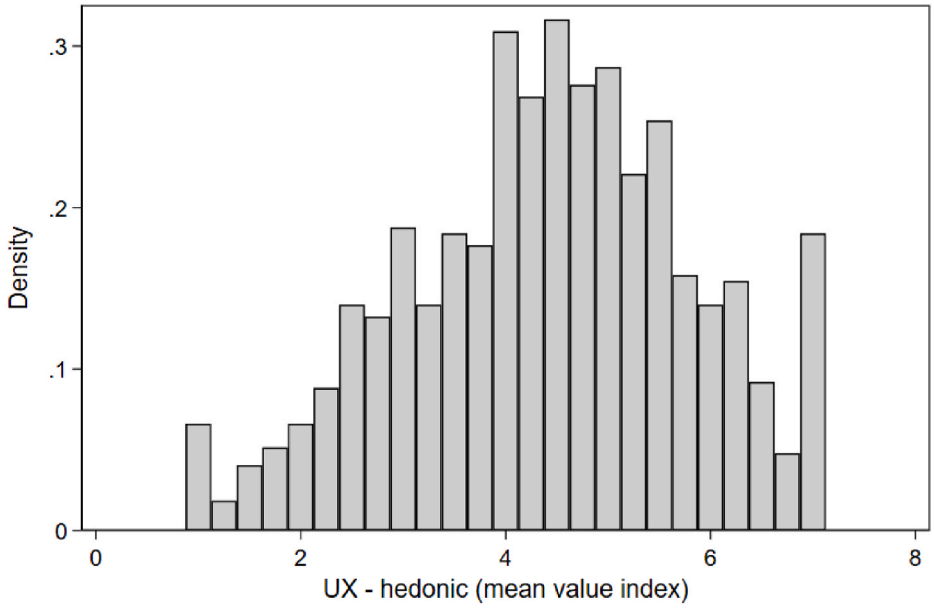


Fig. 14. Distribution of rankings of hedonic quality of UX in e-voting context. N = 1087.

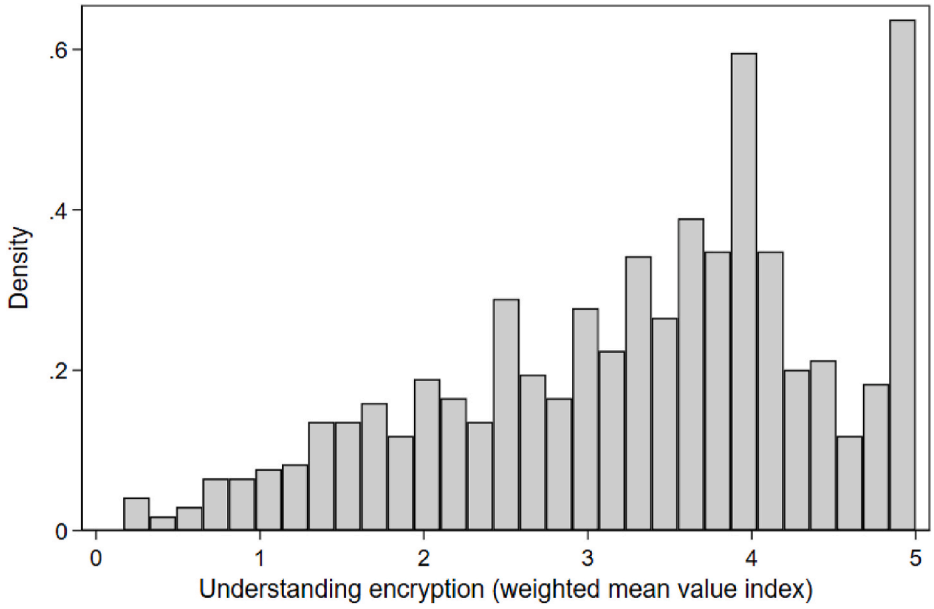


Fig. 15. Distribution of rankings of understanding of encryption in online banking context. N = 1093.

