
From the Quest to Replace Passwords towards Supporting Secure and Usable Password Creation

Dissertation prepared for the degree of Doctor rerum naturalium (Dr. rer. nat.)

Presented by Verena Zimmermann, M.Sc.

Referees:

First Referee: Prof. Dr. Joachim Vogt

Second Referee: Prof. Karen Renaud, PhD

Darmstadt 2020



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Department of Human Sciences
Institute of Psychology
Work and Engineering Psychology

From the Quest to Replace Passwords towards Supporting Secure and Usable Password Creation

Dissertation prepared by Verena Zimmermann, born in Dieburg, Germany

Darmstadt, Technische Universität Darmstadt

Year of the publication of the dissertation via TUprints: 2021

URN: urn:nbn:de:tuda-tuprints-174254

Date of the submission: 29th September 2020

Date of the oral exam: 11th November 2020

Published under the CC BY-NC-ND 4.0 International

<https://creativecommons.org/licenses>

Darmstadt - D17

*Passwords will be with us forever.
(Ronald Rivest, 2005)*

Affirmation

Erklärung gemäß § 9 Abs. 1 S. 6 PO/AT der Allgemeinen Bestimmungen der Promotionsordnung der Technischen Universität Darmstadt

Hiermit erkläre ich, Verena Zimmermann, dass ich die Arbeit - abgesehen von den in ihr ausdrücklich genannten Hilfen - selbstständig verfasst habe. Alle Stellen, die aus anderen Quellen entnommen wurden, sind als solche kenntlich gemacht worden. Die Arbeit ist von mir mit einem Verzeichnis aller benutzten Quellen versehen.

Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Bei der abgegebenen Arbeit stimmen die schriftliche und die zur Archivierung eingereichte elektronische Fassung überein.

Statement pursuant to § 9 subparagraph 1 page 6 PO/AT of the General Provisions of the Doctoral Regulations of Technische Universität Darmstadt

I herewith formally declare that I, Verena Zimmermann, have written the submitted dissertation independently. I did not use any outside support except for the quoted literature and other sources mentioned in the dissertation. I clearly marked and separately listed all of the literature and all of the other sources which I employed when producing this academic work, either literally or in content. The dissertation is provided by me with a list of all used sources.

This dissertation has not been handed in or published before in the same or similar form.

In the submitted dissertation the written copies and the electronic version for archiving are identical in content.

Datum/ Date

Unterschrift/ Signature

Acknowledgement

I would like to express my gratitude for all people that have supported me during the last years as a doctoral researcher.

My special thanks go to my doctorate supervisors Professor Dr. Joachim Vogt and Professor Karen Renaud, PhD. I like to thank them for their support, mentoring, and their helpful, and constructive feedback. I furthermore thank them for the opportunity of a great research stay with Karen Renaud at the Abertay University in Scotland. Finally, I thank both for the engagement in reviewing this dissertation and hope for a continuation of the valuable research collaboration in the future.

Furthermore, I would like to express my appreciation for all my (former) colleagues of the research group Work and Engineering Psychology at Technische Universität Darmstadt and my project partners within the project “User-friendly confidential and authentic communication” within the Center for Research in Security and Privacy (CRISP) that has now been established as ATHENE, the National Research Center for Applied Cybersecurity. I would also like to thank the student researchers Alexandra von Preuschen and Marius Kleboth who did a great job supporting our project work. In addition, I like to thank my cooperation partners from different research groups, especially Nina Gerber, Karola Marky, and Peter Mayer, for the great collaboration.

Finally, I would like to thank my family and friends for their unconditional support. Above all, I would like to mention my husband Tim, my parents Andreas and Christiane, and my sister Katharina.

Parts of this work were accomplished within ATHENE and CRISP, respectively, and were thus financially supported by the German Federal Ministry of Education and Research (BMBF) and by the Hessian Ministry of Science and the Arts.

Abstract

Authentication is an important measure for protecting personal and sensitive information from unauthorised access. Password authentication still is the most widely used form of authentication despite its well-established downsides, including the cognitive load it poses for users and coping strategies resulting thereof. These include the creation of weak passwords or the reuse of passwords across accounts. Alternatives to the knowledge-based password scheme include biometric schemes, such as fingerprint authentication and token-based schemes like chip card authentication. However, attempts to replace the password on a large scale have not yet been successful.

Commencing this research with an extensive rating and comparison of objective features of existing authentication schemes confirmed that the password indeed is not easily replaceable. To shine light on this seemingly intractable issue, a laboratory and an online study were conducted to explore the user perceptions of authentication schemes. Although studied less frequently than technical aspects, user perceptions are highly relevant. First, they can influence acceptance of authentication schemes, and second, mismatches between technical security and security perceptions can ultimately impact security. The two studies revealed a user preference for password authentication across different contexts of use, despite its downsides. While the initial comparison acknowledged the password's persistence with regard to objective features, the studies confirm the relevance of password authentication from a user perspective. Because the security of password authentication largely depends on the password creation and handling of the user, further research was needed to explore measures that support secure and usable password authentication.

A promising approach for encouraging secure choices without constraining the user is provided by the concept of "nudging", as proposed by Thaler and Sunstein. Nudges are small tweaks of the choice architecture that target automatic cognitive processes and that do not limit or significantly influence the cost of the available choices. To support secure password creation, three consecutive field studies analysed the impact of various password nudges on password creation. The first two studies used visual nudges intended to simply encourage stronger passwords and produced insignificant results. Based on the lessons learned, the resulting intervention in the third study combined a nudge with password strength information and compensation for stronger passwords in the form of later password expiry. This intervention indeed encouraged the creation of stronger passwords.

The finding led to the assumption that the combination of a nudge and information provision, a *hybrid nudge*, may be more effective in encouraging secure choices than either intervention on its own. An online study analysed the single and joint effects of nudges and information provision across different security-related decisions including password creation. The findings revealed that the hybrid nudge proved to be most effective across decisions. Furthermore, the combination of transparent nudges with information provision educating users about the reasons for encouraging a particular choice appeared most favourable with regard to ethical considerations. A final online study compared the effects of different hybrid password nudges on password creation, password memorability, and the users' perceptions. It confirmed the effectiveness of the hybrid nudge as compared to exclusive information or nudge interventions on all three counts. Yet, nearly no significant differences between hybrid password nudges emerged, indicating that the type of nudge included plays a minor role compared to the combination as such.

It is concluded that the combination of nudging and information provision constitutes a promising strategy for supporting users in creating secure passwords and in making security-related decisions without enforcing a particular choice. This may further open the path towards a more human-centred approach in cybersecurity as envisioned in a mindset labelled "Cybersecurity, Differently".

The findings are discussed regarding the transferability of the results to real-life settings and their scalability to the large number of accounts users have to manage. Suggestions for future work include field studies on hybrid password nudges, the integration into suitable tools such as password managers to ease the cognitive load, or the development of concepts that especially consider aspects such as account sensitivity or password reuse.

Authentifizierung ist eine wichtige Maßnahme zum Schutz persönlicher und sensibler Informationen vor unbefugtem Zugriff. Nach wie vor ist das Passwort die am weitesten verbreitete Form der Authentifizierung, trotz bekannter Nachteile wie der kognitiven Belastung für die Nutzenden und daraus resultierenden Bewältigungsstrategien. Diese umfassen die Erstellung schwacher Passwörter oder die Wiederverwendung von Passwörtern über Konten hinweg. Alternativen zum wissensbasierten Passwortverfahren sind biometrische Verfahren wie Fingerabdruck-Authentifizierung und gegenstandsbezogene Verfahren wie Chipkarten-Authentifizierung. Bisherige Versuche, das Passwort im großen Maßstab zu ersetzen, waren jedoch nicht erfolgreich.

Beginnend mit einem umfangreichen Bewertungsprozess und Vergleich objektiver Aspekte von existierenden Authentifizierungsverfahren, konnte diese Forschung die Schwierigkeit, das Passwort durch ein anderes Verfahren abzulösen, bestätigen. Um dieses scheinbar unlösbare Problem zu beleuchten, wurden eine Laborstudie und eine Onlinestudie zu den Nutzenwahrnehmungen von Authentifizierungsverfahren durchgeführt. Diese sind relevant, obwohl sie bisher weniger untersucht wurden als technische Aspekte. Erstens können sie die Akzeptanz von Authentifizierungsverfahren beeinflussen und zweitens können Diskrepanzen zwischen technischer Sicherheit und Sicherheitswahrnehmung letztendlich die Sicherheit beeinträchtigen. Die zwei Studien zeigten eine Präferenz der Nutzenden für das Passwortverfahren in verschiedenen Anwendungskontexten trotz seiner Nachteile. Während der initiale Vergleich die Beständigkeit des Passworts im Hinblick auf objektive Aspekte verdeutlichte, bestätigten die beiden Studien die Relevanz des Passwortverfahrens aus Perspektive der Nutzenden. Da die Sicherheit des Passwortverfahrens maßgeblich von der Passwörterstellung und -handhabung durch die Nutzenden abhängt, ist weitere Forschung zur Unterstützung von Nutzenden bei der Erstellung sicherer Passwörter notwendig.

Ein vielversprechender Ansatz zur Förderung sicherer Entscheidungen, ohne die Nutzenden zu beschränken, stellt das Konzept des "Nudging" von Thaler und Sunstein dar. Nudges sind kleine Veränderungen der Entscheidungsarchitektur, die automatische, kognitive Prozesse aktivieren und die weder die Entscheidungsoptionen eingrenzen noch deren Kosten signifikant beeinflussen. Zur Unterstützung sicherer Passwörterstellung wurde in drei aufeinander aufbauenden Feldstudien der Einfluss verschiedener Passwort-Nudges auf die Passwörterstellung untersucht. Die ersten zwei Studien mit visuellen Nudges, die lediglich zu erhöhter Passwortsicherheit ermuntern sollten, erzielten keine signifikanten Ergebnisse. Die aus den Erkenntnissen resultierende Intervention in der dritten Studie kombinierte einen Nudge mit Passwortstärke-Information und der Kompensation stärkerer Passwörter durch eine längere Passwortgültigkeit. Diese Intervention förderte tatsächlich die Erstellung stärkerer Passwörter.

Die Erkenntnisse führten zu der Annahme, dass die Kombination aus einem Nudge und Informationsvermittlung, ein *hybrider Nudge*, wirksamer zur Förderung sicherer Entscheidungen sein könnte als einzelne Maßnahmen für sich alleine. Daher untersuchte eine Online-Studie den individuellen und kombinierten Effekt von Nudges und Informationsvermittlung auf verschiedene sicherheitsbezogene Entscheidungen einschließlich Passwörterstellung. Die Ergebnisse zeigten, dass der hybride Nudge über die Entscheidungen hinweg am effektivsten war. Weiterhin erscheint die Kombination eines transparenten Nudges mit Informationen, die Nutzende über den Grund für die Förderung einer bestimmten Entscheidung aufklären, aus ethischen Gesichtspunkten am günstigsten. In einer abschließenden Online-Studie wurden die Auswirkungen verschiedener hybrider Nudges auf Passwörterstellung, Merkbarkeit und die Nutzenwahrnehmungen verglichen. Die Studie bestätigte die Wirksamkeit hybrider Nudges gegenüber Interventionen, die nur einen Nudge oder nur Informationen beinhalteten, in allen drei Punkten. Allerdings wurden nahezu keine signifikanten Unterschiede zwischen verschiedenen hybriden Nudges gefunden, was darauf hindeutet, dass die Art des verwendeten Nudges eine untergeordnete Rolle im Vergleich zur Kombination als solche einnimmt.

Die Ergebnisse legen nahe, dass die Kombination aus Nudging und Informationsvermittlung eine erfolgsversprechende Strategie zur Unterstützung von Nutzenden im Hinblick auf sichere Passwörterstellung und sicherheitsrelevante Entscheidungen darstellt, ohne eine bestimmte Option zu erzwingen. Dies könnte auch den Weg zu einem stärker menschen-zentrierten Cybersecurity Ansatz öffnen, wie er mit der als "Cybersecurity, Differently" bezeichneten Denkweise vorgestellt wird.

Die Ergebnisse werden unter anderem im Hinblick auf ihre Übertragbarkeit auf reale Situationen und ihre Skalierbarkeit für die große Anzahl von Accounts, die Nutzende verwalten müssen, diskutiert. Die Vorschläge für zukünftige Forschung beinhalten daher Feldstudien zu hybriden Passwort-Nudges, ihre Integration in geeignete Tools wie Passwort-Manager zur Reduzierung der kognitiven Belastung von Nutzenden oder die Entwicklung von Konzepten, die insbesondere Aspekte wie Account-Sensitivität und Wiederverwendung von Passwörtern berücksichtigen.

Contents

Outline and Contribution of the Dissertation	1
Part A: Synopsis	4
Part A.1: Objective Features and User Perceptions of Authentication Schemes	4
1 Introduction to Human Factors in Authentication	4
1.1 Password Authentication	4
1.2 Alternative Forms of Authentication	6
1.2.1 Knowledge-based Authentication	6
1.2.2 Biometric Authentication	9
1.2.3 Token-based Authentication	13
2 Evaluating Objective Features and Subjective User Perceptions of Authentication Schemes	16
2.1 Rating of Authentication Schemes	16
2.2 User Perceptions of Different Authentication Schemes	18
2.2.1 Pilot Study	18
2.2.2 Summary of the Main Study	20
3 The Influence of Context and Type of Scheme on User Perceptions of Authentication Schemes	23
3.1 Pilot Study	23
3.1.1 Pilot Study Method	23
3.1.2 Pilot Study Results and Implications for the Design of the Main Study	23
3.2 Summary of the Main Study	24
3.2.1 Method	24
3.2.2 Results	26
3.2.3 Discussion & Implications	26
4 Interim Conclusion	28
Part A.2: Supporting Secure and Usable Password Creation	30
5 Strategies for Enhancing Secure Password Creation	30
5.1 Constraining Strategies	30
5.1.1 System-generated Passwords	30
5.1.2 Regular Password Expiration	30
5.1.3 Password Policies	31
5.2 Supporting Strategies	31
5.2.1 Approaches for Creating Secure Passwords	31
5.2.2 Approaches for Increasing Memorability	32
5.3 Evaluation of Strategies for Enhancing Password Security	33
6 Password Nudges	35
6.1 Introduction to the Concept of Nudging	35
6.2 Field Studies on the Effectiveness of Password Nudges	36
6.2.1 Method	36
6.2.2 Results	38
6.2.3 Discussion & Implications	38

7	Differentiating the Concept of Nudging from Related Concepts	41
7.1	Dual Process Theories and the Concept of Nudging	41
7.2	Definition of the Nudge Concept	42
7.3	Definition of Related Interventions	42
7.3.1	Code	43
7.3.2	Sludge	43
7.3.3	Information Provision	43
7.3.4	Hybrid Nudge	44
7.4	Implications	44
8	The Ethics of Nudging	45
8.1	Arguments for Nudging	45
8.2	Arguments against Nudging	46
8.3	Guidelines for Ethical Nudging	46
8.4	Implications	48
9	The Influence of Nudge Interventions on Security Decisions and Password Creation	49
9.1	Comparison of Different Nudge Interventions	49
9.1.1	Background	49
9.1.2	Method	50
9.1.3	Results	53
9.1.4	Discussion & Implications	56
9.2	Designing Hybrid Password Nudges for Secure and Usable Password Creation	57
9.2.1	Method	58
9.2.2	Results	61
9.2.3	Discussion & Implications	62
10	Discussion and Reflection	63
10.1	Summary of Findings	63
10.2	Implications & Contribution	65
10.2.1	Rating of Authentication Schemes	65
10.2.2	Studies on User Perceptions of Authentication Schemes	65
10.2.3	Field Studies on Password Nudges	66
10.2.4	The Concept of Nudging	67
10.2.5	Online Studies on Hybrid Password Nudges	67
10.3	Reflection & Limitations	68
10.3.1	Methodological Considerations	68
10.3.2	Content-related Considerations	71
10.4	Outlook & Future Work	72
10.4.1	Considering the Password Creation Context	72
10.4.2	Cybersecurity, Differently	73
10.5	Conclusion	76
Part B: Manuscripts		78
Manuscript 1: “Keep on Rating – On the Systematic Rating and Comparison of Authentication Schemes”		78
Manuscript 2: “The Password is Dead, Long Live the Password – A Laboratory Study on User Perceptions of Authentication Schemes”		79
Manuscript 3: “That Depends – How Context Affects User Perceptions of Authentication Schemes”		80
Manuscript 4: “Nudging Folks towards Stronger Password Choices: Providing Certainty is the Key”		81

Manuscript 5: “Ethical Guidelines for Nudging in Information Security & Privacy”	82
Manuscript 6: “The Nudge Puzzle: Matching Nudge Interventions to Cybersecurity Decisions”	83
Manuscript 7: “Hybrid Password Meters for more Secure Passwords - A Comprehensive Study of Password Meters and Nudges”	84
Manuscript 8: “Moving from a “Human-as-Problem” to a “Human-as-Solution” Cybersecurity Mindset”	85
Bibliography	86
List of Abbreviations	107
List of Figures	108
List of Tables	109
Appendix	110
Appendix A: Author Vita	110
Appendix B: List of Publications	111
Appendix C: Reference List of Authentication Schemes	114

Outline and Contribution of the Dissertation

Overall, this dissertation is structured into two parts: Part A comprises the synopsis of the dissertation. The synopsis includes the theoretical background for motivating the research, connects the different studies and research works within this dissertation, and links the findings and implications in an overall discussion. Of the eight manuscripts included within this dissertation, six have already been published while two are currently being prepared for submission or are under review to make the findings publicly accessible and to allow for peer feedback. The bibliographical information of each publication and manuscript, respectively, is included in Part B of the dissertation. The synopsis briefly summarizes the main findings of each manuscript to allow for reading the synopsis without having to switch between the synopsis and the manuscripts multiple times. As the manuscripts provide much more detail for each study, the reader is invited to take a closer look at each manuscript.

Figure 1 visualizes the logical structure of the dissertation described in the synopsis as a road map and shows which of the manuscripts listed in Part B relates to each research step. This is indicated by a small document symbol with a number that corresponds to the number of the manuscript.

The aim of this dissertation was twofold: first, to explore the user perceptions of different forms of authentication as described in Part A.1, and second, to support user-friendly and secure authentication based on the findings as detailed in Part A.2.

As depicted in Figure 1, the first step towards research aim 1 was to review the literature on human factors in authentication. The insights were incorporated in the introduction of this dissertation. The background colours yellow and blue indicate that this research step considered the objective and technological features of different authentication schemes (yellow) as well as subjective user perceptions of authentication (blue).

Research step 2 involved the identification of numerous, different authentication schemes from the literature review conducted in step 1. These were rated in terms of objective features, namely objective security, deployability, and usability features, using a rating framework developed by Bonneau *et al.* [34]. The rating results were then implemented within an authentication choice support tool named ACCESS [197, 198, 247]. In so doing, this research contributes to enlarging the ACCESS database that supports researchers and practitioners in choosing suitable authentication schemes for their particular context and also provides suggestions for further improving ACCESS [349]. The different schemes in the database were then compared and weighed according to multiple selection criteria to identify the best-rated schemes from different authentication categories, such as knowledge-based or biometric schemes.

In step 3, a laboratory study [347] was conducted to evaluate potential differences in user perceptions of twelve authentication schemes selected in step 2. Furthermore, differences and similarities in terms of the users' perceptions and the objective and technological features of the schemes were compared. This laboratory study, in which people interacted with responsive mock-ups of each authentication scheme, addresses the lack of empirical studies analysing actual as compared to hypothetical user perceptions of authentication schemes (e.g., analysed in a survey). In addition, it contributes by analysing a large number of authentication schemes that were implemented in similar ways to control for external influences, such as different designs. The comparison of technological aspects and user perceptions supports understanding of the user perspective and facilitates the detection of potential mismatches. These are especially important as a mismatch between actual and perceived security may lead users to either refrain from using a technically secure scheme, or to engage in insecure practices by using a technically insecure scheme [140].

In step 4, a consecutive online study was conducted to explore the influence of the type of scheme and the context of use on user perceptions. To do so, the three schemes that were most preferred by the participants in research step 3 were compared across different types of accounts. The study served to validate the assumptions derived in the previous research step and also analysed the impact of potential influencing factors on user perceptions in more detail.

In an interim conclusion, the combination of the rating of objective authentication features with the study results in terms of subjective user perceptions revealed the persisting relevance of secure password creation. Part A.2 of the dissertation thus focuses on supporting secure and usable password creation to work towards research aim 2. The combination of the findings in terms of objective features and subjective user perceptions is depicted by the colour gradient from blue and yellow to green, as shown in Figure 1.

The next step, research step 5, was to review and compare different existing and potential strategies to support secure and usable password creation. Again, the mechanisms (yellow) as well as implications in terms of user perceptions and behaviour (blue) were considered. Based on the review of strategies, "nudging" [303], i.e., subtly changing the choice architecture to encourage secure choices without limiting the users' set of options, was selected as a promising approach.

In three exploratory field studies described in step 6, several password nudges were evaluated in terms of their influence on password security [253, 252]. By so doing, this research contributes to the relatively small number of studies analysing the effects of security-related nudges "in the wild".

Based on the lessons learned in the field studies, a "detour", i.e., an excursus, was made to more clearly define the concept of nudging and to separate it from other interventions such as information provision in research step 7 [249]. Furthermore, the analysis of the nudge concept also led to the exploration of the ethical aspects of nudging. In research step 8, ethical guidelines for deploying nudges in the area of information security and privacy were derived based on psychological guidelines for ethical research [249, 250]. One contribution of this work lies in the differentiation of nudges and related interventions that forms the basis for evaluating their effectiveness and ethical implications. An ethics checklist also serves as an additional aid for other researchers and practitioners aiming to deploy nudges in security, privacy, or related fields.

Afterwards, an online study was conducted dissecting the influences of nudges, information provision, and a combination of nudges and information provision on different security-related decisions. Research step 9 thereby constitutes an important step towards more empirical research analysing *how* nudges exert their influence by examining the single and joint effects of different nudge interventions across different types of decision contexts.

Research step 10 applied the insights to password creation in particular. In an online study, a number of password meters with nudges based on different biases and heuristics were evaluated in terms of their effect on password creation, password memorability, and the users' perceptions. Concrete implications for the design of nudge-related interventions to support secure and usable password creation can be derived from the results.

Finally, research step 10 is followed by a summary and discussion of all research steps presented up to that point. The discussion leads to an outlook on "the road ahead". It describes how future interventions in the wider security and privacy area can be designed in an even more user-centred way to enhance the users' ability to be part of the solution in complex socio-technical environments, instead of being viewed as a potential problem to control. The approach described in research step 11 is labelled "Cybersecurity, Differently" [353].

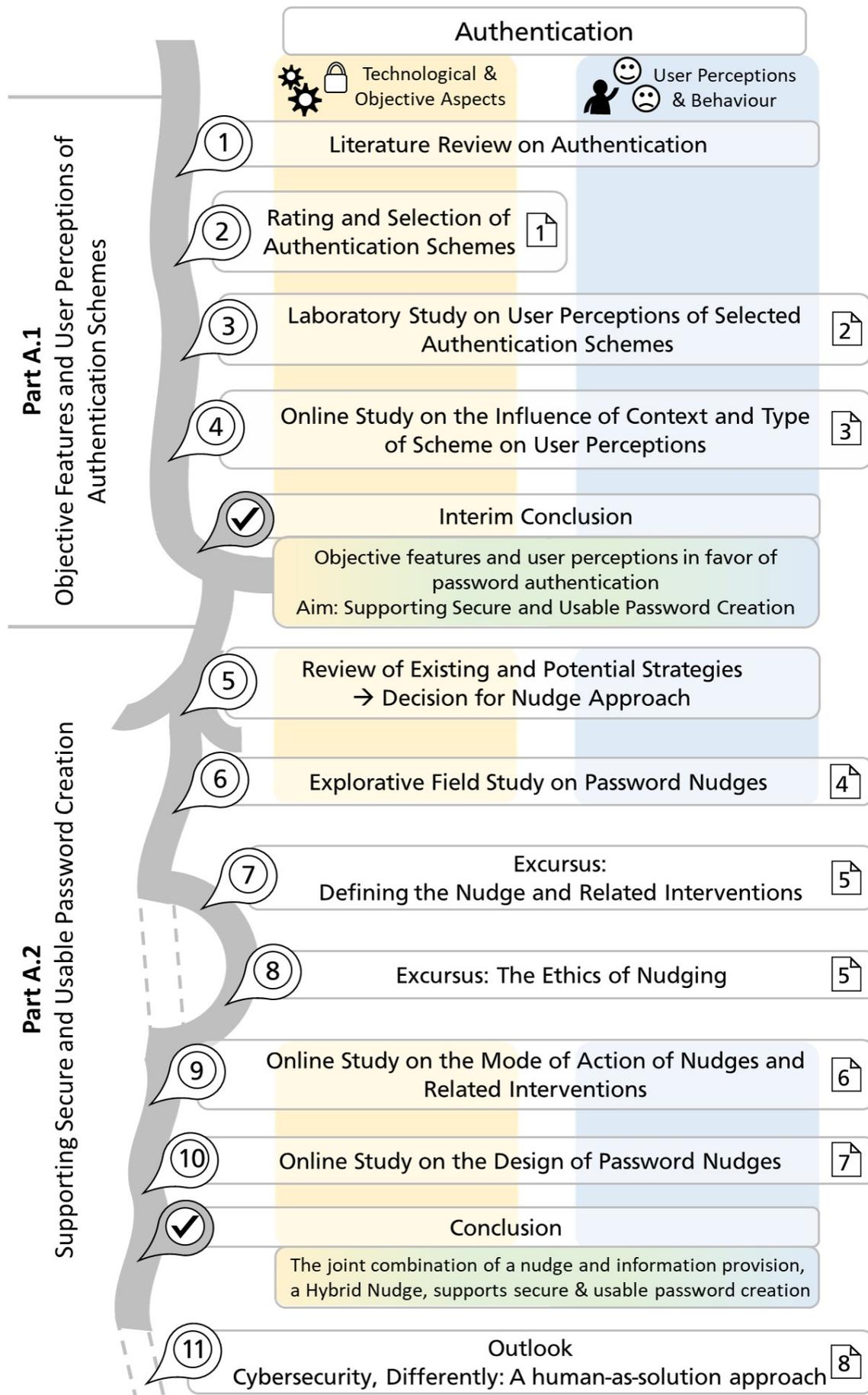


Figure 1: Outline of the dissertation.



Figure 2: Research step 1.

Authentication is an important means to protect data from unauthorized access, that is, to enable access by authorized people and to prevent access by others. When it comes to authentication, three terms need to be differentiated according to Renaud [243]. The first is identification. This step asks a person to identify his- or herself, e.g., by means of a passport, an email address, account number, or an object the person carries. The identity is then compared to the identities in the particular database. It can be compared as one-to-many (identification) or one-to-one (verification) [76]. The second step then is authentication in which the user provides some kind of secret that has been determined in the enrollment phase to provide evidence for the claimed identity. The secret is usually knowledge-based (something the user knows), biometric (something the user is), or token-based (something the user carries). If the provided secret is correct and belongs to the claimed identity, the user is authenticated. Within the service, database, or tool the user authenticated for, he or she may be granted certain permissions to undertake actions. This third step is called authorization.

An authentication system usually consists of the user, the technology, and an interface between the user and the technology, i.e., some mechanism for inputting the authentication secret according to the authentication mechanism.

With regard to authentication, the literature shows that the user is an integral, but also a critical component of the system [243]. The secrets of many authentication schemes, e.g., knowledge-based or cognitive authentication schemes, require cognitive activity [243]. Others require users to perform certain actions (e.g., behavioural biometric authentication schemes) or to provide certain physical or physiological features (biometric authentication schemes). Besides relying on human cognitive processes and behaviour for inputting the secret, the human also plays a vital part in securing the authentication secret. Knowledge-based authentication secrets can be told or discovered, for example, if they are easy to guess for others or physically stored in an insecure place. Some biometric features such as fingerprints could be captured from surfaces by a potential attacker, and tokens could be forgotten, lost, or stolen.

The following sections will provide more detailed information on human factors in authentication based on the reviewed literature. The next section will start with the most commonly used form of authentication, password authentication, before describing alternative forms such as biometric and token-based authentication schemes in more detail.

1.1 Password Authentication

As stated above, password authentication has been and still is the most common form of authentication [292]. A password consists of a sequence of alphanumeric symbols, e.g., uppercase and lowercase letters, numbers, and symbols, that can either be randomly generated by the system or selected by the user [243]. Passwords belong to the group of knowledge-based authentication schemes as the secret, i.e., the password, is supposed to be memorized by the user.

The history of passwords dates back to the early 1960s. At the Massachusetts Institute of Technology (MIT) Computation Center, Fernando Corbató worked on a project concerned with improving computer efficiency through time-sharing [61]. At that time, the user-computer interaction was very slow due to long periods of time required for writing and debugging programs. As a solution to that problem, Corbató suggested time-shared computer usage by making the computer simultaneously available to multiple users in a system called Compatible Time-Sharing System (CTSS) [61].

To protect each user's private files and information from accidental or deliberate manipulation, Fano and Corbotá [93] described a login interface that required the users to provide a project number and a personal password. This was one of the first reported uses of password authentication and originally intended for the highly technically-adept Computation Center personnel. Yet, as computers and the Internet became increasingly available to all kinds of industries and the general public, also password authentication spread across industries and user groups, and from only one to multiple accounts.

Reasons for the password's prevalence might include the following advantages: Compared to other authentication schemes, passwords are relatively easy to deploy, i.e., they are server and browser-compatible and produce low costs per user [34, 36]. On the user side, password authentication has been found easy to learn and efficient to use. Users do not need to buy or carry additional devices and forgotten or stolen passwords can easily be recovered [34], e.g., by sending a password reset link to an associated email account.

However, passwords are not without disadvantages, as also acknowledged by Corbató who has been one of the first persons to use passwords [60]. In terms of security, passwords are static secrets and prone to phishing attacks and keystroke logging, as well as guessing attacks [100, 133]. With increasing computing capacities, passwords need to have a higher entropy to withstand possible offline and online guessing attacks. Information entropy, as introduced by Shannon [278], is a measure for the password's complexity or unpredictability from which it can be calculated how long it would take a potential attacker to guess the password. It is measured in bits and influenced, e.g., by the character sets included in the password and password length. According to the National Institute of Standards and Technology (NIST) [115], passwords should be at least eight characters in length. However, service providers should permit passwords to be at least 64 characters long as length has found to be an important predictor for password entropy or password strength, respectively [163, 171]. In contrast to that, password policies that are frequently deployed by service-providers, i.e., requirements for a password to include certain character types such as uppercase letters, numbers, and symbols, have been found to be less effective in terms of password strength [329]. Furthermore, they also negatively impact memorability and usability [142].

Memorizing a long, complex password as suggested by NIST [115] or required by password policies might not be a problem if a user has one or very few password-protected accounts. Yet, studies found that the number of user accounts increased a lot throughout the years [116]. In 2007, Florencio and Herley [94] studied the password habits of a large group of users and found that users had 25 accounts on average. Some years later, in 2014, Stobert and Biddle [293] conducted an interview study on password habits, in which participants reported having between nine and up to 51 different online accounts such as online shopping accounts, banking accounts, or social network accounts.

Memorizing up to 51 unique and long passwords would pose an enormous cognitive load for the users. To deal with that cognitive effort, users deploy a number of coping strategies:

- *Choosing weak passwords.* Users tend to choose short and easily memorable passwords that, for example, include dictionary words, names, patterns, or birthdays [116, 314, 315]. They are thus easier to guess for a potential attacker. Wei, Golla, and Ur showed that password creation is also influenced by the type of service the password is created for [326]. For example, passwords created for the LinkedIn platform often contained variations of the name "LinkedIn".
- *Reusing password across accounts.* Research found that users frequently reuse passwords across accounts [94, 230, 281, 323], even more so when partial reuse is taken into account [230].
- *Writing passwords down.* Users tend to write passwords down [4, 281, 293, 355] so they might be easily discovered by other present users. However, the practice itself might not pose a security problem in a world where attackers do not have to be present to attack an account. Some researchers even advocate for it [50, 133, 270] given that passwords are stored safely as compared to posted on the screen or keyboard.

Furthermore, research showed that the users' perception of password security differs from technical password security measures. For example, in a laboratory study, Ur *et al.* [315] identified several misconceptions in terms of password strength, such as that adding a symbol at the end of the password makes it secure, or that difficult-to-spell passwords are more secure than others. Expecting only very targeted attacks, participants also believed, e.g., birthdays, to be secure as long as they were not posted online. In an online study analysing perceived and actual password strength, users had misconceptions in terms of the impact of including digits or keyboard patterns in their passwords [314].

With the help of an online game, Seitz and Hussmann [275] compared user ratings of password security to their technical security. They found that users underestimate passphrases by 1.4 points difference on a five-point score.

In some cases, users also seem to lack an understanding of the authentication mechanism that influences their security perception. For example, Bhagavatula *et al.* [27] found that users perceived the security of fingerprint authentication on a mobile phone to be higher than that of a personal identification number (PIN), i.e., a password only consisting of numbers. However, the PIN served as a fall-back mechanism so that the biometric scheme cannot be more secure than the PIN in that case.

Overall, this section showed that password security not only depends on the technical measures used to securely store and transfer passwords but also largely on the users' perceptions, their password creation, and their password handling. To mitigate the security problems arising from the conflict between the high cognitive load for memorizing numerous passwords and the users' strategies to cope with that effort, a number of alternatives to password authentication have been developed. These are presented in the next sections.

1.2 Alternative Forms of Authentication

To overcome the shortcomings of the text password, many alternative forms of authentication schemes have been developed. They can be classified in terms of the type of secret on which they are based. As described above, a common differentiation is between knowledge-based, biometric, and token-based authentication schemes. The following sections will describe the three categories and subcategories of authentication schemes in more detail, along with some examples.

The focus of the following sections will be on single-factor authentication schemes, as compared to two-factor authentication, as two-factor authentication is essentially a combination of any two single authentication schemes. Similarly, special forms of authentication such as single sign-on services and password managers that create and store passwords for the user fall outside the scope of this differentiation as they are based on a single authentication scheme as well, e.g., password managers and many single sign-on services use text passwords as an authentication scheme.

1.2.1 Knowledge-based Authentication

Knowledge-based authentication is based on a secret only known to the person that is authorized to access the data or service, i.e., "something you know". The secret can take the form of text, of which the most popular example is the text password. Besides, it can take the form of, e.g., pictures, procedures, gestures, or formulas used to calculate a secret. The secret can further be recall-based or recognition-based [160, 300]. Figure 3 provides an overview of these variations that are explained in the following sections in more detail.

Like the password, other knowledge-based authentication schemes rely on human cognition and memory. The advantage is that users thus do not need to carry additional tokens. Another advantage is that many knowledge-based secrets are easily revocable [34], which means that users can easily recover from a loss of the secret, e.g., by having a reset link sent to a connected email account. Yet, on the downside, like with password authentication, the cognitive effort increases with the number of accounts and unique secrets. Furthermore, if a secret is actively shared with other users or guessed by an attacker, it is not possible for the system to know who the actual user is. That's because the secret is usually not linked to a person but some kind of identity such as an email address or account name [145].

As the security of knowledge-based authentication very much depends on the scheme itself, e.g., whether it uses a static or dynamic secret, and the handling of the secret by the user, details on technical security are provided along with the descriptions of certain examples for alphanumeric, graphical, and cognitive authentication.

Alphanumeric Authentication

As stated above, the most prominent example of alphanumeric authentication is the password or variants thereof, such as a mnemonic password or a passphrase. As the password has already been described in great detail above, this section describes another form of alphanumeric authentication, namely personal knowledge questions.

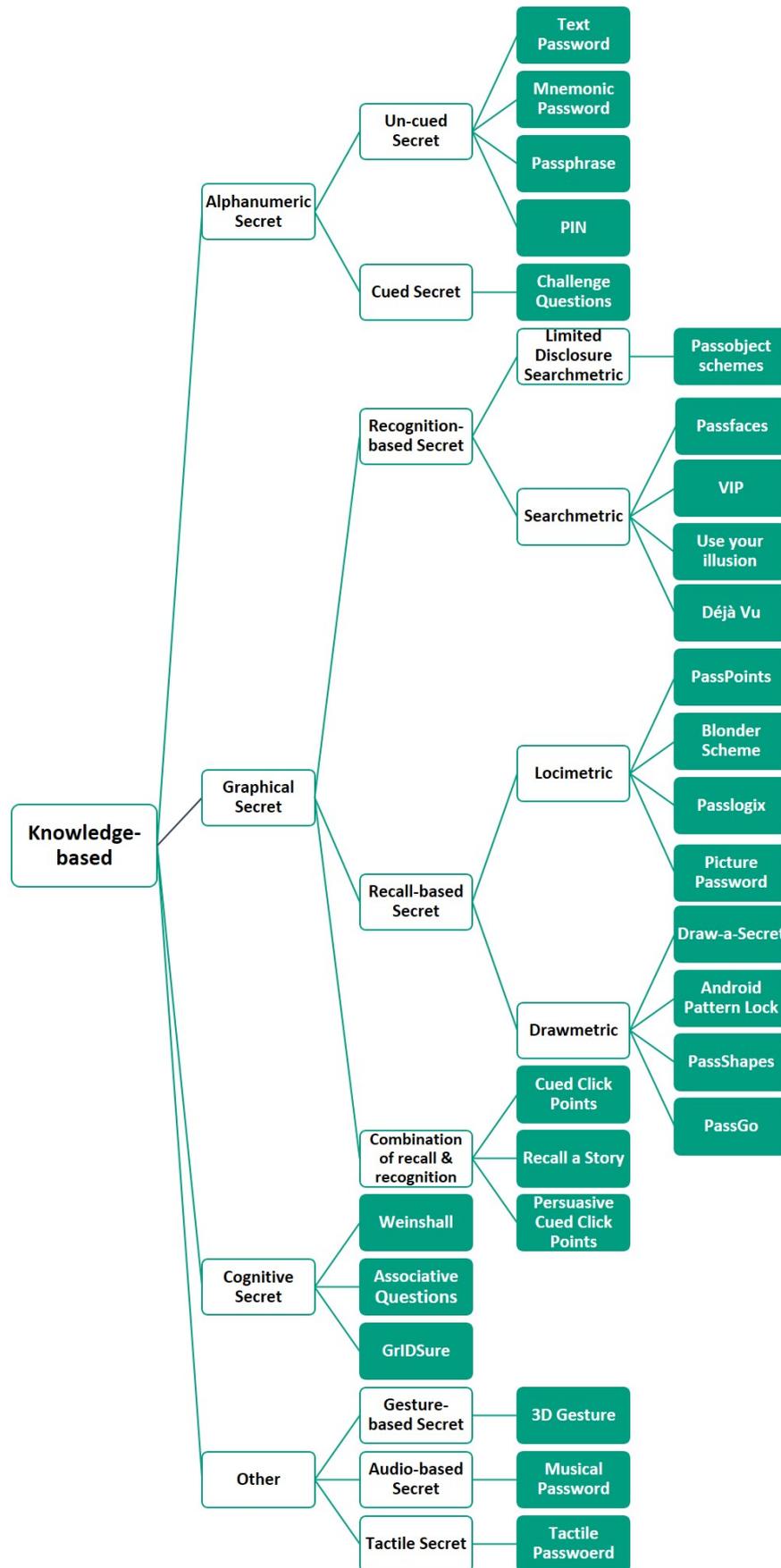


Figure 3: Categorization of knowledge-based authentication schemes. *Note:* The list is not exhaustive. For references to the examples listed in the figure, see Appendix 10.5.

Example: Personal Knowledge Questions

Personal knowledge questions, also known as challenge questions [238] or cultural passwords [243], ask users to provide answers to personal questions, such as the mother’s maiden name or the favourite school subject. The idea behind the scheme is that users should be able to easily recall this kind of personal information with the help of a cue (cued recall). Thus, the scheme has often been used as a reset mechanism for account recovery, e.g., in case of a forgotten password [153]. However, previous studies showed that the answers to many of these questions are either known by relatives or friends of the user, or are easily searchable or guessable. Personal knowledge questions thus provide reduced security as compared to passwords [33, 238]. Furthermore, studies found that many users provide fake answers in order to increase security, but involuntarily reduce security by hardening answers in a predictable way. Even the assumed benefit of high memorability could not be confirmed in the study that instead found that about 40% of users were unable to recall their previously provided answers [33].

Graphical Authentication

Graphical authentication makes use of the fact that people are better in recalling pictures than words [263]. As described by Renaud [244] and illustrated in Figure 3, graphical authentication schemes can be clustered into different categories depending on the action the user has to complete to authenticate, e.g., search for a pre-selected picture among distractors (searchmetric) or mark a certain pre-selected point in a picture (locimetric).

Example: Persuasive Cued Click Points [51, 52]

One example of a graphical scheme is Persuasive Cued Click Points (PCCP) [51, 52]. Upon registration, the user chooses one click point in each of the five subsequent images from a larger set of images. To overcome the finding that users often choose salient and thus relatively easy-to-guess click points in an image [78, 306], PCCP encourages users to choose a "random" click point by restricting choice to a randomly selected area within the image. To log in, the user has to click the pre-selected points on each picture again within a small error margin. Each click point determines the next image that comes up. If the user makes a mistake, the next image is one that was not among the images the user has chosen a click point on. The image thus provides feedback to the user, who can quickly go back to correct the mistake, whereas attackers would not profit from this feedback if they did not know the selected images. The scheme provides many security features such as resilience to targeted impersonation, leaks from other verifiers, phishing, or theft [34]. However, it is not resilient to physical observation, and like other knowledge-based schemes, is not effortless in terms of memorability [34]. In user studies, the scheme was well perceived in terms of usability and perceived security [52] by the participants.