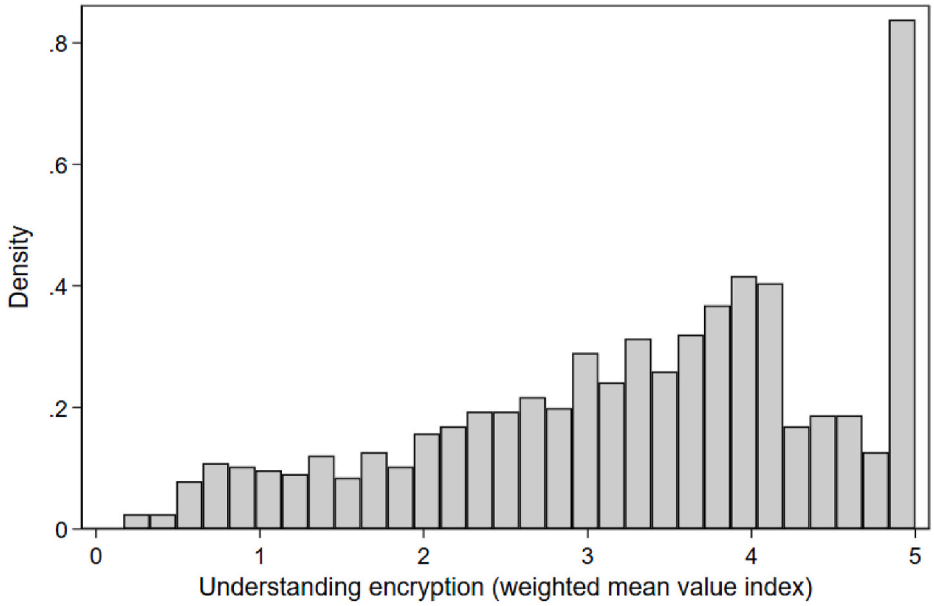


Fig. 16. Distribution of rankings of understanding of encryption in e-voting context. N = 1087.

Regression tables

Table 6
Perceived security – overall effects of textual and visual representation

	Voting		Banking	
Text representation	0.415*	(0.173)	0.590***	(0.165)
Visual representation	−0.143	(0.209)	0.029	(0.162)
Constant	6.882***	(0.230)	7.102***	(0.198)
Observations	1087		1093	

N = 1087 in e-voting, N = 1093 in online banking. Dependent variable: perceived security (scale 1–10). Robust standard errors in parentheses.
$^{+}$ $p < .10$, *$p < .05$, **$p < .01$, ***$p < .001$.

Table 7
Perceived security - single effects of the values of textual and visual representation

	Voting		Banking	
Text: Encrypting your data	0.402^{+}	(0.210)	0.257	(0.192)
Lower complexity description	0.397^{+}	(0.212)	0.741***	(0.189)
Higher complexity description	0.460*	(0.220)	0.696***	(0.192)
Padlock in front of ciphertext	−0.096	(0.260)	−0.052	(0.218)
Vote/Banknote dissolving into ciphertext	−0.191	(0.276)	−0.370^{+}	(0.224)
Vote/Banknote arrow with padlock moving to polling station/bank	−0.024	(0.257)	0.092	(0.208)
Vote/Banknote in envelope	−0.029	(0.277)	0.235	(0.214)
Computer connected to polling station/bank	−0.376	(0.269)	0.142	(0.209)
Constant	6.877***	(0.231)	7.134***	(0.197)
Observations	1087		1093	

N = 1087 in e-voting, N = 1093 in online banking. Dependent variable: perceived security (scale 1–10). Robust standard errors in parentheses.
$^{+}$ $p < .10$, *$p < .05$, **$p < .01$, ***$p < .001$.

Table 8
Regression table for perceived security with interactions

	Model with interactions	
Text: Encrypting your data	0.408^{+}	(0.210)
Lower complexity description	0.401^{+}	(0.212)
Higher complexity description	0.458*	(0.219)
Online banking	0.376^{+}	(0.210)
Text: Encrypting your data # Online banking	−0.139	(0.285)
Lower complexity description # Online banking	0.351	(0.284)
Higher complexity description # Online banking	0.246	(0.291)
Padlock in front of ciphertext	−0.069	(0.169)

(*continued on next page*)

Table 8 (*continued*)

	Model with interactions	
Vote/Banknote dissolving into ciphertext	−0.269	(0.177)
Vote/Banknote arrow with padlock moving to polling station/bank	0.038	(0.165)
Vote/Banknote in envelope	0.110	(0.173)
Computer connected to polling station/bank	−0.107	(0.168)
Constant	6.809***	(0.185)
Observations	2180	

N = 2180. Dependent variable: perceived security (scale 1–10). Robust standard errors in parentheses.
$^{+}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Table 9
Pragmatic quality of UX (UX-PQ) – overall effects of textual and visual representation

	Voting		Banking	
Text representation	0.060	(0.058)	0.036	(0.062)
Visual representation	−0.066	(0.065)	0.006	(0.072)
Constant	6.309***	(0.072)	6.099***	(0.080)
Observations	1087		1093	

N = 1087 in e-voting, N = 1093 in online banking. Dependent variable: UX-pragmatic quality, mean value index based on four items. Robust standard errors in parentheses.
$^{+}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Table 10
Pragmatic quality of UX (UX-PQ) - single effects of the values of textual and visual representation

	Voting		Banking	
Text: Encrypting your data	0.060	(0.070)	0.114	(0.071)
Lower complexity description	0.077	(0.068)	0.045	(0.074)
Higher complexity description	0.028	(0.073)	−0.072	(0.079)
Padlock in front of ciphertext	−0.149^{+}	(0.085)	−0.246*	(0.102)
Vote/Banknote dissolving into ciphertext	−0.020	(0.079)	−0.043	(0.093)
Vote/Banknote arrow with padlock moving to polling station/bank	0.004	(0.081)	0.149^{+}	(0.085)
Vote/Banknote in envelope	−0.073	(0.089)	0.143	(0.092)
Computer connected to polling station/bank	−0.104	(0.089)	0.042	(0.090)
Constant	6.314***	(0.072)	6.102***	(0.079)
Observations	1087		1093	

N = 1087 in e-voting, N = 1093 in online banking. Dependent variable: UX-pragmatic quality, mean value index based on four items. Robust standard errors in parentheses.
$^{+}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Table 11
Hedonic quality of UX (UX-HQ) – overall effects of textual and visual representation

	Voting		Banking	
Text representation	0.040	(0.098)	0.124	(0.090)
Visual representation	0.187	(0.114)	0.030	(0.102)
Constant	4.223***	(0.130)	3.435***	(0.118)
Observations	1087		1093	

N = 1087 in e-voting, N = 1093 in online banking. Dependent variable: UX-hedonic quality, mean value index based on four items. Robust standard errors in parentheses.
$^{+}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Table 12
Hedonic quality of UX (UX-HQ) - single effects of the values of textual and visual representation

	Voting		Banking	
Text: Encrypting your data	−0.003	(0.121)	0.099	(0.111)
Lower complexity description	−0.035	(0.119)	0.141	(0.111)
Higher complexity description	0.134	(0.120)	0.143	(0.112)
Padlock in front of ciphertext	−0.012	(0.147)	−0.078	(0.136)
Vote/Banknote dissolving into ciphertext	0.170	(0.148)	0.147	(0.134)
Vote/Banknote arrow with padlock moving to polling station/bank	0.330*	(0.145)	0.042	(0.132)
Vote/Banknote in envelope	0.244	(0.150)	0.050	(0.132)
Computer connected to polling station/bank	0.217	(0.148)	−0.006	(0.135)
Constant	4.228***	(0.130)	3.433***	(0.118)
Observations	1087		1093	

N = 1087 in e-voting, N = 1093 in online banking. Dependent variable: UX-hedonic quality, mean value index based on four items. Robust standard errors in parentheses.
$^{+}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Table 13
Understanding of encryption – overall effects of textual and visual representation

	Voting		Banking	
Text representation	0.471***	(0.092)	0.600***	(0.088)
Visual representation	−0.057	(0.101)	0.012	(0.090)
Constant	3.015***	(0.120)	2.839***	(0.110)
Observations	1029		1052	

N = 1087 in e-voting, N = 1093 in online banking. Dependent variable: Understanding of encryption, weighted mean value index based on six items. Robust standard errors in parentheses.
$^{+}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Table 14
Understanding of encryption - single effects of the values of textual and visual representation