Cognitive Authentication

Cognitive authentication differs from the other two categories in that the user does not need to memorize the secret per se, but some kind of formula or procedure. Applying this formula leads to the calculation of a new secret each time the user authenticates.

Example: Weinshall [328]

The Weinshall scheme [328] consists of a picture matrix that contains a large number of pictures (between 20 and 80 depending on the configuration). The picture matrix consists of a random selection of a picture set the user has to memorize, and some decoy images. As compared to a picture password, it is not sufficient to recognize the memorized images among the decoys, but the user has to apply a set of rules. The user starts with the picture in the upper left corner of the matrix. If the picture is not among the memorized ones, the user goes to the next picture to the right. Otherwise, the user goes to the picture below. This process is repeated until the user reaches the lower or right side of the matrix and notes the number that is displayed next to the final picture in the matrix. The number is part of the code the has to enter to finally authenticate. To derive the complete code, the user has to repeat the process with as many matrices as numbers in the code before the final code can be entered. If users make a mistake in the process, they have to start anew. The password space, and the scheme’s resistance to brute-force attacks, is calculated with between 2^{47} and 2^{190} password options [328] depending on the number of images, combinations, and the length of the code. However, Golle and Wagner [113] could show that the scheme was not resistant to eavesdropping attacks. In terms of usability, the Weinshall scheme has been found to require user training [328]. In line with that, it was neither rated as easy-to-use, memorywise-effortless, nor efficient by Bonneau *et al.* [34].

The "secrets" in biometric authentication schemes are based on unique features inherent to a person, i.e. "something you are". The word secret is used in brackets here, as biometric characteristics are not per se secret even though they might sometimes be hard to acquire without the person cooperating [218, 324]. That is why some researchers even dismiss biometrics as a form of authentication [269].

Biometric technologies have been defined as "*automated methods of verifying or recognizing the identity of a living person based on a physiological or behavioral characteristic.*" [324, p.1]. According to this definition, biometric authentication has often been divided into schemes based on physiological and behavioural characteristics [10, 28, 145, 324, 325]. Physiological characteristics can further be differentiated from anatomical physical characteristics, which is captured in the definition by Riley *et al.* [255].

While physiological characteristics result from the interaction of physical and (bio-)chemical processes in the organism and can change during measurement [266], anatomical, physical features remain unchanged during measurement. An example of a physical feature would be the fingerprint; an example of a physiological feature would be a person's brainwaves that change in reaction to a cognitive task. A behavioural feature would be the way someone walks. Figure 4 visualizes the different forms of biometric authentication schemes.

Biometric authentication can further be classified in terms of whether it is based on static information (e.g., a facial image) or dynamic information (e.g., blinking movement) [284], whether the information is examined once per session or continuously, or whether the information is provided actively or passively.

An advantage of biometric authentication is that, in contrast to knowledge-based and token-based schemes, biometrics allow for recognition of a specific person [145] and require personal presence [76]. Furthermore, the "secret" cannot be forgotten like a password or lost like a token as people carry the characteristics with them in any case [145]. In that, they increase convenience for users. Mimicking or stealing biometric information is not impossible, e.g., a fingerprint can be lifted from a surface, but very difficult [145] especially as current sensors often include liveness checks. Thus, they cannot be circumvented by just using images of, e.g., a face or fingerprint. In addition, the storage of the sensitive biometric information can be protected by not storing the biometric information in its original format, but as an encrypted digital representation [145]. These so-called templates could, e.g., consist of certain points extracted from a fingerprint instead of a complete fingerprint image. The security of biometric systems could be further enhanced by using multimodal biometric systems [145].

To be suitable for authentication, a biometric feature should fulfill the following requirements according to Jain, Ross, and Prabhakar [145]:

- *Uniqueness*: The manifestation of the characteristic should be sufficiently different between individuals to be able to distinguish them.
- *Universality*: All people must possess the biometric characteristic.
- *Permanence*: The biometric characteristic should not be affected by time or age, respectively.
- *Measurability*: The acquisition of the biometric feature should be easy and allow for further processing, e.g., by using feature extraction.
- *Performance/ Reliability*: The technology used to extract the feature should be accurate, fast, and robust.
- *Circumvention*: It should be hard to substitute or imitate the characteristic.
- *Acceptability*: The use of the characteristic for authentication and the corresponding technology used to capture the data should be acceptable for a large group of the relevant population.

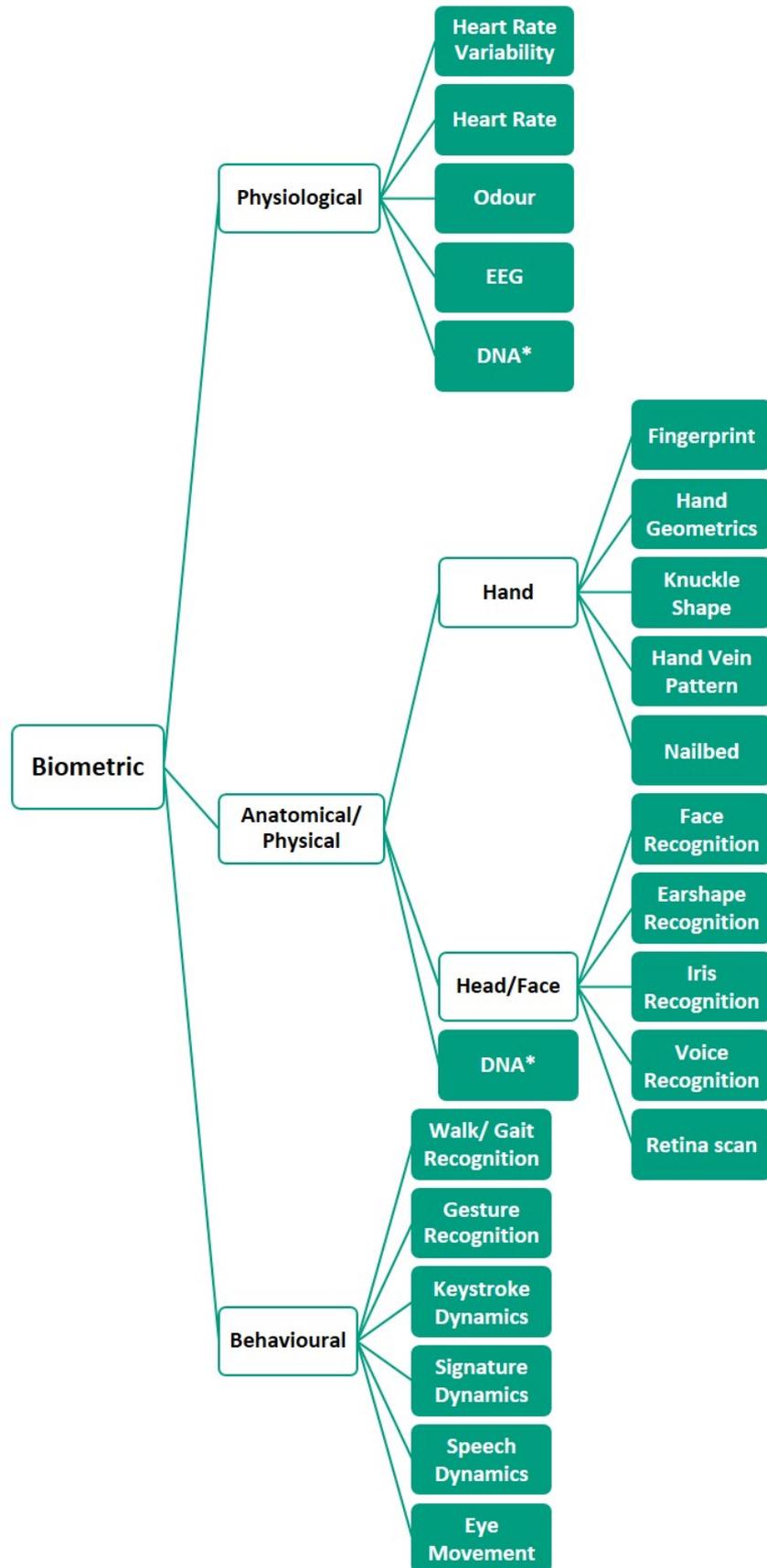


Figure 4: Categorization of biometric authentication schemes. *Note:* The list is not exhaustive. For references to the examples listed in the figure see Appendix 10.5. *DNA can be clustered differently.

While these requirements are important, they also pose some challenges. The following list describes the challenges connected to each requirement that apply to biometric authentication in general.

- *Uniqueness*: As stated above, biometric features should be unique for each person and thus might allow a link between the biometric feature and the person's identity. This, on the one hand, can be viewed as a security feature, but also raises privacy concerns on the other hand [63, 308]. Privacy concerns about how the personal and unique biometric information is stored and used are most prevalent. Even though not all biometric information can be used to reveal the user's identity, the fear is that biometric information could be linked to a person [324]. Face recognition could, for example, be used in combination with security cameras to detect and trace criminals, or to exert control over employees [261]. Other biometrics, such as iris images, can be used to derive personal information such as certain diseases [260]. Furthermore, if biometric information is successfully imitated or stolen, the information could be misused for identity theft. Apart from that, once stolen information cannot be easily revoked or replaced [324]. Another challenge results from the fact that even though the biometric characteristics might be distinct, the information collected by a sensor, e.g., certain points within a fingerprint, might bear inter-class similarities reducing the practical distinctiveness of the biometric feature [145].
- *Universality*: Even though biometric characteristics such as the fingerprint might be universal, not every person might be able to use them, e.g., due to a disease or injury. For example, fingerprint authentication might be temporarily or permanently unsuitable for people with burns or cuts, fingerprints worn down by chemicals or labour, or in case of genetically indistinct fingerprints [145, 261]. Thus, there is a small percentage of people of which the biometric characteristic cannot be captured [261].
- *Permanence*: Similar to the aspect stated above, many biometric characteristics might well be affected by age, disease, or injury over time, causing intra-class variations. Consider, e.g., the change in male voices in puberty, changes in an older person's gait, or difficulties in tracing an iris affected by cataract.
- *Measurability*: Not all biometrics are as easily observable and measurable as the fingerprint. For example, the extraction of someone's DNA or the capturing of someone's brainwaves during a certain task might be more intrusive and effortful. Additionally, the features that are more easily observable might also be more prone to "theft". For example, in the past, fingerprint images have been used to circumvent a sensor [45]. Another problem affecting the measurement process is noise in the data, e.g., caused by cuts in a finger or a dirty sensor [145].
- *Performance/ Reliability*: In general, the biometric technology performs the more accurate, the more data points it uses. However, this also increases the cost and processing time of technology [261]. Apart from that, for each level of accuracy, a trade-off has to be made. If the number of people who are falsely accepted is to be minimized, this increases the number of people falsely rejected. For example, a study cited by Schneier [267] calculated that a one percent increase of the false rejection rate would increase the throughput time of passengers at an airport by 45 minutes.
- *Circumvention*: As a general rule, Sasse [261] states that the better biometric technologies are protected against potential attacks, the more expensive they are. This might not only refer to the financial cost, but also to time or effort for the user. And even though sensors are getting more sophisticated, e.g., by using liveness detection, previous research has shown that it is possible - while cumbersome - to circumvent, e.g., a fingerprint sensor [14, 196], or to replace or steal the information stored in the database [205].
- *Acceptability*: The acceptability of using certain biometric characteristics for authentication might depend on a variety of factors, including culture and religion. For example, some religions prohibit the photographing of eyes or face or the touching of objects that have been touched by the opposite gender [261]. Research indicates that general acceptability might be higher for applications with a high perceived security need or when user effort is considerably reduced, e.g., when replacing multiple passwords with fingerprint authentication [261]. Finally, the acceptability among experienced users is higher than amongst inexperienced users [261].

To make these considerations more graspable, examples for physiological, physical and behavioural biometric authentication schemes are described in more detail along with their advantages and challenges in the following.

Physiological Biometric Features

Example: Body Odour Authentication

Body odour authentication is based on the fact that a component of each human's body odour is unique. The body odour can be captured non-intrusively by sensors that analyse the chemicals the smell consists of, known as volatiles [28]. As the composition of chemicals changes with recent activities and medications, it is, for example, possible to diagnose certain diseases or recent activities such as sexual activity apart from authenticating people. This raises some privacy issues [28]. Further, the analysis of the body odour might be impacted by chemicals in the surrounding environment, such as deodorants [76, 145].

Anatomical/Physical Biometric Features

Example: Fingerprint Authentication

Fingerprint authentication, which concerns the patterns of ridges and valleys on a fingertip [145], is the oldest form of biometric authentication [28]. While in the past ink has been used to capture and compare fingerprints, nowadays, sensors based on optical, thermal, silicon, or ultrasonic principles are used of which optical sensors are the most common [28].

With the progressive development of fingerprint sensors, their acquisition cost has been considerably reduced in the last years [234], making it publicly available in laptops and smartphones. Furthermore, the maturity and accuracy of fingerprint sensors has been shown to be very high [144, 240], which further contributed to the spread of fingerprint authentication. To date, it is the most commonly used biometric authentication scheme [292]. Still, as stated above, fingerprint authentication is not available for a small percentage of the population due to injuries, worn-down, or genetically indistinctive fingerprints [145, 261]. The performance of the authentication process can further be impacted by dirty fingers or sensors, and too wet or too dry fingers that influence their capacitance [28, 76]. Apart from that, the use of the personal fingerprint for authentication has been connected to privacy concerns (e.g., the fingerprint being misused for identity theft) [49, 68], safety concerns (e.g., being cut-off the finger for circumventing the system) [261], concerns in terms of hygiene (e.g., the transmission of viruses on shared sensors) [210], and associations with criminals as the fingerprint has often been used for forensic purposes [144]. Studies analysing user perceptions of fingerprint authentication have often found fingerprint authentication to be highly rated in terms of preference, convenience, or security perception by many participants [49, 81, 139, 193], but also disliked by some for privacy reasons [49, 68].

Behavioural Biometric Features

As indicated by the name, behavioural authentication relates to the behaviour of a person [28]. Anatomical/physical and physiological characteristics "are innate or naturally grown to; and behavioral biometrics are mannerisms or traits that are learned or acquired." [159, p.1566]. Examples include gait, gesture, signature dynamics, and keystroke dynamics recognition the last of which will be described in more detail below.

Example: Keystroke Dynamics

The underlying assumption of keystroke dynamics recognition is that each person types in a particular way. Even though the typing might not be exclusively unique to a person, it seems to be sufficiently distinct to differentiate users from one another [145]. The keystroke dynamics could be unintrusively traced for typing a specific password (static keystroke analysis) or for general typing over a longer time period (dynamic keystroke analysis) [159]. The method measures the time taken to type a certain word, the speed of typing, typing errors, the time between hitting keys, or the pressure applied [76, 159]. This method is relatively cheap as no external sensor or hardware is needed [28, 159]. Methods for classifying keystroke dynamics include statistical methods, neural networks, pattern recognition techniques, and hybrid approaches [159]. As described above, keystroke dynamics can be captured unintrusively and transparently. The advantage is that users might not need to be interrupted for authentication and are already familiar with typing and typing-based authentication, such as passwords, which might increase the acceptability of keystroke dynamics recognition [159]. However, the fact that keystroke dynamics can also be collected without the user being aware of it might also raise privacy concerns.

In token-based authentication schemes, the secret is a token or “something you have”, respectively. First of all, token-based schemes can be classified in terms of whether the token is linked to a person’s identity, e.g., an identity card, or whether it can be used by any person, e.g., a key or a transponder. In that case, the mere possession of the token enables a person to authenticate. Second, token-based schemes can be electronic, e.g., a hardware token, or non-electronic, e.g., a paper token. Sometimes, the differentiation is not entirely clear as also some of the non-electronic tokens require some electronic processing. Consider, for example, a paper card with a personal bar code that requires a bar code scanner or a list of codes that needs to be entered in an online system. Figure 5 provides an overview of different forms of token-based authentication schemes.

An advantage of token-based schemes is that they do not require any mental effort of the user, as no secret has to be remembered. However, a disadvantage is that tokens have to be carried and can thus easily be left or stolen. Furthermore, in the case of non-personalized tokens, the token can be used by any other person to authenticate. This reduces the security of single-factor token-based schemes [243, 261]. Apart from that, revoking or replacing lost or stolen tokens is often connected to high costs [261], as may be the general acquisition of hardware needed for the authentication process, e.g., in the case of smart card readers.

These disadvantages can at least be partially mitigated by storing multiple credentials on one token [261], miniaturizing tokens [261], or choosing tokens that users always carry with them, such as their smartphones. Another option is to combine token-based schemes with a secondary factor (two-factor authentication), e.g., a knowledge-based password that protects access to the information in case the token is stolen [261] or a secret that is directly incorporated in a configurable 3D object [194]. Then again, this involves some mental effort for the user similar to that of knowledge-based authentication. An example is the YubiKey [343] token, of which some configurations require a different password for each verifier. Previous research on user aspects of token-based authentication suggest that tokens such as smart cards or radio-frequency identification (RFID) tags were less accepted by users as compared to knowledge-based or biometric schemes [101, 152].

In the following, one example each will be provided for personalized and a non-personalized token.

Personalized Tokens

Example: Identity Cards

A prime example of personalized tokens are passports or identity cards, often including a photograph of the person, which are issued to citizens by states, but are also used by insurance companies or private organisations such as fitness studios issuing personalized membership cards that users have to carry. In the increasingly digital world, efforts are undertaken to combine the multiple identities a user has for different services, e.g., for accessing public services, travelling, health insurance, the driver licenses, or parking permits into one identity card or an electronic smart card, respectively. Examples of nations including several identities in one card, as listed in [26], are Greece, Malaysia, and Portugal. The challenges of these efforts include the large cost of the infrastructure for introducing, maintaining, and revoking the identity cards [26]. In addition, there might be privacy concerns resulting from storing many digital identities in one place and challenges in terms of trust in the issuing state or organisation [26].

Non-personalized Tokens

Example: Transponder

An example of a non-personalized electronic token would be a transponder that communicates with an electronic door or car lock [136]. The word transponder is a combination of *transmitter* and *responder* because a transponder, upon reception of an electronic signal, emits a response. In essence, a transponder is an electronic equivalent to a key with the advantage that for many different doors, a person would not need to carry a number of keys but that all required access codes can be stored on one transponder. It can take many different forms, such as a key chain, a chip card, e.g., as described in [136], or could be integrated into a personal watch or other objects such as suggested by [181]. The transponder being non-personalized means that if the transponder is lost, stolen, or passed to another person, the other person gains access to the room or car the same way the primary user would.

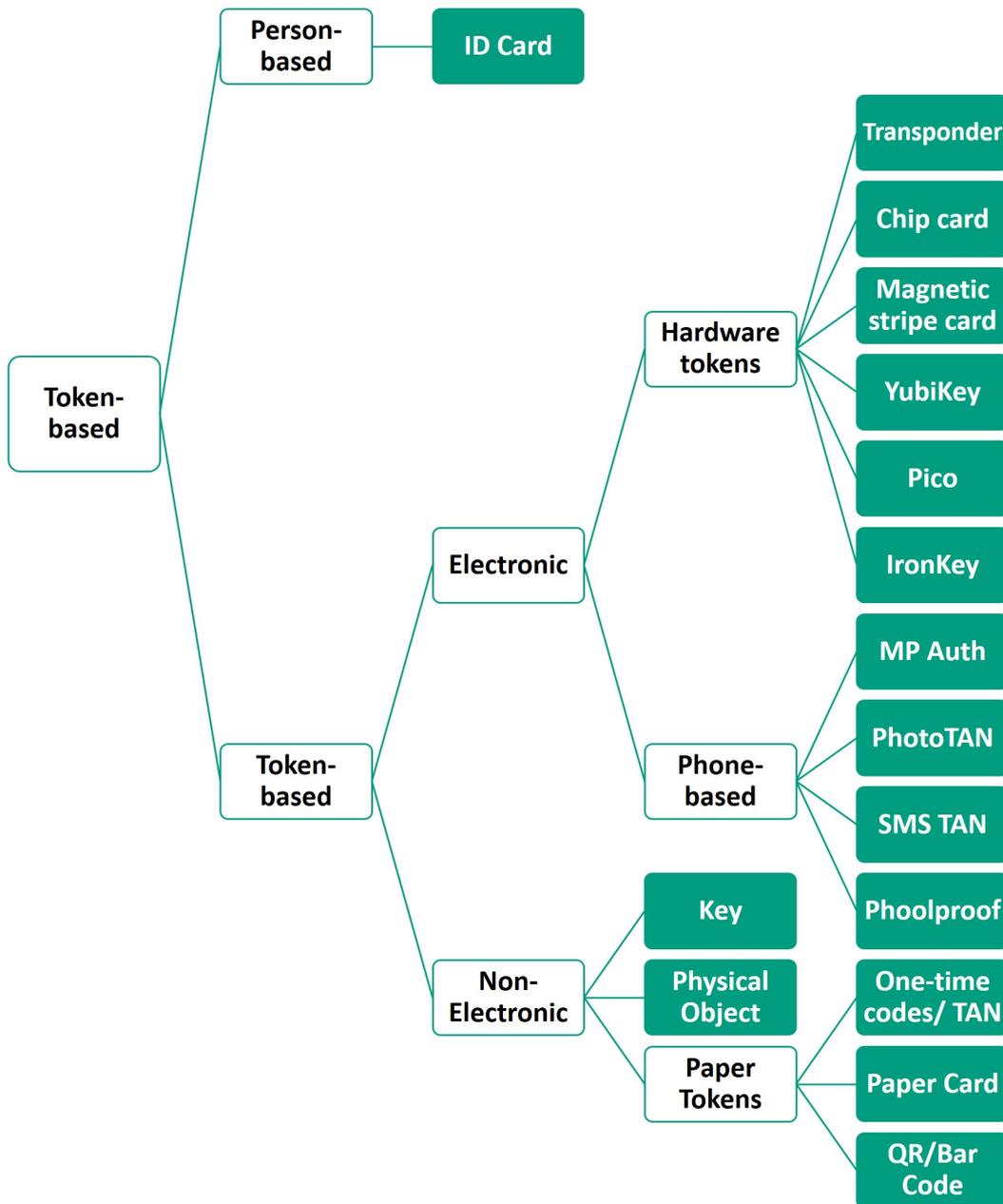


Figure 5: Categorization of token-based authentication schemes. *Note:* The list is not exhaustive. For references to the examples listed in the figure see Appendix 10.5.