	Voting		Banking	
Text: Encrypting your data	0.291**	(0.112)	0.386***	(0.108)
Lower complexity description	0.437***	(0.111)	0.659***	(0.104)
Higher complexity description	0.697***	(0.109)	0.780***	(0.101)
Padlock in front of ciphertext	−0.029	(0.130)	0.139	(0.117)
Vote/Banknote dissolving into ciphertext	−0.143	(0.131)	0.044	(0.127)
Vote/Banknote arrow with padlock moving to polling station/bank	0.057	(0.128)	0.035	(0.119)
Vote/Banknote in envelope	−0.031	(0.137)	−0.105	(0.119)
Computer connected to polling station/bank	−0.130	(0.133)	−0.103	(0.119)
Constant	3.008***	(0.119)	2.841***	(0.110)
Observations	1029		1052	

N = 1087 in e-voting, N = 1093 in online banking. Dependent variable: Understanding of encryption, weighted mean value index based on six items. Robust standard errors in parentheses.
$^{+}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Table 15
Regression table for understanding of encryption with interactions

	Model with interactions	
Text: Encrypting your data	0.296**	(0.112)
Lower complexity description	0.435***	(0.111)
Higher complexity description	0.699***	(0.109)
Online banking	−0.115	(0.111)
Text: Encrypting your data # Online banking	0.085	(0.155)
Lower complexity description # Online banking	0.219	(0.152)
Higher complexity description # Online banking	0.074	(0.148)
Padlock in front of ciphertext	0.057	(0.087)
Vote/Banknote dissolving into ciphertext	−0.051	(0.091)
Vote/Banknote arrow with padlock moving to polling station/bank	0.049	(0.087)
Vote/Banknote in envelope	−0.070	(0.090)
Computer connected to polling station/bank	−0.116	(0.089)
Constant	2.982***	(0.100)
Observations	2081	

N = 2180. Dependent variable: Understanding of encryption, weighted mean value index based on six items. Robust standard errors in parentheses.
$^{+}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

References

Abu-Salma, R., Krol, K., Parkin, S., Koh, V., Kwan, K., Mahboob, J., Traboulsi, Z., & Sasse, M. A. (2017). The security blanket of the chat world: An analytic evaluation and a user study of telegram. In *Proceedings 2nd European workshop on useable security*. Paris, France: European Workshop on Usable Security. https://doi.org/10.14722/eurousec.2017.23006.

Abu-Salma, R., Redmiles, E. M., Ur, B., & Wei, M. (2018). Exploring user mental models of end-to-end encrypted communication tools. In *Proceedings of the 8th USENIX workshop on free and open communications on the internet* (p. 8). FOCI.

Acemyan, C. Z., Kortum, P., Byrne, M. D., & Wallach, D. S. (2014). Usability of voter verifiable, end-to-end voting systems: Baseline data for Helios, Prêt à Voter, and Scantegrity II. *The USENIX Journal of Election Technology and Systems, 2*(3), 26–56.

Adams, A., & Sasse, M. A. (1999). Users are not the enemy. *Communications of the ACM, 42*(12), 40–46. https://doi.org/10.1145/322796.322806

Al-Natour, S., Cavusoglu, H., Benbasat, I., & Aleem, U. (2020). An empirical investigation of the antecedents and consequences of privacy uncertainty in the context of mobile apps. *Information Systems Research, 31*(4), 1037–1063. https://doi.org/10.1287/isre.2020.0931

Alvarez, R. M., Hall, T. E., & Trechsel, A. H. (2009). Internet voting in comparative perspective: The case of Estonia. *PS: Political Science & Politics, 42*(3), 497–505. https://doi.org/10.1017/S1049096509090787

Auspurg, K., & Hinz, T. (2015). *Factorial Survey Experiments*. SAGE Publications.

Blattner, M. M., Sumikawa, D. A., & Greenberg, R. M. (1989). Earcons and icons: Their structure and common design principles. In *Human-computer interaction* (Vol. 4, p. 11). Computers & Applied Sciences Complete, 1.

Busse, K., Schäfer, J., & Smith, M. (2019). *Replication: No one can hack my mind revisiting a study on expert and non-expert security practices and advice. Fifteenth symposium on useable privacy and security*. SOUPS 2019.

Chan, A. H. S., & Ng, A. W. Y. (2010). Investigation of guessability of industrial safety signs: Effects of prospective-user factors and cognitive sign features. *International*

Journal of Industrial Ergonomics, 40(6), 689–697. https://doi.org/10.1016/j.ergon.2010.05.002

Cox, A. L., Gould, S. J. J., Cecchinato, M. E., Iacovides, I., & Renfree, I. (2016). Design frictions for mindful interactions: The case for microboundaries. In *Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems - CHI EA '16* (pp. 1389–1397). https://doi.org/10.1145/2851581.2892410

Cranor, L. F. (2021). Informing California privacy regulations with evidence from research. *Communications of the ACM, 64*(3), 29–32. https://doi.org/10.1145/3447253

Damghanian, H., Zarei, A., & Siahsarani Kojuri, M. A. (2016). Impact of perceived security on trust, perceived risk, and acceptance of online banking in Iran. *Journal of Internet Commerce, 15*(3), 214–238. https://doi.org/10.1080/15332861.2016.1191052

Distler, V., Lallemand, C., & Koenig, V. (2020a). Making encryption feel secure: Investigating how descriptions of encryption impact perceived security. In *The 5th European workshop on useable security (EuroUSEC)* (p. 10).

Distler, V., Lenzini, G., Lallemand, C., & Koenig, V. (2020b). The framework of security-enhancing friction: How UX can help users behave more securely. In *New security paradigms workshop 2020* (pp. 45–58). https://doi.org/10.1145/3442167.3442173

Distler, V., Zollinger, M.-L., Lallemand, C., Roenne, P. B., Ryan, P. Y. A., & Koenig, V. (2019). Security—visible, yet unseen?. In *Proceedings of the 2019 CHI conference on human factors in computing systems* (pp. 1–13).

Dosso, C., & Chevalier, A. (2021). How do older adults process icons during a navigation task? Effects of aging, semantic distance, and text label. *Educational Gerontology, 47*(3), 132–147. https://doi.org/10.1080/03601277.2021.1886634

Dourish, P., Grinter, R. E., de la Flor, J. D., & Joseph, M. (2004). Security in the wild: User strategies for managing security as an everyday, practical problem. *Personal and Ubiquitous Computing, 8*.

Dourish, P., & Redmiles, D. (2002). An approach to useable security based on event monitoring and visualization. In *Proceedings of the 2002 workshop on new security paradigms* (pp. 75–81). https://doi.org/10.1145/844102.844116

Edwards, W. K., Poole, E. S., & Stoll, J. (2008). Security automation considered harmful?. In *Proceedings of the 2007 workshop on new security paradigms - NSPW '07* (p. 33). https://doi.org/10.1145/1600176.1600182

Egelman, S., Kannavara, R., & Chow, R. (2015). Is this thing on?: Crowdsourcing privacy indicators for ubiquitous sensing platforms. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems* (pp. 1669–1678). https://doi.org/10.1145/2702123.2702251

Fahl, S., Harbach, M., Muders, T., Smith, M., & Sander, U. (2012). Helping Johnny 2.0 to encrypt his Facebook conversations. *Proceedings of the Eighth Symposium on Usable Privacy and Security - SOUPS, '12*, 1. https://doi.org/10.1145/2335356.2335371

Fassl, M., Gröber, L., & Krombholz, K. (2021). Exploring user-centered security design for useable authentication ceremonies. In *Proceedings of the 2021 CHI conference on human factors in computing systems* (p. 15).

Felt, A. P., Ainslie, A., Reeder, R. W., Consolvo, S., Thyagaraja, S., Bettes, A., Harris, H., & Grimes, J. (2015). Improving SSL warnings: Comprehension and adherence. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems - CHI '15* (pp. 2893–2902). https://doi.org/10.1145/2702123.2702442

Felt, A. P., Reeder, R. W., Ainslie, A., Harris, H., Walker, M., Thompson, C., Acer, M. E., Morant, E., & Consolvo, S. (2016). Rethinking connection security indicators. In *Twelfth symposium on useable privacy and security (SOUPS 2016)* (pp. 1–14). https://www.usenix.org/conference/soups2016/technical-sessions/presentation/porter-felt.

Gallagher, K., Patil, S., & Memon, N. (2017). New me: Understanding expert and non-expert perceptions and usage of the tor anonymity network. In *Thirteenth symposium on useable privacy and security (SOUPS 2017)* (pp. 385–398). https://www.usenix.org/conference/soups2017/technical-sessions/presentation/gallagher.