This might have benefits in terms of usability, e.g., when a person aims to borrow someone's car, but impacts security if the loss is not noticed in time [261]. On the other hand, a loss of a physical object such as a transponder might be noticed earlier than the stealing of a password by an attacker [218], and a missing transponder can be revoked quickly and remotely via a web interface. This benefits security and also decreases cost and effort compared to the loss of a physical key that requires exchanging the lock cylinder and all distributed keys.

After introducing password authentication as well as existing alternatives including other knowledge-based schemes, biometric, and token-based authentication schemes, the next section analyses and compares the objective features and well as the user perceptions of these different forms of authentication.

This chapter first details the identification and rating of different authentication schemes in terms of objective features (see section 2.1). Afterwards, the results of the rating process were used for the selection of authentication schemes that are analysed in terms of the users' perceptions in a laboratory study (see section 2.2). Finally, the findings of the laboratory study led to a follow-up study focusing on the influence of the context of use on users' perceptions of authentication schemes that is described in section 3.

2.1 Rating of Authentication Schemes

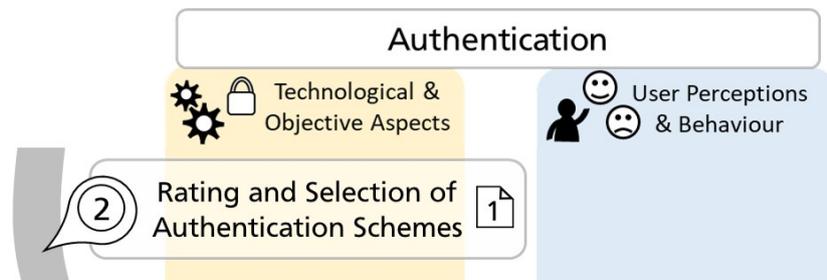


Figure 6: Research step 2.

Manuscript 1: Zimmermann, V., Gerber, N., Mayer, P., Kleboth, M., von Preuschen, A., and Schmidt, K. Keep on rating—on the systematic rating and comparison of authentication schemes. *Information & Computer Security* 27, 5 (2019), 621–635. doi:10.1108/ICS-01-2019-002

Based on a systematic literature review Velásquez, Caro, and Rodríguez [319] identified eight decision frameworks aiming to facilitate comparison and selection of authentication schemes. Some of these consider certain, specialized contexts or requirements. For example, one framework concerns Internet Protocol Multimedia Subsystems [85], one focuses on supporting the design of knowledge-based authentication schemes [97], and a third one compares paid user authentication methods in Korea in terms of managers' preference [166]. The comparison of schemes in terms of their resistance to certain types of attacks by Wang *et al.* [321] concerns two-factor authentication, whereas this research was foremost interested in single factor-authentication. Altinkemer and Wang [11] analysed the costs and benefits associated with implementing a new authentication scheme or a combination of schemes from an economic perspective, e.g., in terms of implementation costs and market share. Guel [118] suggests a framework mainly focusing on technical requirements in terms of authentication, authorization, and server attributes.

Other frameworks can be more broadly applied. For example, Palmer [225] proposes an approach for selecting the most suitable automated personal identification mechanism, short ASMSA, as a decision framework for authentication schemes. It considers the organisational and user perspective and consists of three stages: 1) understanding of strategic goals, 2) effectiveness of requirements, and 3) efficiency of solutions. For each step, a list of criteria is identified that should be considered in the decision process, such as the task environment characteristics or the stakeholders' compromise ability. While very comprehensive and helpful, the framework is also very extensive, including over 200 criteria that need to be discussed in a given organisational context for each scheme. This makes its application to a large number of schemes rated outside of a certain context impracticable.

Thus, to assess and compare objective features of different authentication schemes, the framework proposed by Bonneau *et al.* [34] was chosen. This framework includes 25 rating features categorized into three dimensions: usability, deployability, and security. Each feature is described along with rating criteria detailing when a certain feature is given or not. An example is provided by the feature "Memorywise-Effortless" belonging to the usability features. It is given if the user does not have to memorize any secrets, *quasi*-given if the user has to memorize one secret, and not given if the user has to memorize more than one secret. Furthermore, the rating can be conducted independently of a certain organisational context.

Instead, different contexts and requirements resulting from different contextual factors can later be applied to judge which of the rating features are especially important and to exclude authentication schemes not possessing the features.

Manuscript 1 "*Keep on Rating - On the Systematic Rating and Comparison of Authentication Schemes*" by Zimmermann *et al.* [349], that is an extended version of Zimmermann *et al.* (2018) [348], details how various authentication schemes were chosen from a literature review and rated using an adapted version of the rating framework by Bonneau *et al.* [34]. The adapted version by Mayer *et al.* [197] included a refinement of the rating criteria and a procedure for pairwise comparisons of schemes falling into the same category. The results were integrated into the database of ACCESS¹ [198, 247], an authentication choice support system, that builds on the refined rating system. Thereby, the number of schemes in the database was increased from 45 to 85 schemes. The process is graphically depicted in Figure 7.

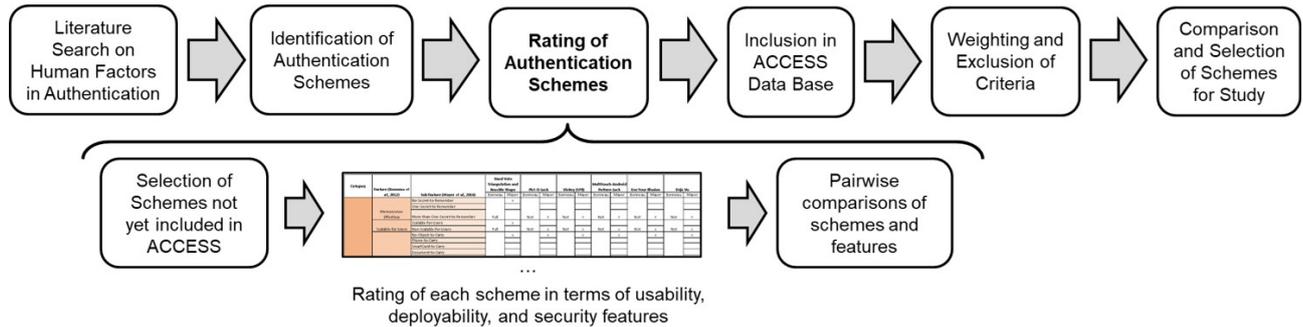


Figure 7: Procedure of the rating process conducted to compare and select authentication schemes for this research as described in [348].

As described in Manuscript 1, ACCESS allows for comparing authentication schemes against each other in terms of the included rating features. It is not possible to quantify which of the schemes performs "best" overall, but to compare the schemes in terms of their suitability for a certain use case. To do so, it is possible to specify requirements, e.g., by excluding rating features that are not important for a certain use case, or by giving more weight to rating features that are deemed more important than others. Given the requirements specification, ACCESS then calculates a performance score for all included schemes so that these can be ranked.

For this research (also see section *Application* of Manuscript 1), the rating was conducted considering a user with a laptop and mobile phone available. In addition to the included usability, deployability, and security criteria, the following requirements were applied for the rating:

- The scheme should produce no or low cost for the users to not reduce acceptance or adoption rates due to financial aspects.
- The scheme should be deployable on web browsers to prevent users from having to rely on additional software or hardware and from not being able to authenticate from different places.
- The scheme should not require users to carry additional items besides the assumed laptop and the mobile phone to decrease the burden for the user and the reliance on additional items that need to be present for each login attempt. Besides, laptops and mobile phones are nowadays often equipped with a variety of sensors, such as cameras, microphones, and fingerprint sensors, allowing for a variety of authentication schemes.
- The feature "accessible" was excluded from the rating, even though it was deemed important. The idea was to avoid bias by excluding schemes not yet providing alternatives for all kinds of impairments in their current form. This is especially important as the rating also included theoretical concepts of schemes that did not have the maturity to provide alternatives for different impairments. Nevertheless, this feature should be considered for the chosen schemes.
- The feature "non-proprietary" was excluded from the rating, as it was deemed irrelevant for identifying the most suitable schemes from a research perspective.

¹ available from: <https://access.secuso.org/>

- The features "resilient-to-throttled guessing" and "resilient-to-unthrottled-guessing" detailing the number of guesses needed to compromise the authentication secret were excluded from the rating. They are implementation-specific and can be adapted to fulfill the required security needs when setting up the chosen authentication scheme. For example, login attempts could be throttled to three unsuccessful attempts before an account is locked for a certain time frame.
- Likewise, the features "no-trusted-third-party" and "server-compatible" were excluded assuming that potential service providers or the researchers within, e.g., the research project "Secure and user-friendly authentication and communication" would be able to set up the chosen scheme without having to rely on external servers or third parties.

The rating resulted in a list of all applicable authentication schemes within the ACCESS database sorted by their performance score. Again, it has to be noted that the schemes on top of the list can not be viewed as the "best" schemes overall, but just as the most suitable schemes given a certain use case, certain weighting and selection criteria, and specific implementations or configurations of schemes. The rating overall confirmed the difficulty to replace the password as none of the authentication schemes included in the database stood out as being superior to password authentication in all regards and across all contexts of use.

Thus, for the next research step, from the resulting rating list, the top two authentication schemes of five different authentication categories, such as knowledge-based and biometric schemes, were chosen. These are listed in section 2.2.2 and described in detail in Manuscript 2. The ten schemes selected from the list were supplemented by the two "challengers" password authentication and fingerprint authentication resulting in twelve schemes overall. These twelve schemes were compared in terms of the users' perceptions in a laboratory study described in section 2.2.2.

2.2 User Perceptions of Different Authentication Schemes

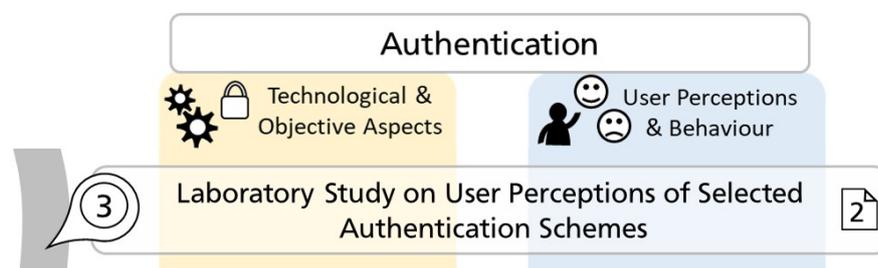


Figure 8: Research step 3.

Manuscript 2: Zimmermann, V., and Gerber, N. The password is dead, long live the password—a laboratory study on user perceptions of authentication schemes. *International Journal of Human-Computer Studies* 133 (2020), 26–44. doi:10.1016/j.ijhcs.2019.08.006.

To compare the twelve schemes in terms of the users' perceptions, first, the experimental setup was tested and refined in a pilot study before conducting the actual laboratory study. The next sections describe the pilot study setup and implications for the main study before summarizing the key points of the main study.

2.2.1 Pilot Study

A pilot study was conducted to test the intended experimental setup, procedure, and the convincibility of mock-ups used to simulate the authentication schemes in the laboratory study. The authentication schemes were chosen to be simulated rather than using existing implementations for several reasons:

First, for privacy reasons, no participant's authentication data should be stored or processed by third parties. This was especially relevant in terms of biometric data.

Second, the use of a simulation allowed for controlling differences between the authentication schemes not related to the authentication mechanism itself, e.g., different maturity levels, error rates, brands and designs.

The primary focus of the pilot study was not on the results in terms of user perceptions but on identifying ways to improve further the study design and the simulation of authentication schemes to increase the validity of the results in the main study. The pilot study will thus be presented with this focus here. The following sections refer to the article "*If it wasn't secure, they would not use it in the movies' - Security Perceptions and User Acceptance of Authentication Technologies*" by Zimmermann and Gerber (2017) [346]. For further details concerning the method and results of the pilot study, the reader is referred to the article.

Pilot Study Method

To evaluate the study design, an ad-hoc selection of eight authentication schemes covering knowledge-based and biometric authentication was analysed: text password, graphical password, gesture recognition, fingerprint recognition, face recognition, iris recognition, speech recognition, and ear shape recognition.

Procedure

After an introduction to the study and the provision of an informed consent sheet, the participants interacted with all eight authentication schemes in a within-subject design. The password authentication served as a baseline; the order of the seven remaining schemes was randomized. The user interaction with the authentication schemes was implemented in the following way: The participants interacted with an interface showing instructions and providing feedback, as well as with the technologies required for authentication (e.g., a fingerprint sensor). The system was designed to convey the intended functionality. However, it was actually operated by the experimenter from a remote work station. No technical functionality of the authentication schemes was implemented. This usability method is also known as "Wizard of Oz" study [65]. The instructions and the system's feedback were presented within a PowerPoint presentation. For all authentication schemes, the design and layout were equal. On the experimenter's screen, the participant's screen was duplicated so that the experimenter could react to the participant's input in time. To maintain the convincibility of the simulation, the experimenter's screen was not visible to the participants. The authentication schemes were simulated using the following apparatus:

- An eye and facial expression tracking system called "FaceLAB" [273] used to simulate the face, iris, gesture, and ear shape recognition,
- A microphone for simulating speech recognition,
- A PowerPoint feature to simulate a graphical password similar to PassPoints [333],
- And a built-in fingerprint sensor of a Sony VAIO notebook to simulate fingerprint recognition.

Following each interaction with an authentication scheme, the participants rated the scheme's perceived security, effort, and cost-benefit ratio, as well as expected usage problems, and intention to use the scheme.

After having interacted with all schemes, the participants were asked to rate the schemes against each other in terms of preference and privacy concerns and to answer some questions in a semi-structured interview. The interview aimed to identify reasons for different ratings, to control for experience with biometrics, and to inform users about the simulation approach used in the study. Finally, the participants were asked whether they had been aware of the simulation and to provide ideas for further improvements.

Sample

The sample consisted of $N = 35$ German undergraduates studying either psychology (29) or psychology in IT (6). Of these, 24 identified as female, eleven as male. The participants were between 19 and 47 years old ($M = 23.09$, $SD = 5.38$). Seventeen out of 35 participants had never used biometric authentication schemes before. The participation was compensated with course credit.

Pilot Study Results and Implications for the Design of the Main Study

First of all, the results revealed differences in user perceptions in terms of privacy concerns and expected problems (see [346]). This can be viewed as an indication that the type of authentication scheme is a relevant factor for influencing user perceptions.

Second, the perceived effort did not differ between authentication schemes. This might be due to all authentication schemes being designed in a similar way and with a zero-error rate to avoid influences resulting from different stages of maturity.

While this can be viewed as an indication for successful unification of authentication schemes, current implementations of authentication schemes (e.g., fingerprint sensors vs. face recognition) used in everyday life do differ in their error rates. Therefore, the main study implemented realistic error rates for the analysed authentication schemes to increase the external validity of the findings.

Third, the simulation approach was efficient in that a majority of $n = 22$ participants did not see through the simulation and felt like interacting with actual authentication schemes.

Eight participants mentioned doubts after having been informed about the simulation, five participants raised questions during the experiment or clearly stated to have been aware of the simulation in the interview. Four of the five participants had previous experience with biometrics. Some participants felt that the authentication process had been too “smooth” to be realistic. Thus, to make the simulation more realistic, especially for people having experience with different authentication schemes, the main study used enhanced HTML mock-ups that reacted to the participants’ input instead of the experimenter’s. Further, this once more encouraged implementing realistic error rates and authentication-scheme specific problems.

Fourth, the authentication schemes analysed in the pilot study had been selected ad hoc informed by related work. However, to be able to compare objective features with subjective perceptions, and to include relevant schemes for researchers and practitioners, the choice of authentication schemes for the main study was based on the rating process described in Zimmermann *et al.* [349] (see Manuscript 1).

2.2.2 Summary of the Main Study

Please note that the study and its results are described in detail in Manuscript 2 *“The password is dead, long live the password – A laboratory study on user perceptions of authentication schemes”* by Zimmermann and Gerber (2020) [347] that has been published in the International Journal of Human-Computer Studies.

As described above and in Manuscript 2, twelve authentication schemes representing different types of authentication were compared in terms of the users’ perceptions in a laboratory study:

- *“Challengers”*: password authentication, fingerprint authentication
- *Knowledge-based (text)*: personal questions, e.g. [264], preference-based authentication [147]
- *Knowledge-based (graphical)*: PassPoints [333], Persuasive Cued Click Points [52]
- *Cognitive*: Weinshall [328], associative questions [143]
- *Biometric*: signature recognition, e.g. [208], keystroke dynamics, e.g.[42]
- *Token-based (phone)*: PhotoTAN [220], MPAAuth [191]

Method

In a between-subject study, $N = 41$ participants interacted with and rated realistic, interactive HTML mock-ups of all twelve schemes. Mock-ups rather than actual schemes were chosen to a) control for different designs, b) control for different maturity levels of schemes, and c) to respect the participants’ privacy in that no personal data was collected or transferred by the schemes.

Procedure

After receiving information on the study and agreeing to participate, the participants completed twelve rounds of registration and authentication with the mock-ups of the selected authentication schemes in a randomized order. To increase the realism of the study and as authentication is a secondary task, the participants were asked to authenticate for a fictional email account. They had to complete a number of short, different tasks, such as checking a date in an email or replying to a question in an email. After each task that required authentication with another scheme, the participants rated the scheme in terms of preference, security, security-readiness [295, 296], privacy, usability, effort(-benefit ratio), expected problems, and intention to use. After having interacted with all twelve schemes, the participants were furthermore asked to rate the schemes against each other and to provide their opinions in a short follow-up interview. The study procedure is graphically depicted in Figure 9 along with an exemplary screenshot of the mock-up for the scheme Persuasive Cued Click Points [52].

Sample

The $N = 41$ participants were undergraduate students with majors in psychology and psychology in IT that were compensated with course credit. The participants' mean age was $M = 21.8$ ($SD = 2.82$) years. A total of 33 participants identified as female, eight as male. Twenty-four participants reported previous experience with biometrics.

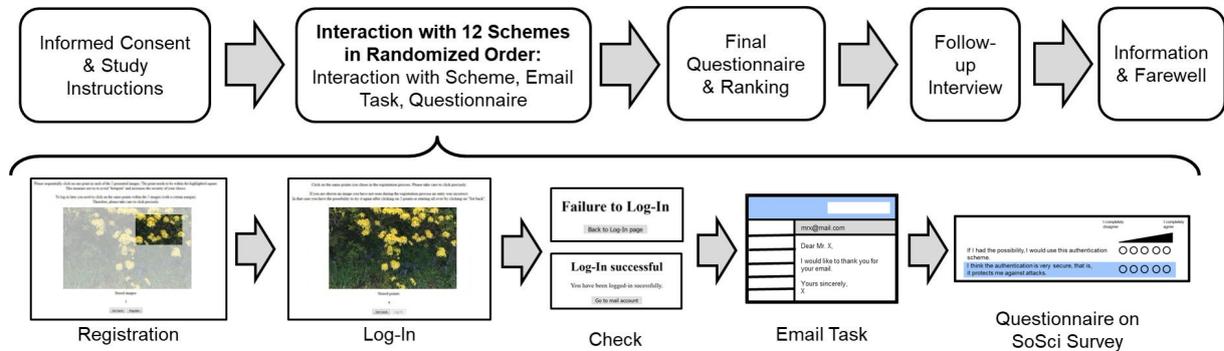


Figure 9: Study procedure of the laboratory study on user perceptions of authentication schemes, adapted from [347].

Results

Conducting repeated-measures analyses of variance (RM ANOVAs), significant differences were found in the user perceptions of the twelve schemes in terms of preference, perceived usability, problem expectation, perceived security, privacy concerns, intention to use, security readiness, perceived effort, and perceived effort-benefit ratio.