Gaw, S., Felten, E. W., & Fernandez-Kelly, P. (2006). Secrecy, flagging, and paranoia: Adoption criteria in encrypted email. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 591–600). Association for Computing Machinery. https://doi-org.proxy.bnl.lu/10.1145/1124772.1124862.

Hassenzahl, M. (2001). The effect of perceived hedonic quality on product appealingness. *International Journal of Human-Computer Interaction, 13*(4), 481–499. https://doi.org/10.1207/S15327590IJHC1304_07

Hassenzahl, M. (2008). User experience (UX): Towards an experiential perspective on product quality. In *Proceedings of the 20th conference on l'Interaction homme-machine* (pp. 11–15).

Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In G. Szwillus, & J. Ziegler (Eds.), *Vol. 57. Mensch & computer 2003* (pp. 187–196). Vieweg+Teubner Verlag. https://doi.org/10.1007/978-3-322-80058-9_19.

Hassenzahl, M., Eckoldt, K., Diefenbach, S., Laschke, M., Len, E., & Kim, J. (2013). Designing moments of meaning and pleasure. Experience design and happiness. *International Journal of Design, 7*(3).

Huang, H., Yang, M., Yang, C., & Lv, T. (2019). User performance effects with graphical icons and training for elderly novice users: A case study on automatic teller machines. *Applied Ergonomics, 78*, 62–69. https://doi.org/10.1016/j.apergo.2019.02.006

International Organization for Standardization. (2018). Ergonomics of human-system interaction—Part 11: Usability: Definitions and concepts (standard No. 9241-11: 2018). ed-2:v1:en https://www.iso.org/obp/ui/#iso:std:iso:9241:-11.

Ion, I., Reeder, R., & Consolvo, S. (2015). No one can hack my mind": Comparing expert and non-expert security practices. In *Eleventh symposium on useable privacy and security (SOUPS 2015)* (pp. 327–346). https://www.usenix.org/conference/soups2015/proceedings/presentation/ion.

Khan, I. U., Hameed, Z., & Khan, S. U. (2017). Understanding online banking adoption in a developing country: UTAUT2 with cultural moderators. *Journal of Global Information Management, 25*(1), 43–65. https://doi.org/10.4018/JGIM.2017010103

Krug, S. (2000). *Don't make me think!: A common sense approach to web usability*. Pearson Education India.

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. In A. Holzinger (Ed.), *Vol. 5298. HCI and usability for education and work* (pp. 63–76). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-89350-9_6.

Lausch, J., Wiese, O., & Roth, V. (2017). What is a secure email?. In *Proceedings 2nd European workshop on useable security. European workshop on useable security*. https://doi.org/10.14722/eurousec.2017.23022. Paris, France.

Lim, S. H., Kim, D. J., Hur, Y., & Park, K. (2019). An empirical study of the impacts of perceived security and knowledge on continuous intention to use mobile fintech payment services. *International Journal of Human-Computer Interaction, 35*(10), 886–898. https://doi.org/10.1080/10447318.2018.1507132

Mahlke, S. (2008). *User experience of interaction with technical systems*. Doctoral dissertation.

Majrashi, K. (2020). Performance of mobile users with text-only and text-and-icon menus in seated and walking situations. *Behaviour & Information Technology*, 1–19. https://doi.org/10.1080/0144929X.2020.1795257

Marky, K., Kulyk, O., Renaud, K., & Volkamer, M. (2018). What did I really vote for?. In *Proceedings of the 2018 CHI conference on human factors in computing systems - CHI '18* (pp. 1–13). https://doi.org/10.1145/3173574.3173750

McDougall, S., Reppa, I., Kulik, J., & Taylor, A. (2016). What makes icons appealing? The role of processing fluency in predicting icon appeal in different task contexts. *Applied Ergonomics, 55*, 156–172. https://doi.org/10.1016/j.apergo.2016.02.006

Murillo, A., Kramm, A., Schnorf, S., & Luca, A. D. (2018). "If I press delete, it's gone"—user understanding of online data deletion and expiration. In *Fourteenth symposium on useable privacy and security (SOUPS 2018)* (pp. 329–339). https://www.usenix.org/conference/soups2018/presentation/murillo.

Özlen, M. K., & Djedovic, I. (2017). Online banking acceptance: The influence of perceived system security on perceived system quality. *Journal of Accounting and Management Information Systems, 16*(1), 164–178. https://doi.org/10.24818/jamis.2017.01008

Pagter, J. I., & Petersen, M. G. (2007). A sense of security in pervasive computing—is the light on when the refrigerator door is closed? *International Conference on Financial Cryptography and Data Security*, 383–388.

Panditrao, S., O'Brien, D., & Stark, E. (2021, July 14). Increasing HTTPS adoption. *Chromium Blog*. https://blog.chromium.org/2021/07/increasing-https-adoption.html.

Petitpas, A., Jaquet, J. M., & Sciarini, P. (2020). Does E-Voting matter for turnout, and to whom? *Electoral Studies*, 102245. https://doi.org/10.1016/j.electstud.2020.102245

Rogers, Y. (1989). Icons at the interface: Their usefulness. *Interacting with Computers, 1*(1), 105–117. https://doi.org/10.1016/0953-5438(89)90010-6

Ruoti, S., Kim, N., Burgon, B., van der Horst, T., & Seamons, K. (2013). Confused Johnny: When automatic encryption leads to confusion and mistakes. In *Proceedings of the ninth symposium on useable privacy and security - SOUPS '13* (p. 1). https://doi.org/10.1145/2501604.2501609

Ryan, P. Y., Rønne, P. B., & Iovino, V. (2016). Selene: Voting with transparent verifiability and coercion-mitigation. In *International conference on financial cryptography and data security* (pp. 176–192).

Schechter, S. E., Dhamija, R., Ozment, A., & Fischer, I. (2007). The emperor's new security indicators. https://doi.org/10.1109/SP.2007.35, 51-65.

Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Design and evaluation of a short version of the user experience questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence, 4*, 103. https://doi.org/10.9781/ijimai.2017.09.001

Sheldon, K. M., Elliot, A. J., Kim, Y., & Kasser, T. (2001). What is satisfying about satisfying events? Testing 10 candidate psychological needs. *Journal of Personality and Social Psychology, 80*(2), 325.

Spero, E., & Biddle, R. (2020). Out of sight, out of mind: UI design and the inhibition of mental models of security. In *New security paradigms workshop 2020* (pp. 127–143). https://doi.org/10.1145/3442167.3442174

Vance, A., Lowry, P. B., & Eggett, D. (2015). Increasing accountability through user-interface design artifacts: A new approach to addressing the problem of access-policy violations. *MIS Quarterly, 39*(2), 345–366. https://doi.org/10.25300/MISQ/2015/39.2.04

Vassil, K., Solvak, M., Vinkel, P., Trechsel, A. H., & Alvarez, R. M. (2016). The diffusion of internet voting. Usage patterns of internet voting in Estonia between 2005 and 2015. *Government Information Quarterly, 33*(3), 453–459. https://doi.org/10.1016/j.giq.2016.06.007

Vaziripour, E., Wu, J., O'Neill, M., Whitehead, J., Heidbrink, S., Seamons, K., & Zappala, D. (2017). Is that you, Alice?. In *A usability study of the authentication ceremony of secure messaging applications. Thirteenth symposium on useable privacy and security (SOUPS 2017)* (pp. 29–47). https://www.usenix.org/conference/soups2017/technical-sessions/presentation/vaziripour.

Wallander, L. (2009). 25 years of factorial surveys in sociology: A review. *Social Science Research, 38*(3), 505–520. https://doi.org/10.1016/j.ssresearch.2009.03.004

Whitten, A., & Tygar, J. D. (1999). A usability evaluation of PGP 5.0. In *Proceedings of the 8th USENIX security symposium* (pp. 169–183).

Wiedenbeck, S. (1999). The use of icons and labels in an end user application program: An empirical study of learning and retention. *Behaviour & Information Technology, 18* (2), 68–82. https://doi.org/10.1080/014492999119129

Wogalter, M. S., Silver, N. C., Leonard, S. D., & Zaikina, H. (2006). Warning symbols. In M. S. Wogalter (Ed.), *Handbook of warnings* (pp. 159–176). CRC Press. https://doi.org/10.1201/9781482289688.