The findings further revealed, that despite its shortcomings in terms of the cognitive load it poses for the users, the "challenger" password authentication was still rated highest in terms of preference, usability, intention to use, and lowest in terms of expected problems and effort. As reasons for preferring password authentication, the users often mentioned familiarity with the scheme ($n = 11$), its ease of use ($n = 4$), perceived security (depending on the password creation and handling) ($n = 4$), and also memorability ($n = 4$).

The second highest preference rating was received by the fingerprint scheme that also received favourable ratings in terms of usability, effort, expected problems, and intention to use. Furthermore, the fingerprint scheme perceived the highest security rating, but also scored high in terms of privacy concerns.

While usability, effort, intention to use, and problem expectation ratings correlated with the users' preference, security and privacy ratings were not related to preference. Nevertheless, when asked for the reasons for preferring or disliking a certain scheme, security and privacy were often mentioned. From the privacy ratings across all schemes, it seems that biometric schemes, and schemes asking for the users' personal information, such as the personal knowledge questions, raised greater privacy concerns than the other schemes. Furthermore, while biometrics were rated as very secure, often privacy issues were mentioned in the follow-up interviews as a reason for not intending to use biometrics.

Discussion & Implications

Overall, the results provided relevant insights in terms of user preferences and comparisons of subjective perceptions with objective features from the previous rating. For example, the password and fingerprint were not selected for the study based on their high performance score in the rating, but as wide-spread "challengers". Yet, that they were more preferred than the ten higher-scoring schemes from the rating indicates that they live up to their label as "challengers", at least from a user perspective.

One reason for the finding that security and privacy were uncorrelated with user preference might be that security and privacy are complex and often invisible constructs. Thus, it might be difficult for users to include these aspects in their decision process. The finding thus points to the need to make security and privacy features of schemes visible for the users.

Aside from this, the study also led to the assumption of a perceived security-privacy trade-off of biometric and knowledge-based schemes. That is, while biometric fingerprint authentication was well-liked for its perceived ease of use and security features (uniqueness of information, security against theft and forging) by some participants, it was also often disliked because of perceived privacy concerns for providing personal information.

Thus, fingerprint authentication was the only scheme ranging both among the most preferred ($n = 6$) and the least preferred schemes ($n = 5$). Instead, password authentication was rated as less secure, but also as less privacy-invasive as no personal information was required.

However, it has to be considered that the laboratory study only included a relatively small and homogeneous sample of participants. Thus, to shine light on this seeming trade-off, a follow-up study was conducted to analyse security and privacy perceptions of biometric and knowledge-based authentication schemes in more detail and with a larger, more heterogeneous sample.

Another factor that seemed to influence the users' perceptions of authentication schemes was the context in which the schemes were used. For example, fingerprint authentication was most often reported for unlocking the mobile phone and known from border control when travelling. In contrast, password authentication is widely implemented in online services of all kinds, such as email accounts, social networks, or online shopping. Thus, the follow-up study considered the context of use as a second aspect to analyse in more detail.

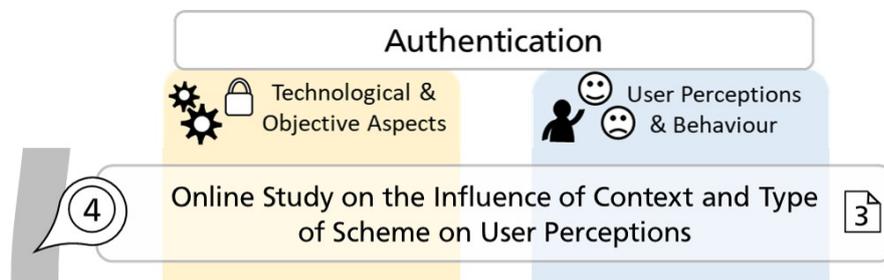


Figure 10: Research step 4.

Manuscript 3: Zimmermann, V., Gerber, P., and Stöver, A. That Depends – How Context Affects User Perceptions of Authentication Schemes. (*in preparation*).

This section details the follow-up study to analyse the impact of the context of use and the assumed security-privacy trade-off on the users' perceptions. The first part describes the pilot study, its setup, and implications before the main study and its findings are described after that.

3.1 Pilot Study

To test a study design for analysing the users' perceptions in terms of security and privacy of authentication schemes, and the users' perceptions in terms of different contexts of use, a pilot study was conducted. It was completed with the support of a student group within the course "Project and process management" that is part of the psychology master's study program at Technische Universität Darmstadt. The student group was supervised by Nina Gerber and Verena Zimmermann, and the study results incorporated in the publication "*Security vs. privacy? User preferences regarding text passwords and biometric authentication*" that has been presented at the conference "Mensch und Computer" [105].

3.1.1 Pilot Study Method

Procedure

In an online survey implemented on SoSci Survey, the $N = 129$ participants were first provided with some general information on text passwords and biometric authentication. The participants were then asked to indicate their preference for the two and to provide reasons for their choice. Afterwards, seven contexts of use were described in a randomized order, among others, online banking, social networks, cloud services, and notebook unlocking. For each context of use, the participants were again asked to indicate their preference in terms of text passwords and biometric authentication and provide reasons in a text field. All participants that were unfamiliar with at least one of the contexts of use or indicated not to have any preference were excluded from the further analysis resulting in $N = 95$ participants.

Sample

A total of 55 out of the 95 people identified as female, 40 as male. The medium age was $M = 28.71$ ($SD = 11.88$). Most participants (51.54%) were students, followed by employees (43%), and people being self-employed (3%).

3.1.2 Pilot Study Results and Implications for the Design of the Main Study

Not given any context, $n = 52$ out of the 95 people preferred password authentication over biometric authentication. Categorizing the qualitative responses following an open coding approach with two raters revealed efficiency and perceived security as the main reasons for the users' preference.

Interestingly, the $n = 43$ participants preferring biometric authentication did so mainly for the same reasons. In all contexts of use, the password was more often preferred than biometric authentication. The descriptive differences were smallest for online banking and online shopping and largest for social networks and email accounts. McNemar tests for dependent samples revealed differences between the users' overall preference and their preference indicated for the following contexts of use: social networks, cloud services, and email accounts. The protection of personal data, i.e., privacy, was rated most important in contexts involving financial information such as online banking and online shopping.

First of all, the roughly 50/50 split of participants overall preferring password authentication or biometrics indicates that there might be different user groups with different needs that may not be fulfilled by only one standard authentication scheme for all users. Thus, online service providers might consider offering multiple authentication schemes or combinations thereof.

Second, while the findings suggest that the context of use may indeed influence the users' perceptions of authentication schemes and the perceived importance of privacy, privacy concerns were not among the main reasons for preferring text passwords over biometrics. However, this may also be due to the format of the questions asking for reasons for the users' preference rather than justifying why the other option was not chosen. Apart from that, it may have been difficult for some users to rate text passwords as one concrete authentication scheme against the group of biometrics that were not further limited in the study to one concrete scheme.

Furthermore, it was found that habit or familiarity with a scheme in a certain context of use may act as a predictor for preference, as the contexts with the highest preferences for text passwords, e.g., social networks and email accounts, were the ones that often use password authentication as the standard scheme. This could also be partially confirmed by the users' qualitative responses.

The findings led to the following implications for the main study:

- First, to complete the picture also token-based schemes were included in the main study to cover the three main categories of authentication schemes.
- Second, for each authentication category, a specific authentication scheme was introduced. These were the three schemes scoring highest in terms of the users' perceptions in the previous laboratory study: passwords (knowledge-based), fingerprint authentication (biometric), and PhotoTAN authentication (token-based).
- Third, familiarity with each scheme was measured to analyse its role as a factor influencing user preference.
- Fourth, user perceptions, including privacy and security perceptions, were collected for all schemes, also the less preferred, to shine light on the assumed privacy-security trade-off.

3.2 Summary of the Main Study

Please note that the study and its results are described in detail in Manuscript 3 *"That Depends - How Context Affects User Perceptions of Authentication Schemes"* by Verena Zimmermann, Paul Gerber and Alina Stöver [351] that was being prepared for submission at the time of the publication of this dissertation.

The manuscript describes an online study in which $N = 202$ participants rated their perceptions of different authentication schemes across different contexts of use. The study aimed to analyse the influence of the type of scheme and the context of use on the user's perceptions in general, and on the assumed privacy-security trade-off in particular.

3.2.1 Method

The online study was conducted using the platform Clickworker and used a within-subjects design. Thus, all participants rated all authentication schemes in all contexts of use.

Based on the results of the laboratory study described in Manuscript 2, the three best-rated authentication schemes in terms of user perceptions from each authentication category were chosen, resulting in the following selection:

- *Password authentication:* As described in the previous sections, password authentication is a representative of the knowledge-based authentication schemes. The user memorizes and types in an alphanumeric secret to authenticate.
- *Fingerprint authentication:* Fingerprint authentication belongs to the group of biometric authentication schemes. To authenticate, the user needs to register the fingerprint or certain points thereof that are compared with the fingerprint upon login using a fingerprint sensor.
- *PhotoTAN authentication:* This scheme is a token-based scheme using the smartphone as a token [220]. An individual secret key is generated and stored on the smartphone. The key and thus the smartphone is necessary to extract a one-time code from a colored pattern that is created by the website or service the user aims to access. The user is authenticated upon entering the correct one-time code on the website.

The following contexts of use were selected based on their spread and relevance, and also in terms of the different types of data protected by authentication in each context:

- *Email Account:* Email accounts can include personal and work communication of varying sensitivity, and are often used as a possibility to reset forgotten passwords of other accounts.
- *Social Network:* Social networks can be used to share personal and social information including text messages, pictures or locations with friends or a wider range of people, e.g., on Facebook or Instagram. Other networks are specialized in sharing certain types of social information such as work-related information, e.g., on LinkedIn or Xing, or academic information, e.g., on Google Scholar or ResearchGate.
- *Online Banking:* Online banking accounts generally provide access to personal financial information such as the account balance and types of transactions, and also provide functionalities such as initiating transactions or withdrawals of money.
- *Smart Home:* Smart home technologies and devices intelligently connect household appliances, sensors, and consumer electronics [350]. Exemplary smart home devices include smart TVs, smart speakers, smart heating, or lighting. Smart home devices can hold personal information of varying degrees of sensitivity such as motion detection from movement sensors, personal routines such as wake-up times, personal images from security cameras, or health-related data from smart health devices.

Procedure

At the beginning of the study, the participants were presented with some information on the study and an informed consent sheet. After having accepted to participate, the participants were provided a textual and visual depiction of the registration and login process for each of the three schemes in a randomized order. This part was followed by a rating of the familiarity with and the frequency of use of each scheme on a 100-point visual analogue scale.

In the main part of the study, the participants were presented scenario descriptions of the four contexts of use and asked to rate their preference for using each scheme for the respective context of use on a 100-point scale. Afterwards, the participants could select qualitative reasons for their preference rating from a matrix that was based on the findings from a literature review and previous studies [105, 346, 347]. For example, the participants had the option to explain their preference rating by choosing low vs. high security, learnability, or intrusiveness of the scheme. Besides preference, the participants were asked to rate the perceived security, privacy concerns, effort-benefit ratio, and their intention to use a certain scheme within a certain context of use on a visual analogue scale ranging from 1 to 100. The sequence of the contexts of use was randomized to balance sequential effects.

Finally, the participants were asked for their general perceptions of all three schemes independent of a certain context of use. Besides the measures already mentioned, this part included a rating of the subjective usability of the scheme using the System Usability Scale (SUS, [38]) and a rating of the scheme's efficiency. The study ended with the collection of basic demographic information, the participants' technological affinity [98], and general privacy concerns based on Smith, Milberg and Burke [286] and adapted by Malhotra, Kim and Agarwal [190].

For an overview, the study procedure is depicted in Figure 11.

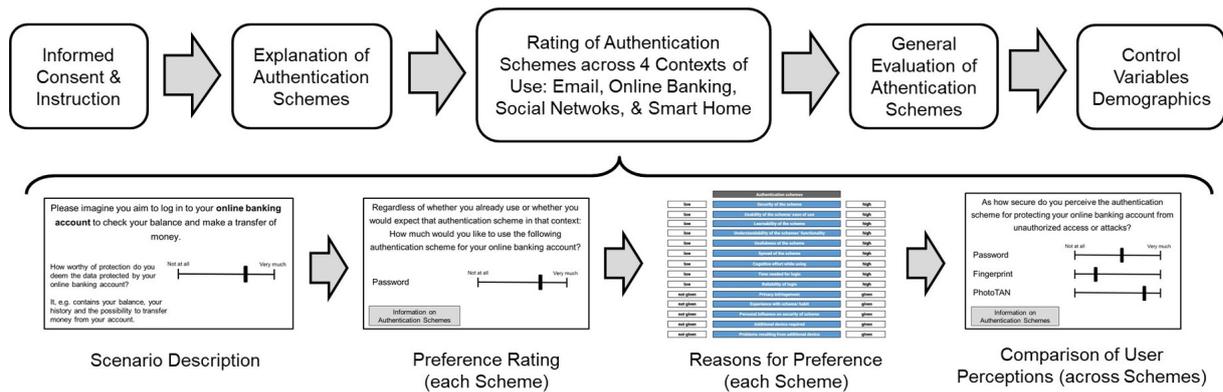


Figure 11: Study procedure of the online study on the influence of the context and type of scheme on user perceptions of authentication schemes [351].

Sample

The sample consisted of $N = 202$ German participants that were recruited via the online survey panel Clickworker and compensated according to the platform's suggestion and above minimum wage. A total of 98 identified as female, 103 as male, and one as diverse. The sample's age distribution was as follows: $n = 25$ were between 18 and 24, $n = 86$ were between 25 and 34, $n = 22$ were between 45 and 54, and $n = 15$ were between 55 and 64 years old.

3.2.2 Results

Repeated-measures multivariate analyses of variance (RM MANOVAs) were conducted to analyse differences in the users' perceptions across schemes and contexts. Linear regressions were used to analyse the influence of certain predictors on the users' preference for a certain scheme. Due to deviations from normality, robust procedures based on trimmed means and bootstrapping were used for the subsequent analysis wherever possible.

Despite its downsides in terms of cognitive effort, that was even acknowledged by the participants, the password was most preferred in general and across all contexts of use. The relatively new PhotoTAN scheme that is increasingly deployed in the banking sector (for which it also received the highest relative preference rating) was the least preferred overall. Preference ratings concerning the fingerprint scheme were ambiguous. While some people highly preferred the scheme, others highly rejected it. Furthermore, the fingerprint scheme was rated as the most secure but also received the highest values in terms of privacy concerns.

The results revealed that the type of scheme as well as the context individually influenced the participants' perceptions of the schemes. However, while all measured perceptions, including privacy and security, differed across the schemes, the privacy and security perceptions did not differ across contexts of use. The finding was mirrored when analysing the interaction effects of scheme and context on the participants' perceptions that were significant except for security and privacy perceptions. When analysing the extent to which certain measures influenced the participants' preference ratings, it was found that security perception, the perceived effort-benefit ratio, and perceived usability were relevant predictors across a number of schemes and contexts.

3.2.3 Discussion & Implications

The high preference ratings in terms of password authentication confirm the findings from the laboratory study on user perceptions [347] and show that users prefer the password despite its downsides in terms of cognitive load. Likewise, the ambiguity in fingerprint preference ratings hints at a split in the user group already found in previous research [49, 68, 254, 347].

The high level of security perceptions and privacy concerns at the same time associated with fingerprint authentication offer an explanation for the ambiguity in terms of preference. This can also be viewed as an indication of the assumed security-privacy trade-off.

Furthermore, the finding that security and privacy perceptions differed across schemes but not contexts indicates that these perceptions are closely tied to the scheme. Likely, security and privacy may somehow be traded-off with other aspects such as usability or an estimation of whether the security and privacy level are "acceptable" for the respective context of use. An indication for this trade-off might be that security perception was among the significant predictors for preference ratings, along with usability perceptions and the effort-benefit ratio. The effort-benefit ratio itself constitutes a trade-off that might include other aspects not directly measured and bears similarities to related concepts such as "security readiness" [295, 296], the effort people are willing to take for security given differing levels of data sensitivity.

Finally, the research revealed indications for the influence of familiarity on the participants' perceptions, yet, these could not be confirmed when analysing the predictors influencing user perceptions. Reasons for this finding might be that familiarity has been measured in general as opposed to context-specific, and that the variability in familiarity was constrained as, e.g., most people were highly familiar with password authentication.

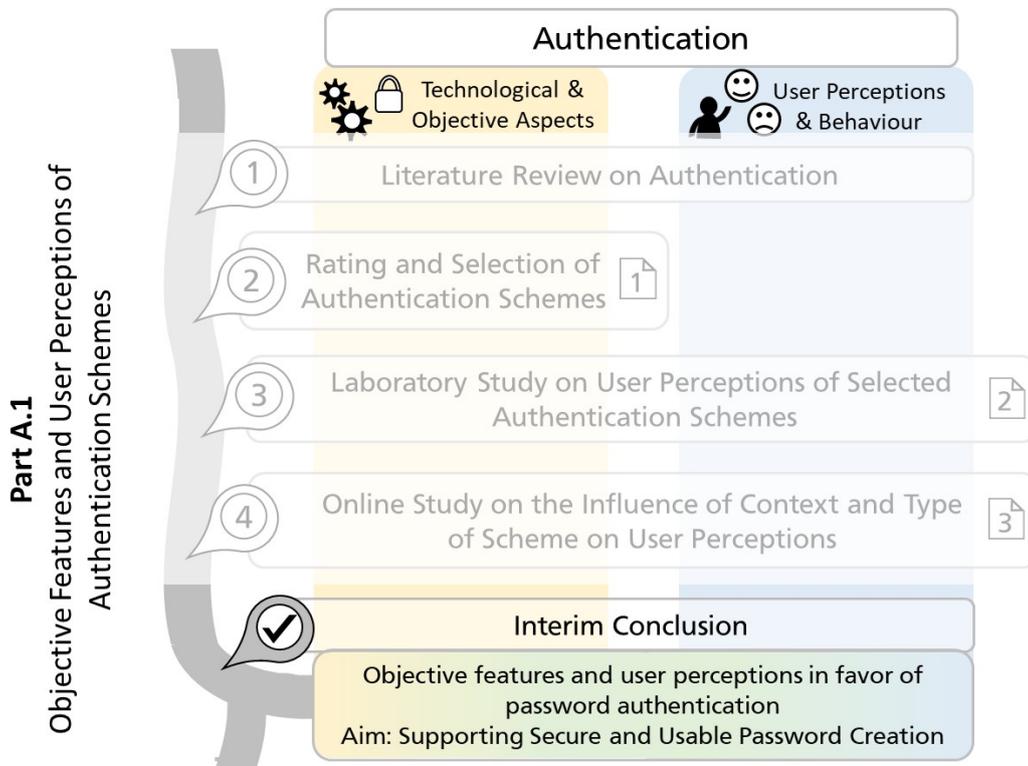


Figure 12: Interim conclusion.

This research started with a rating of objective features and an analysis of subjective perceptions of a large number of authentication schemes with the initial aim of finding or developing a secure, while usable alternative to password authentication. Yet, the findings revealed that the quest to replace passwords persists:

The extensive rating of authentication schemes in terms of objective usability, deployability, and security features updated and extended the original database of Bonneau *et al.* [34], but did not lead to the identification of an authentication scheme outranking password authentication across all rating criteria. Rather, it became obvious that all classes of schemes have certain advantages and shortcomings. This is also true for password authentication that has, for example, often been criticized for the memory load it poses for users and for being susceptible to guessing, phishing, and keylogging attacks on the one hand [133]. On the other hand, password authentication provides some important advantages. It does not require client hardware as, for example, fingerprint authentication or smart cards do [133]. Passwords do not require tokens that would need to be carried, and that are often expensive. Forgotten passwords can be easily replaced using automated self-service resets via secret questions or by sending reset links to an associated email address. Passwords have shorter login times than other forms of authentication, e.g., some graphical or cognitive schemes [187]. They allow for setting up an account at once without having to wait for some digital validation or setting up a physical device. Finally, passwords are browser-compatible, allowing access to accounts from anywhere in the world. Thus, the additional security features offered by other forms of authentication might not always justify the cost [133].

The subsequent comparison of subjective user perceptions of authentication schemes that have rarely been subject to laboratory studies involving actual interaction before, shed light on the seemingly intractable issue. However, the findings were different than expected. It was found that the password, despite its shortcomings, especially with regard to the cognitive load, was preferred by the users. This finding could be confirmed not only in the laboratory study but also in two consecutive online studies (pilot and main study) analysing users' preferences across different contexts of use.

Thus, the "*overwhelming hatred from the users*" [282, p.13] attested by other researchers could not be confirmed in this line of research.

Instead, the participants provided valid reasons for preferring the password, at least when compared with other authentication schemes that all have their advantages and shortcomings. In the laboratory and online studies, password authentication was rated high in terms of usability and low in terms of privacy concerns, effort, or expected problems. The users' preferred password authentication out of habit and familiarity with the scheme, but also - contrary to the expectation - perceived it as relatively secure and memorable. The online study further revealed that the preference for password authentication was stable across different contexts of use when compared to fingerprint and PhotoTAN authentication.

Furthermore, the users' positive attitude towards and their high intention to use passwords are both predictors for acceptance and actual use following the Technology Acceptance Model (TAM) [66]. It is thus likely that users will continue to use password authentication for some more time when given the choice. In line with that, researchers acknowledged the persistence of passwords [133], even if for other reasons: Herley and van Oorshot argue that security experts aim to replace passwords without understanding what is required to do so, or what the actual achievement will be. Significantly improving usability and security at the same time without causing disruption is deemed unlikely. Given the variety of account types and assets protected by them, Herley and van Oorshot also view it as naive to expect a single solution for all use cases. Another problem concerns the difficulty of estimating harm caused by compromised passwords, but also the harm prevented by and cost for implementing alternative schemes. Given the lack of a "silver bullet" meeting "*all goals in all situations*" [133, p.32], best-fit solutions should be explored. For many authentication needs, these might include password authentication. That said, Herley and van Oorshot rate the belief that passwords are dead as incorrect and harmful, and instead, encourage a research agenda for improving password authentication. In line with that, Siddique, Akhtar, and Kim [282] find that promoting biometrics as the ultimate in authentication is unfounded and even suppressed valuable password-related research.

Acknowledging passwords as the most widely-spread authentication scheme, with a number of advantages in terms of objective features, and as positively perceived by the users, the second part of this research will focus on password authentication. This is not to say that password authentication was the "best" or should be the sole form of authentication. Indeed, the previous studies also revealed preferences for other forms of authentication (especially fingerprint authentication), and different groups of users preferring different authentication schemes. Furthermore, differences in user perceptions were apparent across different contexts of use in the online study. Aside from legal, technical, or security-related requirements, these user-centred aspects should be taken into account when choosing authentication schemes for a certain service or context.

Password authentication will not only be focused on in the following because it is wide-spread and well-perceived but because improving password security remains an open challenge. Human factors play an important role therein: Aside from technical measures, the security of password authentication is largely impacted by the users' decisions and behaviours, i.e., their password creation and password handling. As formulated by Jain "*the security of the entire system is only as good as the weakest password*" [145, p.14]. The password's shortcomings that especially concern human factors, i.e., human cognition, perception, and behaviour, remain unsolved and require further research. Therefore, instead of developing "*yet another authentication scheme*" [133, p.29], the second part of this research focuses on supporting users in creating secure and memorable passwords.

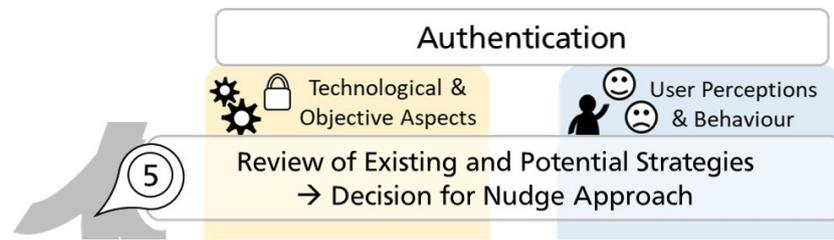


Figure 13: Research step 5.

A variety of measures have been developed to increase password security. While some of the strategies rely on increasing security by constraining the users to comply with a set of rules or technical measures, others aim to increase the security by reducing cognitive load or by providing support in creating secure passwords. All the strategies have advantages and disadvantages that will be discussed in the following.

Please note that this section serves to review common strategies and to explain the reasons for focusing on a certain type of measure in the second part of this dissertation. However, it has not been included in one of the manuscripts.

5.1 Constraining Strategies

Constraining strategies focus on providing technical measures or rules to prevent users from creating weak passwords.

5.1.1 System-generated Passwords

Early on, users were often provided with system-generated passwords [5] that are generally stronger than user-generated passwords [203]. Yet, system-generated passwords consisting of a random sequence of alphanumeric symbols are also generally harder to memorize as they bear no meaning for the user, and lead users to write passwords down [5, 224, 355]. Even the technique of chunking, as tested for PINs of different length, had promising but only limited impact on the memorability of system-generated PINs [141]. For these reasons, user-generated instead of system-generated passwords became most widely used in the long-term [5].

5.1.2 Regular Password Expiration

Regular password expiration requires users to change their passwords at regular intervals with the idea to limit the opportunity to expose the password or the damage an attacker could do to a system with a compromised password, respectively [50].

Even though a long-upheld practice in industry and universities [95], empirical studies suggest that regular password expiration does not achieve the security aims hoped for and should be called into question [53, 344]: Zhang *et al.* [344] studied a large data set of previous university account passwords that had to be changed every three months and found that for 17% of the passwords an attacker knowing the last password would be able to guess the new password in less than five attempts. Furthermore, 41% of the accounts could be compromised within three seconds in an offline guessing attack. The finding that users tend to use slight variations of previous passwords as new passwords, e.g., by appending numbers to previous passwords [119], making it easier for attackers to guess passwords based on prior knowledge, was also confirmed by Chiasson and Van Oorshot [53]. Furthermore, it may be possible for an attacker that had once gained access to an account to install malware or a keylogger allowing him or her constant access even when passwords are changed [64].

Apart from that, regular password change has also been found to be a "source of frustration" for users [262, p.ES-2] and a "usability disaster" [133, p.34]. A recent online study on the users' self-reported behaviour and attitudes concerning mandated password change [119] indicated that the practice might neither negatively impact password strength of new as compared to previous passwords, nor does it increase password or account security overall as password changing strategies were found to be predictable.

Given the questionable benefit in terms of security and the downsides in terms of usability, organisations such as the National Institute of Standards and Technology (NIST) [115] now suggest that passwords do not need to be changed unless there is evidence of compromise.

5.1.3 Password Policies

Password policies are a form of restricting user-generated passwords to include a minimum number of characters and certain types of character sets. Often password policies also forbid the use of frequently used words collected on blacklists. A password policy that has been widely used is *comprehensive8* that requires the user to create a password of a minimum eight characters length, an uppercase letter, a lowercase letter, a number, a symbol, and no dictionary words. Yet, this strategy to increase password security and "randomness" has not had the effects hoped for. First of all, the security benefits of applying password policies were smaller than expected [329]. Proctor *et al.* [237] furthermore found that password policies increase the length of the password creation and login procedure and lead to systematic password creation patterns. In line with that, Inglesant and Sasse [142, p.383] have identified several usability issues related to password policies and conclude that existing policies "are too inflexible to match their capabilities, and the tasks and contexts in which they operate". Therefore, also NIST [115] calls for focusing on password length as the primary predictor for password strength [163, 171, 237] as opposed to password composition.

5.2 Supporting Strategies

Supporting strategies focus on lessening the security-usability trade-off or overcoming mismatches between user perceptions and technical security. These can again be divided into approaches aimed at supporting secure password creation and approaches aimed at decreasing the memory load.

5.2.1 Approaches for Creating Secure Passwords

Password Information

One approach for supporting users in creating strong passwords is to educate users about what makes strong passwords, e.g., by presenting information texts next to a password entry field. Exemplary advice might include: "Avoid basing your passwords on easily-guessed information, such as names, birthdays, or things you like (e.g., favourite animal or song). This would make it easy for attackers to guess." The idea is that improved decision making is possible when the user has all the facts [43], i.e., is aware of potential threats and has the knowledge to counteract them.

Yet, research has shown that security advice and security-related information, in general, are often not followed. For example, empirical studies show that privacy policies or terms of service are not read by the majority of people [217] and that users were still found to use weak passwords despite many years of password advice [96]. Potential reasons for not following the advice or applying coping strategies include the overwhelming amount of existing advice, their benefits being (or appearing to be) hypothetical [132], and conflicting advice across websites and services [211, 342].

However, to date, few studies specifically examined the construction of password advice and the passwords created with it [342]. For example, the recent NIST guidelines from 2017 [115] suggesting to no longer enforce password composition policies but to allow for very long passwords as an important predictor for password strength has yet to be implemented by many providers and developers. Thus, evaluations of its effectiveness are still rare [342].

Password Meters

Password meters provide the user with an indication of their created password's strength to allow for aligning the user's perception with technical security measures and for motivating the user to increase password strength without enforcement. Password meters can, e.g., take the form of a coloured feedback bar or text such as "Your password is strong". They are frequently implemented in practice on various websites and in different forms, among them eBay, Facebook, Google, or Twitter [99]. In 2015, Van Acker *et al.* [317] found password meters to be deployed in about a third of the top 250 Alexa domains. In practice, password meters are sometimes combined with additional elements such as nudges (see next section), suggestions for improving password strength, or password policies.

Even without enforcing any rules or minimum requirements, some studies analysing different types of password meters with different user groups found an increase in users' password strength when a password meter was deployed (e.g., [84, 102, 167, 232, 318]). Furthermore, memorability rates did not seem to be negatively affected [84, 167, 232]. At the same time, a number of studies did not reveal significant improvements for some or all of the tested password meter conditions (e.g., [84, 112, 274, 317]) so that findings in terms of the effectiveness of password meters remain mixed.

Furthermore, analysing different password meters implemented in practice revealed the importance of consistent measures and feedback across accounts. Different cracking algorithms and strength estimations used by different researchers and websites can produce inaccuracies and inconsistencies [111, 316] that might lead to confusion among users that receive varying feedback for similar passwords across accounts.

Password Nudges

Nudges aim to encourage users to choose the wise or secure option, respectively, by making small changes to the choice architecture that do not limit the choice set and that activate automatic cognitive processes such as biases and heuristics [303]. The term nudge was introduced by Thaler and Sunstein in their seminal book "Nudge" in 2008 [303]. It describes multiple examples of successful applications across various domains, such as default nudges to increase the number of organ donors or an image of a fly in a urinal to reduce spilling.

Emerging from the field of behavioural economics, nudging has also been applied to the digital world, including cybersecurity in general and password creation in particular. As described above, password meters are often combined with nudges such as colour-coding nudges [313], nudges invoking social norms [84], or fear appeals [318].

Password nudge examples aside from password meters are still rare: Kaleta, Lee, and Yoo [157] nudged users to focus on the desirability of creating strong passwords to make users take a high construal level perspective. They found that passwords were indeed stronger when users were induced to think at a high as compared to low construal level. A study by Kankane, DiRusso and Buckely [158] aimed to increase users' intention to change auto-generated passwords by, e.g., using default and salience nudges. Nicholson *et al.* [214] made use of three interventions, including an incentive, a length instruction nudge, and a priming nudge. The priming nudge consisted of an image of a man in a library carrying many books along with the instruction to create a password based on that image. However, similar to password meters, the findings are mixed: Kankane, DiRusso and Buckely [158] found one nudge to significantly reduce comfort levels with the auto-generated password. Nicholson *et al.* [214] found no significant differences for the priming nudge, but for the length instruction and incentive.

5.2.2 Approaches for Increasing Memorability

Passphrases and Mnemonics

Passphrases and mnemonics are exemplary strategies aiming to enhance the memorability of strong passwords. Passphrase approaches suggest using a complete sentence or a combination of words as a password as compared to one single password [162]. An example would be: "Let'seatsoameicecreamatthemalltomorrowat3pm!". That way, long passwords can be created with the idea that these meaningful sentences might be more memorable than meaningless passwords [235]. The increase of password length is aimed at increasing the time necessary to guess a password, thereby enhancing its security [162].

Mnemonics are a variant of passphrases that suggest thinking of a sentence and then take the first character of each word as a password [180]. The exemplary passphrase above would result in the following mnemonic: "Lesicatmta3p!".

Empirical research found that the strength of random passwords and mnemonics were similar and that both were stronger than user-generated passwords [340]. At the same time, mnemonics were not harder to remember than user-generated passwords [340]. Similar results were found for the passphrases. The memorability of passphrases did not significantly differ from user-generated passwords [162, 280]. However, the use of passphrases increased the number of typographical errors and login times, resulting in negative user perceptions [162, 280]. More recent studies found that typographical errors could be mitigated by using well-designed passphrases that were consistent with a normal word processing mode [161], or by using an algorithm that accepts the most common typing errors [215].

Password Managers

Password managers and single-sign-on services are technical approaches for decreasing the cognitive load for the user [182]. They do so by storing the users' multiple passwords for them. In general, they only require one master password to gain access to all the stored passwords.

Password managers that can be browser-based, cloud-based, or a client application, offer to relieve the cognitive load of memorizing multiple passwords from users [133]. Yet, password managers also face some challenges, including security issues and user adoption rates.

Security threats are posed by a password manager being a single point of failure and the frequent lack of malware-resistance [133]. In addition, Li *et al.* [182] identified vulnerabilities in terms of the use of bookmarklets, web browsers, authorization, and the user interface.

Aside from the security-related challenges, adoption rates are still low despite the promise of lessening the cognitive burden for the user [231, 251]. Among the factors hindering adoption, researchers identified a lack of trust in the security of password managers and an underestimation of the threat of password loss [16, 17, 231]. A study by Alkaldi, Renaud and Mackenzie [7] suggests that meeting self-determination needs, such as autonomy, relatedness, and competence, might positively affect adoption.

In terms of the effectiveness of password managers, Pearman *et al.* [230] found, at least until the point of their study, that the use of password managers or auto-fill tools neither had a significant effect on password strength nor reuse. The findings of Lyastani *et al.* [186], who conducted a large-scale study on the impact of password managers on password reuse and password strength, are mixed. They found that password managers, at least if they include password generators, can indeed positively influence password strength and reduce reuse. Yet, the effect also depends on the user's password strategies, i.e., whether the user relies on the technical support for password creation.

5.3 Evaluation of Strategies for Enhancing Password Security

The review of strategies for secure password creation shows that the strategies can be classified as constraining and supporting strategies. From the comparison above, it appears that strategies supporting users without enforcing rules are not necessarily less effective in increasing password strength as compared to constraining strategies such as mandatory password change or strict password policies. Making users fit technical requirements without further consideration of the human factor might instead lead to frustration or resistance among users, and perhaps even to the use of security-weakening workarounds. This is supported by the finding that the security benefit of the constraining strategies is questionable [53, 329, 344] and accompanied by considerable usability problems [133, 142, 237, 262].

Furthermore, the human-centred stance taken in this research suggests focusing on the strategies that support users in creating and managing secure passwords by considering human factors such as users' knowledge, perceptions, and cognition. For a detailed discussion of the stance taken towards the user in this research also see Manuscript 8 "Moving from a "Human-as-Problem" to a "Human-as-Solution" Cybersecurity Mindset" [353]. Instead of considering the human a weak link or a problem to control, it is assumed that users may well be able and willing to contribute to security, given that they are sufficiently considered and supported.

Thus, focusing on supporting approaches, a range of choices remain. While password information, password meters, and password nudges aim to primarily support password creation, passphrases and password managers primarily constitute strategies for facilitating the memorability of strong passwords.

However, these strategies are not mutually exclusive. For example, password information can contain the suggestion to create passphrases or hint at the use of password managers. Yet, the three strategies for supporting secure password creation address the first, relevant step towards secure passwords. A password manager would only be of limited help when used to store only weak passwords. Thus, password information, password meters, and password nudges remain in the selection process.

While all three strategies appear promising and provide potential for future research, they are also accompanied by mixed findings. Password information is often not followed [96], and password meters and password nudges are sometimes found to be ineffective [84, 112, 214, 274, 317].

The mixed findings in terms of information provision are likely to result from decades of overwhelming and contradictory security advice [132, 211, 342]. Even though empirical research on individual sets of password information is still rare [342], it is unclear whether adding a new set of guidelines to the large amount of advice that also quickly changes with technological developments adds an important benefit for users. Contrary to information provision, the application of nudges to the digital space, including password nudges in cybersecurity, is a relatively new field [327]. Thus, the mixed findings might be due to the still small number of empirical studies.

While some studies across cybersecurity domains revealed the nudge's potential (e.g., [55, 310, 320, 322]), some aspects of nudging are not yet well understood. For example, further research is needed to understand *how* nudges exert their influence and *why* some nudges are effective while others are not [46].

As described above, password meters in practice are often combined with a nudge (e.g., [84, 313, 318]). Thus, it remains unclear whether their (in)effectiveness results from the password feedback alone or from the nudges targeted at motivating users to increase password strength.

Therefore, the second part of this research will focus on designing password nudges as a new and promising field of research. Yet, as already described, the strategies to enhance password security are not mutually exclusive. Thus, after first undertaking efforts to design effective password nudges, consideration will be given to the combination with potentially beneficial aspects of other strategies such as password meters, password information, or password managers.

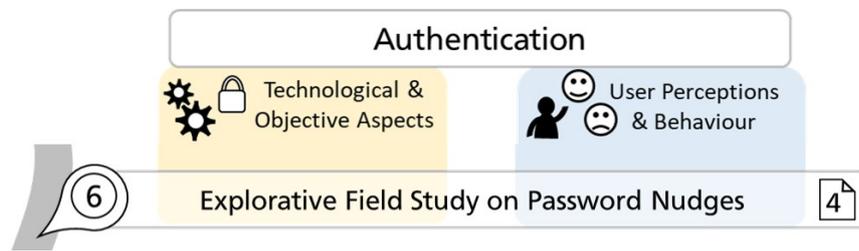


Figure 14: Research step 6.

Manuscript 4: Renaud, K., and Zimmermann, V. Nudging folks towards stronger password choices: providing certainty is the key. *Behavioural Public Policy* 3, 2 (2019), 228–258. doi:10.1017/bpp.2018.3.

Following the decision to focus on designing password nudges as a promising and new area of research, this chapter describes three consecutive field studies that aimed to compare the influence of different variants of password nudges on password creation.

First, the concept of nudging will be briefly introduced. For a detailed introduction to nudging, the reader is referred to Manuscript 5 *“Ethical guidelines for nudging in information security & privacy”* that has been published in the *International Journal of Human-Computer Studies* by Renaud and Zimmermann [249], and Manuscript 6 *“The Nudge Puzzle: Matching Nudge Interventions to Cybersecurity Decisions”* by Zimmermann and Renaud that is has been published by the journal *ACM Transactions on Computer-Human Interaction* [354].

Second, the study design, results, and implications of the field studies will be described. This part is a brief summary of Manuscript 4 that has been published by the journal *Behavioural Public Policy* by Renaud and Zimmermann [252]. Specific aspects of the three field studies are furthermore published as *“Lessons Learned from Evaluating Eight Password Nudges in the Wild”* in the Proceedings of the LASER Workshop by Renaud, Zimmermann, Maguire, and Draper [253], as well as *“Enriched Nudges Lead to Stronger Password Replacements ... but Implement Mindfully”* by Renaud and Zimmermann in the Proceedings of the Information Security for South Africa (ISSA) conference [248].

6.1 Introduction to the Concept of Nudging

With their seminal book *“Nudge”*, Thaler and Sunstein introduced the concept of nudging in 2008 [303]. According to the authors, a nudge describes *“any aspect of the choice architecture that alters people’s behavior in a predictable way without forbidding any options or significantly changing their economic incentives”* [303, p.6].

The term *“choice architecture”* broadly refers to any context in which a decision is made, such as physical contexts but also digital decision interfaces. As stated in the definition, nudges are supposed to preserve all choice options and to not unbalance the choice by introducing significant incentives, not only in terms of money, but also time, effort, or social sanction [129]. Thus, no choice should be enforced. An example would be that of a store in which healthy snacks are placed next to the checkout to encourage healthier choices while people still have the opportunity to buy unhealthy snacks [173].

Nudges are supposed to work by activating certain automatic cognitive processes such as heuristics or biases [43, 122, 129]. Heuristics are mental shortcuts that ignore part of the information to facilitate decision-making [107]. An example would be the availability heuristic that describes how cognitively easily accessible information also comes to mind more easily [107]. A widely-cited example is that of an experiment in which English words with a certain letter in the first position were more easily available than words with this letter in the third position even though these were more frequent [311]. Cognitive biases include some systematic distortion from an objective aspect [126]. For example, the hindsight bias describes the tendency of seeing past events as predictable even though the information was not given before the event.

Among the large number of choices people are confronted with on a daily basis, it is estimated that about 95% are guided by automatic cognitive processes [46], introducing numerous possibilities for nudging.

The nudge concept originates from the area of behavioural economics, where the concept has been applied with the intention to encourage citizens to choose the "wiser" option. Exemplary applications include prompts to encourage stair use as physical activity [29, 165], the use of social comparisons to encourage energy saving [8], or the variation of food proximity and serving tools at a salad bar to reduce food-intake [259]. As of today, several governments including the UK, the USA, and New South Wales in Australia [22, 31, 80, 304] embraced the idea and established units to analyse the effects of nudging for the public welfare.

Beyond the physical contexts, nudging has already been applied to the digital area, labelled "digital nudging" [327]. This concerns a variety of choices, e.g., taken on websites, during smartphone configuration, or software installation. One important area of digital nudging is the field of human-centred security and privacy [2]. For example, Choe *et al.* [55] used a framing nudge to increase users' awareness of the privacy implications of smartphone applications. In a study by Almuhimedi *et al.* [9], users were presented with notifications about how often and with which applications the user's location has been shared to encourage users to review and adapt their privacy-settings. Raja *et al.* [239] used images representing the term "firewall", such as a brick wall, to encourage users to undertake protective measures. Von Zezschwitz *et al.* [320] used background images to encourage users to choose stronger Android Unlock Patterns based on the image.

Finally, as already described in section 5.2.1, initial studies have trialled a number of password nudges such as default or salience nudges [158], a priming image next to the password entry field, or changing the wording of the password creation instruction [214]. Kaleta, Lee, and Yoo [157] designed a technique to nudge users to focus on the desirability of creating strong passwords as compared to feasibility based on construal level theory. Different variants of password nudges such as fear appeals [318] or nudges appealing to social norms [84] have also been used within password meters.

Because these initial studies produced some promising but also mixed results [84, 158, 214, 318], the field studies described in the next section aim to shine light on the effectiveness of password nudges by testing several variations in the wild.

6.2 Field Studies on the Effectiveness of Password Nudges

To analyse the impact of different password nudges on password creation, three consecutive field studies have been conducted. All three studies have a similar setup, were conducted in the same environment, and the design of the password nudges builds on the findings of the respective previous study. Thus, the studies will be described and discussed together.

6.2.1 Method

The studies were conducted within a university web application, which allowed the students to view coursework deadlines and timetable information, as well as to access coursework grades and to submit requests. To gain access to the system, students had to authenticate with an identifier and a password within the university network. For the purpose of the study, different types of password nudges were displayed on the login interface to encourage stronger passwords. Each of the three studies ran for an academic year between October 2014 and April 2017.

Password strength was measured with the zxcvbn.js password strength estimator [331]. Zxcvbn is available open-source and heuristically estimates password strength, e.g., by checking for common words, patterns, and repetitions in passwords. The tool works completely client-side so that the transmission of unhashed passwords to the server was avoided. Among others, zxcvbn provides a password score between 0 and 4 that indicates whether the number of guesses needed to break the password is less than 10^2 , 10^4 , 10^6 , 10^8 or above. Furthermore, password length was collected as the number of characters included in the password.

The following list details the password nudges and the control condition deployed in the studies. Screenshots of the images displayed in the different conditions are shown in Figure 15.

- *IV0 Control*: The control condition consisted of the standard login interface asking the participants to "Choose a password". The condition was deployed in studies 1 and 2.
- *IV1 Priming*: To encourage people to choose another and more secure password than the frequently chosen term "password", this condition asked people to "Choose a secret". The idea was to prime people by the use of the wording based on the priming effect [134]. The condition was included in study 1.
- *IV2 University Context*: Aimed to activate an expectation effect [258], the participants were presented with a static graphic that compared average and expected password strength in the university context. The condition was used in studies 1 and 2.
- *IV3 School Context*: This condition targeted the identification with the peer group by adapting the static graphic of IV2 to compare average passwords with the stronger passwords of the school of computer science students to which most participants belonged. The idea was to encourage participants to identify with and increase their password strength to that of their peers based on the findings by Brewer [37] and Castano *et al.* [47]. The condition was used in studies 1 and 2.
- *IV4 University Context and Feedback*: In addition to IV2, this condition displayed interactive feedback on password strength in relation to the static graphic as used in password meters (e.g., [84, 313]). The condition was deployed in study 1.
- *IV5 School Context and Feedback*: In addition to IV3, this condition displayed and interactive feedback on password strength in relation to the static graphic as in IV4. The condition was deployed in study 1.
- *IV6 University Context and Reflection*: This condition showed the static graphic used in IV2 along with an item asking people to indicate how strong they perceived their password to be on a 5-point scale ranging from very weak to very strong. The idea was to make people reflect on password strength. The condition was included in study 2.
- *IV7 School Context and Reflection*: This condition showed the static graphic used in IV3 along with the item from IV6 asking people to indicate how strong they perceived their password to be. This time participants were asked as computer science students to facilitate identification with the peer group. The condition was included in study 2.
- *IV8 Social Norm*: Condition IV8 that was used in study 2 aimed to activate social norms by inducing the feeling of being watched. It was implemented with an image of a pair of watching eyes that has been shown to activate norms and increase pro-social behaviour [21].
- *IV9 Hybrid Authentication Nudge*: This intervention used in study 3 consisted of three parts: 1) an image of a long dachshund aiming to invoke the association of "the longer, the stronger", 2) a compensation of the effort associated with stronger passwords provided by increasing the duration a password can be retained before it had to be changed, and 3) a reminder of the current password expiry date. When users created the password, they received dynamic feedback in terms of their password expiry date as a password strength estimation.

Procedure

For an overview, the study procedure is depicted in Figure 15. The figure includes screenshots of the password nudges deployed in the three studies.

When participants started the registration process for the university web application, they were first presented an informed consent form for participation in the study. When students decided not to participate, they could still use the web application.

When agreeing to participate, students were randomly assigned to one of the password nudge conditions or a control condition that simply asked participants to create a password in the first two studies. The respective password nudge was then displayed on the login page where the students created their passwords.

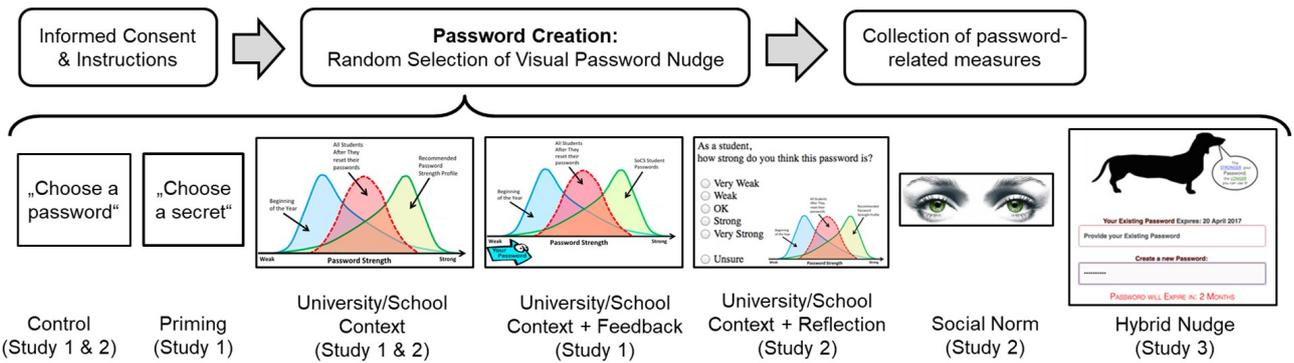


Figure 15: Study procedure of the three field studies exploring the influence of password nudges on password creation as described in [252]. Images included with permission of © 2018 Cambridge University Press.

Based on the requirements of the ethics committee to treat all students equally, and as the nudge in study 3 impacted the password changing policy, all participating students in the third study were assigned to the same password nudge. Thus, no control condition existed in the third study.

Sample

The participants were students who created passwords for their actual university account. The majority was enrolled in technical courses, mainly computer science. Due to the requirements of the university’s ethics committee the collection of further demographic information such as gender or age, and other variables, was not possible. Even though participation in the study was voluntary, the majority decided to participate. In study 1 $N = 497$ of the 587 students that registered for the application participated, study 2 comprised $N = 776$ out of 816 students, and study 3 included $N = 672$ participants out of 918.

6.2.2 Results

For the analysis of the results of studies 1 and 2, password strength and password length of the experimental password nudge conditions were compared to the control condition. As password strength was measured on an ordinal scale and password length not normally distributed, Benjamini-Hochberg corrected [25] Mann-Whitney-U tests were conducted.

The results of study 1 revealed that none of the experimental conditions (IV1, IV2, IV3, IV4, IV5) differed from the control group (IV0), neither in terms of the password strength score nor the password length.

Likewise, comparing password strength and password length of the experimental conditions (IV2, IV3, IV6, IV7, IV8) with the control group (IV0) in study 2 did not reveal any significant differences.

Due to ethical considerations, no control group existed in study 3 that used IV9. Thus, as a substitution, the results of study 3 were compared to the control and experimental conditions of study 2 that had been conducted the previous year. As the sample consisted of natural groups of students, some students used the application for more than a year and took part in both of the studies. This was acknowledged by separating the participants that had previously participated from those that had participated only once. Accordingly, separate analyses were conducted for independent and repeated measures. The results revealed that password strength in study 3 was significantly higher compared to the previous year’s control group and all experimental conditions. The same was true for password length and the independent as well as the repeated measures analyses. The only insignificant difference was between the password length of IV9 and IV7 in the analysis of the independent measures.

6.2.3 Discussion & Implications

The finding that none of the password nudges in study 1 led to stronger or longer passwords than the control condition was unexpected as all password nudges were based on effects previously successful in other contexts.

In the case of failed nudges Sunstein [299, p.4] suggests to "*nudge better (or differently)*" when there is good reason to believe that the user's choice might be biased or based on misconceptions. Indeed, previous studies could show that users' knowledge and perceptions in terms of passwords include misconceptions and deviations from technical measures [275, 314, 315]. Thus, the password nudges per se seemed to be warranted but might need to be redefined to be effective.

Thus, study 2 tested a different set of nudges. These partially included interventions from study 1 to exclude influences that might have affected the time frame of study 1 or the specific sample, and also new interventions to test alternative password nudges. Yet again, no significant differences were detected. This led to a discussion and reflection on the non-significant results of studies 1 and 2, as described by Renaud *et al.* [253]. The reflection included methodological aspects and password-related considerations.

In terms of methodology, password strength was measured using an artificial password score ranging from 0 to 4 that might have considerably reduced the variance of the data as, e.g., all passwords requiring between 10^4 and 10^6 guesses to be broken would be assigned the same score. Next, the use of hashed passwords for privacy reasons later prevented the calculation of different metrics. Furthermore, the required application of a non-parametric procedure might have slightly reduced the test power.

Finally, while the application of the password nudges in real university accounts with actual passwords had certain benefits in terms of the external validity of the findings, it also suffered from some limitations. These included ethical considerations in terms of the collection of demographic and additional variables, but also the lack of possibility to control for external influences or password reuse.

In terms of password-related aspects, it was acknowledged that the participants' password creation context might not have been sufficiently considered, i.e., that authentication is a secondary aim and a complex task influenced by many factors such as time, cognitive effort, knowledge, and experience.

It was furthermore found that at least some of the deployed nudges did not sufficiently enhance understanding of what makes a good password. This would have been especially important considering the deviations of user perceptions from technical measures in previous studies [275, 314, 315]. For example, the priming nudge using the word "secret" (IV1) or the social norm nudge showing a pair of watching eyes (IV8) simply aimed to encourage stronger password choices but did not indicate how to do so or whether the participants achieved that aim. In contrast, the university context nudge (IV4) displaying a graphic and interactive feedback provided the participants with an indication of actual strength. That this might be more helpful is supported by the fact that this condition had the highest descriptive values even though the difference was deemed non-significant after applying a correction for multiple testing.

In line with that reasoning, also Nicholson *et al.* [214] did not find an effect of a simple, visual password nudge on password strength. In a between-subjects study, users' were presented an image of a man in a library carrying a lot of books in conjunction with the instruction to create a password based on that image. The authors did not find a main effect of the image on password creation. In another condition, Nicholson *et al.* [214] changed the wording in the instruction, which positively impacted password creation. When participants were instructed to create a long password, the password strength increased. In contrast to the use of the word "secret" in this research, the word "long" seems to have provided users with a hint of what to do as length is an important predictor for password strength.

Following these considerations and the suggestion of Sunstein, failed but warranted nudges could be handled by enhancing the effects of the nudge through incentives, mandates, or bans [299]. This led to the design of the password nudge for the third study.

It was decided that the study 3 nudge should not only encourage stronger choices, labelled a *simple nudge*, but also provide participants with an indication of password strength, labelled a *hybrid nudge*. Furthermore, the participants should be provided an incentive or rather a compensation for the increased effort associated with creating stronger passwords to acknowledge the complexity of password creation. To directly relate the incentive with password creation, stronger passwords were compensated with an increased password expiry date before it had to be changed, as suggested by Seitz *et al.* [276]. Thus, the hybrid nudge finally consisted of an image of a long dachshund as a simple nudge, a dynamic indication and a reminder of password strength in the form of the current expiry date, and the related compensation to keep strong passwords longer than weak ones.

As no control group existed due to ethical considerations, the results of the previous year's study served as a comparison. Even though not ideal, the comparison provided a strong indication that the hybrid nudge was more successful in encouraging the creation of strong passwords than the previous simple nudges.

Overall, reflecting on the findings of the exploratory field studies on password nudges led to further consideration in several regards:

First, the potential decrease in variance caused by the applied 5-point zxcvbn password score [331] led to a consideration of alternative measures for future password nudge studies.

Second, while beneficial in terms of external validity, the field study setting prevented the collection of additional demographic or control variables to analyse or control for potential influencing factors. Thus, future studies should systematically examine the effectiveness of different nudges in controlled settings before transferring the findings to practical applications.

Third, the insignificant findings in studies 1 and 2 might be explained by the potential lack of feedback and password information in these *simple nudges*. Indeed, the assumed benefit of providing participants with feedback and compensation for strong password creation became visible in study 3, applying a *hybrid nudge*. To analyse the assumed distinction between simple and hybrid nudges and to test differences in terms of their effectiveness, it would be beneficial to more clearly distinguish the concept of nudging from other interventions such as feedback or information provision.

Fourth, beyond a more clear-cut definition of nudges, the design of future nudges should also be based on a better understanding of how nudges exert their influence.

The first and second consideration resulted in adaptations of the study design in that it was decided to use an artificial but controlled setting for future password nudge studies, and to include more fine-grained password scores and measures such as password entropy or the number of guesses required to crack a password.

The third and fourth consideration led to an in-depth examination of the existing nudge literature to better understand *what* a nudge is compared to related interventions, *how* nudges exert their influence, and *when and where* nudges are not only successful but also suitable. Finally, this also led to a detailed engagement with the ethical aspects of nudging. The next sections thus differentiate the nudge concept already described in section 6.1 from related interventions and discuss their ethical implications.

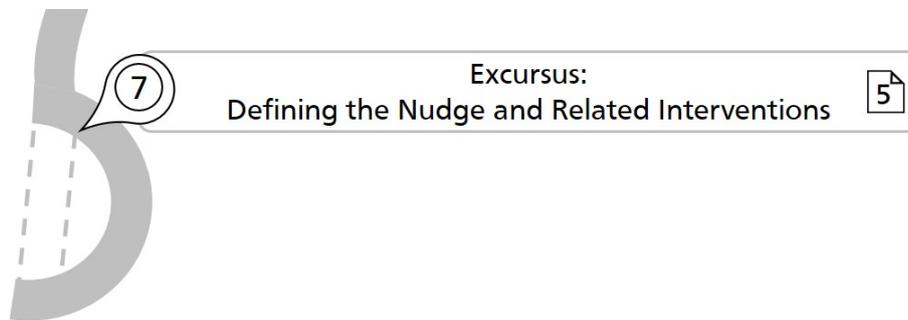


Figure 16: Research step 7.

Manuscript 5: Renaud, K., and Zimmermann, V. Ethical guidelines for nudging in information security & privacy. *International Journal of Human-Computer Studies* 120 (2018), 22–35. doi:10.1016/j.ijhcs.2018.05.011.

Reflecting on the mixed findings of the field studies revealed the importance of having a clear understanding of what constitutes a nudge and how a nudge can be separated from related interventions. For example, one of the nudge conditions studied by Nicholson *et al.* [214] included a financial incentive, even though the original definition by Thaler and Sunstein excludes “significantly” changing the economic incentives for certain options [303]. Likewise, information provision has sometimes been labelled a nudge [44, 59, 146]. Yet, even though not directly excluded from Thaler and Sunstein’s definition, researchers pose the question, what would be new about nudging if mere information provision was classified a nudge [122] and deem such a broad definition problematic in terms of the falsifiability of the nudge concept [223]. Finally, also a discussion about whether or not the proposed ban of 16-ounce drinks in the United States fulfils the definition of preserving all options as people would be able to refill smaller cups [192] demonstrates some ambiguity in the nudge concept.

Based on a review of the nudge literature, this section will thus provide a more clear-cut definition of the nudge concept and related concepts. It will first briefly introduce Dual Process Theories that have often been used to explain how nudges exert their influence [123, 204, 303]. For more detail, the reader is referred to the article “*Ethical guidelines for nudging in information security & privacy*” by Renaud and Zimmermann [249] of which the next section is a brief summary.

7.1 Dual Process Theories and the Concept of Nudging

Psychological Dual Process Theories [88, 156, 222, 285, 290] differ in terms of certain details and the labels provided for different concepts. However, most share the distinction of System 1 and System 2 information processing of the brain. These two types of reasoning can be broadly differentiated as follows:

System 1 can be described as a “*form of universal cognition shared between humans and animals*” and comprises several, partially autonomous subsystems [88, p.454]. The System 1 mode of reasoning is referred to as instinctive, fast, parallel, automatic, and requiring little capacity [88, 290]. With regard to consciousness, it is assumed that System 1 processes are rather implicit [90] and that only the final product of the process is conscious [88]. However, other researchers do not detail the question of consciousness [285]. System 1 is associated with heuristic processing and intuition [290].

System 2 is supposed to be evolutionary recent and specifically attributed to human information processing [88]. This mode of reasoning is often described as slow, sequential, and limited in terms of capacity [88]. System 2 processes are supposed to be explicit and conscious [88]. This type of reasoning allows for analytic processing, explicit learning, and rational decision-making [290].

Researchers conducted studies based on the deductive reasoning paradigm to demonstrate the existence of the two modes of reasoning [87, 89] and also provided indications through neuropsychological studies [109, 110].

However, the nature of the relationship between the two systems is not yet entirely clear. Assumptions include an interaction of the two systems [285], parallel processing [90], and serial processing [108]. Others suggest that the two systems complement one another with System 1 being the default system of which the outputs are monitored by System 2 [91, 155]. Some researchers propose single system frameworks in which the type of reasoning varies along a continuum and in relation to the quality of the cognitive representations [222]. Despite the different assumptions in terms of the two systems' relationship, many researchers agree that there is some kind of connection between System 1 and 2, and that "a fully dissociationist view of System 1 and System 2 is not adequate for capturing the complexities in which decision-making processes operate" [183, p.295].

The nudge concept is related to Dual Process Theories in that these have often been used to explain how nudges exert their influence [123, 204, 303]. Following prominent definitions of nudging that include the activation of automatic cognitive processes [43, 122, 129], there is a relation between nudges and automatic System 1 information processing. Thus, nudges are supposed to encourage a certain choice by purposefully tapping System 1 information processing and by making use of the automatic, quick, and rather implicit processes associated with it. This also implies that the nudges are not necessarily aware of the intervention that takes place. Consider for example the previously mentioned study in which food intake was reduced by varying the proximity of certain food choices and the serving tools at a salad bar [259], or the use of smaller plates [123]. Without adequate information, customers might not be aware of the intervention or its purpose.

Taking into account some connection between System 1 and System 2, some authors differentiate between Type 1 and Type 2 nudges [123, 183]. While Type 1 nudges primarily target System 1 information processing as described above, Type 2 nudges also activate System 2 information processing. Examples include traffic light labelling [183] or the use of framing [123]. According to Lin *et al.* Type 2 nudges aim to encourage a reevaluation by disrupting coherence between the evidence base for decision-making and the actual choice. Hansen and Jespersen's [123] differentiation bears some similarities in that a Type 2 nudge is supposed to attract reflective attention so that both definitions include a reflective, and thus conscious, aspect. However, a Type 2 nudge does not directly activate System 2 but does so indirectly via activating System 1 [123]. This implies that both Type 1 and Type 2 nudges make use of automatic cognitive processes but vary in the degree in which they target System 2 and conscious reflection, respectively.

7.2 Definition of the Nudge Concept

Based on the considerations above and the reviewed literature, the following criteria for an intervention counting as a nudge were derived.

- *Predictability*: Nudges should influence nudges in a predictable way and towards a predicted outcome [303].
- *Automatic cognitive processes*: Nudges activate the automatic cognitive processes used by System 1 information processing such as well-known biases and heuristics [43, 122, 129].
- *Equality of costs*: No choice should be more costly financially or economically, or in terms of time, effort, or social sanction [129, 303].
- *Preservation of choices*: The nudge should not remove or ban any pre-nudge choice [129, 303]. Here, it is important to disentangle the concept of choices (e.g., the possibility of eating as much as one likes staying with the food-intake example from above) from that of options (e.g., smaller plates) [192].
- *Nudge for good*: Nudges should be deployed for the good of the nudgee [303].

7.3 Definition of Related Interventions

Throughout the literature review, some related concepts emerged that are differentiated from the nudge in the following. Table 1 provides a short overview of the aspects in which these interventions differ from the criteria for nudges.

		RELATED INTERVENTIONS			
NUDGE CRITERIA		Code	Sludge	Information Provision	Hybrid Nudge
	Predictability				
	Automatic Processes			X	(X) targets automatic AND reflective cognitive processes
	Equality of Costs	X			
	Preservation of Choices	X			
	Nudge for Good		X		

Table 1: Differences between the criteria for a nudge and related interventions.

7.3.1 Code

According to Calo [43], a code constitutes an intervention designed to make the undesired behaviour more difficult by making changes to the environment. Speed bumps on a road that require drivers to slow down would be an example for a code. Like nudges, a code can activate automatic cognitive processes to reduce the possibility of users choosing the "unwise" option. An equivalent to the physical speed bump example would be speed bumps drawn on a street to create the optical illusion of speed bumps to make drivers slow down [15, 117]. However, there are two main differences to a nudge: First, while codes mainly focus on decreasing the undesired choice, nudges generally focus on facilitating the desired choice. For example, while a code would aim to prevent speeding, a nudge would encourage slow driving by motivating drivers behaving accordingly with a happy face traffic sign [3]. Second, codes also include interventions that make choosing another option unduly difficult, "*illegal or next to impossible*" [43, p.778], e.g., by applying sanctions or physical barriers. This type of intervention falls outside the definition of the nudge as costs for different choices are not equal, and the preservation of choices might be endangered.

7.3.2 Sludge

As originally intended by Thaler and Sunstein nudges should be used "for good" and for encouraging people to make "better" choices as judged by the nudges themselves [303]. Nevertheless, as with many insights or interventions, it is also possible to misuse nudges for personal or economic benefits rather than for the good of the individual nudgee or the general public. An example would be some low-cost airlines that make use of nudges to encourage customers to purchase unnecessary options for the company's financial gain [327]. This type of intervention would be labelled a *sludge* [302] and does not comply with the definition of the nudge. What is different from the definition of the nudge is not the intervention per se, but the purpose the concept is used for.

7.3.3 Information Provision

Information provision, also labelled a notice by Calo [43], is aimed at supporting better decision-making by providing users with the relevant facts. The idea is to bridge a potential knowledge gap between some party (e.g., a service provider or organisation) on the one side and the user on the other side [43].

Information provision can take the form of various educational elements, such as reports on certain topics, explanations, notifications, warnings, or feedback. Examples include terms of use or warnings on dangerous products [43]. Information provision is widely used as a regulatory intervention, yet criticised for seldom working in practice [43]. Potential reasons for the failure of information provision include an overload of information [24, 132], high complexity of the provided information, and difficulties with making sense of the information for decision-making [24].

The main difference from a nudge is the cognitive process the intervention targets. While nudges target automatic cognitive processes in System 1 (or indirectly System 2 via System 1 in case of Type 2 nudges), information provision primarily and directly targets active cognitive involvement and reflective reasoning, i.e., System 2 processing.

7.3.4 Hybrid Nudge

The combination of a nudge with additional educational elements was termed a *hybrid nudge* in the field study described in section 6.2.

The exemplary hybrid nudge described in the study comprised two simple nudge elements: 1) an image of a long dachshund to automatically activate the association of "the longer, the stronger", and 2) the compensation of a stronger password with a later password expiry date. The latter intervention was supposed to nudge users towards stronger passwords not by incentivizing, but by compensating the increased effort for secure password creation with a password-related feature rather than with financial or other economic incentives. By so doing, the choice of weak versus strong passwords should become more balanced rather than imbalanced by the compensation. Yet, it could be argued that this intervention is positioned at the limits of the simple nudge definition that includes no "significant" change of the economic incentives [303].

The hybrid nudge furthermore included two educational elements: 1) feedback on password strength supposed to educate users on their current password strength, and 2) a reminder of the password expiry date coupled to the password's strength.

The main difference to a nudge is that educational elements enhance the hybrid nudge. It is thus a combination of two types of interventions: the nudge and information provision. As such, it is supposed to activate automatic cognitive processes (System 1) to encourage a certain choice, but also to target reflective reasoning (System 2) to enhance the users' awareness and understanding of the intervention.

7.4 Implications

This section differentiated the concept of nudging from related interventions. It has been shown that the code, sludge, and information provision fall outside the definition of the nudge. The combination of a nudge and information provision has been labelled a hybrid nudge. The field studies provided indications for the effectiveness of this type of joint intervention.

Thus, apart from the nudge as a main focus of this research, the hybrid nudge will further be discussed and analysed in the following. Likewise, information provision as a supplement to the nudge in the combined intervention hybrid nudge will be considered in the next sections even though outside the nudge definition. This is relevant for analysing the single versus joint effects of the nudge interventions across the dual processing model, as suggested by Dolan *et al.* [80].

The code and sludge, however, will not be considered further, as they may limit the choice set and are not implemented "for good", respectively. They are thus not in line with the aim of this research to design user-centred interventions that support users without restricting choices or enforcing strict rules.

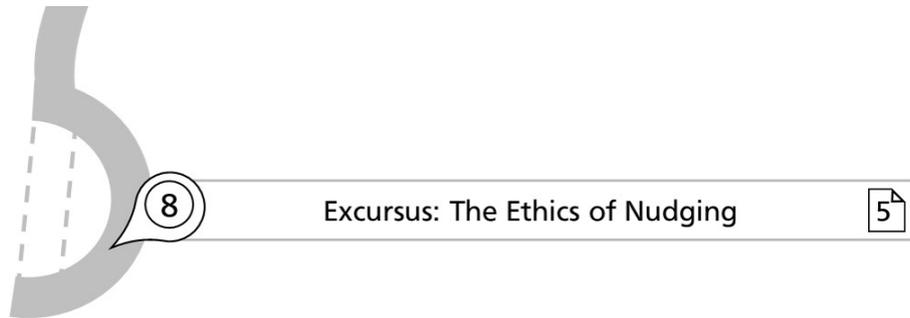


Figure 17: Research step 8.

Manuscript 5: Renaud, K., & Zimmermann, V. (2018). Ethical guidelines for nudging in information security & privacy. *International Journal of Human-Computer Studies* 120, 22-35, <https://doi.org/10.1016/j.ijhcs.2018.05.011>.

Despite its spread and numerous applications, the concept of nudging has not enjoyed unanimous success. This section briefly summarizes the arguments of the supporters and critics of the nudge concept before discussing the ethical implications for the nudge, the hybrid nudge, and information provision as differentiated in the previous section. The ethical considerations led to the derivation of guidelines for ethical nudging based on established guidelines for ethical psychological research that are summarized at the end of the section. The process is depicted in Figure 18.

Please note that the discussion surrounding nudging and ethics is described in more detail in the article "*Ethical guidelines for nudging in information security & privacy*" by Renaud and Zimmermann [249] of which this section is a brief summary.

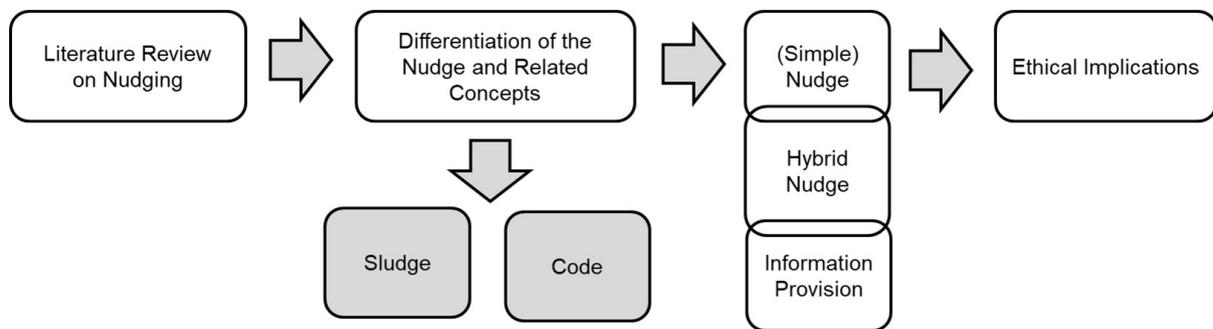


Figure 18: Graphical depiction of the process to differentiate the nudge from related interventions before analysing ethical implications for the selected intervention types as described in [249].

8.1 Arguments for Nudging

The arguments of the supporters of nudging can broadly be classified into four aspects.

The first argument is that there is no such thing as a neutral choice architecture and that nudging is thus inevitable [2, 39, 298]. Any design decision, such as colouring, positioning, or wording, will influence the nudgee. Instead of nudging unawares and with unintended and possible negative side effects, supporters of nudging argue for designing the choice architecture deliberately and ethically.

The second argument extends the first by not only acknowledging that nudges are inevitable but also beneficial [79]. Not actively and ethically deploying nudges might open the path for the deployment of nudges for malign purposes [298]. Nudging "for good" instead can counteract unethical nudges [298] and support civic behaviour as a government's responsibility [151].

Third, advocates for nudges argue that humans are confronted with a plethora of often difficult choices with many options [124]. Thus, nudges that support the nudgees by facilitating choice may well be considered helpful [32, 57] rather than compromising autonomy.

Autonomy is also the topic of the fourth argument. Autonomy is referred to as "*the quality or state of being self-governing*" in general and "*self-directing freedom and especially moral independence*" on the level of personal autonomy [77]. Some supporters of nudging believe that nudging does not infringe human autonomy as the original choice set is preserved and nudgees are free to ignore the influence of the nudge and to choose otherwise [298].

8.2 Arguments against Nudging

The main arguments against nudging can be clustered into four groups.

First, to pick up on the last argument of the nudge advocates, its critics question that autonomy is preserved when nudges are applied [164]. For example, Mitchell [206] states that "libertarian paternalism", a term intertwined with nudging and used to describe the idea of authorities influencing towards the "right" choice without constraining or coercing [297], is an oxymoron. He explains that freedom of choice, and thus perhaps also autonomy, cannot be assumed in contexts in which decision-making is known to be influenced by biases and heuristics [206]. Other researchers suggest that nudges can be manipulative if they lead to choices people would not have made without the nudge [336]. Furthermore, the removal of the need to think about a choice might lead to excessive convenience [271].

Second, the critics of nudging argue against the "covert" influence of nudges that primarily target automatic and perhaps subconscious cognitive processes [123]. According to them, the transparency of nudges is a prerequisite for their ethical deployment [216]. Indeed, also Thaler and Sunstein agree that the disclosure of the nudge and the defence of its "goodness" is warranted for an ethical deployment of nudges [303].

A third argument against nudging concerns the power of the choice architect. Researchers are concerned that the choice architects, like the nudgees, succumb to automatic cognitive processes so that their decision-making might be flawed [43]. Thus, the qualification of choice architects and the trust in their ability to know "better" is questioned [212], especially given changing opinions and uncertainty on what is deemed the "best" choice.

Fourth, mismatched nudges that do not align with the targeted group, context, or choice can produce unanticipated and perhaps adverse side effects. For example, the use of pictorial warnings on cigarette packs to reduce smoking increased anxiety and craving in heavy smokers [185].

8.3 Guidelines for Ethical Nudging

Ethical principles should guide research as well as the practical application of research. Well known sets of principles for ethical psychological research include that of the British Psychological Society (BPS) [305], the American Psychological Association (APA) [12], the European Federation of Psychologists' Association (EFPA) [86], or the Belmont report [74].

A comparison of the principles revealed some similarities so that the main principles can be summarized as follows:

- *Respect for Persons*: People should be respected regardless of cultural and individual differences such as gender, race, religion, or disabilities. No person should be treated in an unfair, prejudiced or discriminatory way. Autonomy and self-determination of the person should be maximised while recognising potential limits. Personal data should be treated confidentially and anonymously.
- *Beneficence*: Research should be aimed at contributing to the common good. All people involved in the research should be protected from potential harm or risks.
- *Justice*: Research should be just in that all people should equally be entitled to access and benefit from it.
- *Scientific Integrity*: Researchers should comply with ethical and scientific standards to ensure the quality and contribution of the research.

-
- *Social Responsibility:* The social responsibility of research as well as the expected, and perhaps unexpected, outcomes and consequences of research should be considered.

Comparing these principles with the arguments against nudging, it can be seen that some of the objections raised by the nudge critics do not align with the principles for ethical research.

For example, the criticism that nudges impair personal autonomy affects the principle to respect people and to maximise their autonomy and self-determination. The assumed impact on autonomy is also connected to the second criticism that nudges target cognitive automatic processes, and by so doing influence "covertly". An increase in transparency, as called for by some researchers [123, 216, 272, 303], would be in favour of the respect principle and could be viewed as an important step towards autonomy. If users are aware of the nudge and its purpose, it might be easier for them to ignore the nudge's influence. An important distinction in this regard again is the concept of Type 1 and Type 2 nudges as described by Hansen and Jespersen [123]. From their point of view, the use of transparent Type 2 nudges that not only target automatic cognitive processes but also indirectly activate reflective reasoning are most favourable in terms of ethics.

The third argument of nudge critics concerned the power and role of the choice architect. The actions and decisions of the choice architect can, in the worst case, negatively affect a number of principles. For example, the beneficence principle can be compromised if the nudge is not intended for the good of the nudgee, the social responsibility principle is affected if the nudge produces adverse and unintended side effects, and scientific integrity can be endangered if the use of the nudge cannot be justified.

Finally, the last argument against nudging also concerns scientific integrity in that the quality of research is endangered if the employed nudge does not fit its intended purpose. Furthermore, potential unintended side effects may not only affect the social responsibility principle but also justice if the effects concern certain groups of people in particular such as in the example where anxiety and craving was especially increased in heavy as compared to light smokers [185].

These considerations lead to several implications and guidelines for the ethical deployment of nudges:

- *Suitability:* In light of the concerns for autonomy and regarding the principle respect, researchers should only use a nudge if warranted. In some contexts, other forms of intervention with no implications for personal autonomy, such as information provision, might be more suitable.
- *Consideration of Ethics:* When using a nudge, researchers, i.e., the choice architects in that case, should generally reflect on its ethical implications. They can, for example, discuss implementation plans with other researchers or ethical review boards to ensure the intervention respects people and is scientifically sound.
- *Justification:* Researchers should be able to justify their decision to nudge and the selected type of nudge to ensure scientific integrity.
- *Nudge for Good:* Even though already a criterion for nudging as such [303], nudging "for good" should also be the intended purpose for the ethical deployment of nudges. This guideline is supposed to ensure the beneficence and respect principles.
- *Preservation of equally costly Choices:* To comply with the definition of nudges and with regard to autonomy, no choices should be banned or significantly altered in terms of financial or other economic incentives.
- *Transparency:* The nudge itself should be transparent so that the nudgees are aware of its existence and the purpose of the intervention [123, 216]. By increasing transparency the potential "covert" influence of the nudge should be counteracted. If transparency is not possible for a good reason, people should at least be debriefed.
- *Matching:* In order to ensure scientific integrity and quality of research, the nudge should be designed to match the targeted group, context, and choice.
- *Avoid Unintended Side Effects:* The consideration of the nudge's effects is supposed to reduce unsuccessful nudges and unintended side effects. This implication serves the social responsibility and justice principles and aligns with the nudge's definition to nudge in a predictable way and towards a predicted outcome.

8.4 Implications

Regarding the interventions selected for further analysis, the nudge, information provision, and the hybrid nudge, the ethical considerations have the following implications:

The discussion summarized in this section and the guidelines derived from it primarily concern the nudge. Apart from following the nudge definition, the guidelines derived from the discussion suggest that future research on nudging should carefully consider the suitability of the nudge and its ethical implications. The type of nudge should be matched with the respective decision, and the effects of nudging should be considered. Furthermore, the design of the nudge should be transparent in terms of the presence of the nudge and its intended purpose. In this regard, Type 2 nudges that indirectly activate reflective reasoning in System 2 via System 1 seem to be ethically favourable over Type 1 nudges [123].

The use of information provision in research and practice does not seem to be accompanied by ethical concerns as long as no sanctions or coercion are applied [249]. The literature analysing the reasons for unsuccessful information provision [24, 132] suggests that information should be designed in a way that it is not overwhelming, of little complexity, and clear in terms of the implications for the decision.

As information provision directly targets reflective reasoning in System 2, it is supposed that the combination of a nudge and information provision in a hybrid nudge further contributes to the transparency of the nudge. The educational elements of information provision could provide users with information beyond the existence and purpose of the nudge. They could, for example, be used to educate users *why* a certain choice is deemed "better" or *how* people can make "better" decisions in the long term when the nudge may not be present. In the context of password creation, information provision could be used to inform users of password strength but also to educate users on how to improve password strength or how to manage multiple, strong passwords. The combination of the power of the nudge targeting automatic cognitive processes with information provision that has sometimes been shown to be ineffective on its own [219] might further lead to synergistic effects as indicated by the results of the hybrid nudge in the field studies [252].

However, other researchers raise the question whether the power of the nudge that is drawn from the activation of automatic cognitive processes might be decreased by increasing its transparency and thus the degree of reflection [192]. It may well be that the first trialled hybrid nudge used in the field study (see section 6.2) was only effective because of the certain context, sample, or combination of elements.

Thus, after this "detour" to better understand the concept of nudging and its ethical implications, the next section details two studies designed to explore the single and joint effects of nudging and information provision in hybrid nudges, first, in different cybersecurity-related decision contexts, and second, in terms of password creation in particular.

9 The Influence of Nudge Interventions on Security Decisions and Password Creation

The following sections summarise two consecutive studies on the effectiveness of nudging in cybersecurity with a focus on password creation. The first study analyses the single versus combined effect of nudging and information provision on different types of cybersecurity decisions, including password creation. The second study specifically applies the findings of the first study to password creation and compares differences in password nudges based on different heuristics, biases, and norms in detail.

9.1 Comparison of Different Nudge Interventions

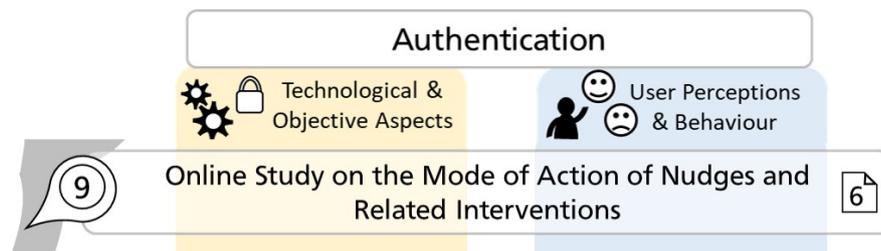


Figure 19: Research step 9.

Manuscript 6: Zimmermann, V., and Renaud, K. The Nudge Puzzle: Matching Nudge Interventions to Cybersecurity Decisions. *ACM Transactions on Computer-Human Interaction* 28, 1 (2021), 7:1-7:45. doi:10.1145/3429888.

This section is a brief summary of Manuscript 6 "The Nudge Puzzle: Matching Nudge Interventions to Cybersecurity Decisions" by Zimmermann and Renaud [354]. The focus of the summary is on password creation. For further details and findings, the reader is referred to the manuscript.

9.1.1 Background

From the discussion around the ethics of nudges primarily targeting automatic System 1 information processing (see section 8) it was concluded that nudges should be transparent for the user. Transparency in that sense can concern two aspects: (a) a nudge intervention can be visible so that the user is aware of the intervention and its purpose, and (b) beyond that, a nudge can be transparent in terms of its reasons so that the user can understand why they are being nudged towards a certain option.

In terms of the first aspect, the nudge itself can be more or less transparent following the Type 1 and Type 2 nudge differentiation by Hansen and Jespersen [123]. For example, a positioning nudge changing the sequence of options is invisible to the user and probably goes unnoticed. Instead, transparent footprints leading to a garbage bin [123], or an emoticon posted next to a certain option, is visible. However, in both cases, it might not become clear for the user *why* they are being nudged towards a certain option.

One way to achieve this aim, and address the second aspect might be an intervention clearly targeting System 2 information processing, a prime example of which is information provision. It is assumed that the nudge's transparency, and thereby its acceptance, can be increased by combining it with information provision. That way, the nudge is supposed to encourage a secure choice while the information educates users about why it is deemed the favourable option. Education might also support users in making better related or subsequent decisions in which the intervention is not present. Furthermore, being informed about the reasons for the choice, it might also be easier for users to resist the nudge should they decide to choose another option. Yet, it remains unclear whether the inclusion of information provision would diminish the nudge's influence [192]. Therefore, this study analysed the single and combined impact of a nudge and information provision as the first aspect of interest. For an easier differentiation from the hybrid nudge, a nudge was labelled a *simple nudge* in this study.

The hypotheses guiding the study were formulated as follows:

Hypothesis H₁: Hybrid nudges, i.e., the combination of a simple nudge and information provision, are most effective in encouraging secure choices, as compared to no intervention, a simple nudge, or information provision on their own.

Hypothesis H₂: Information provision and hybrid nudges are more effective in helping people to choose the secure option in subsequent decisions where no nudge intervention is present.

As a second aspect, the single versus the combined impact of simple nudges and information provision was analysed across four different types of decisions. One of the decisions was password creation as the focus of this research. The type of decision was differentiated in terms of its frequency (high vs low) and complexity (high vs low).

9.1.2 Method

In an online study conducted via Amazon Mechanical Turk $N = 450$ people were asked to take four cybersecurity-related decisions that included different types of nudge-related interventions. In a follow-up study two weeks later, the participants took the same security decisions in the control condition variant to check the long-term effects of the interventions on decision-making when the nudge intervention was not present.

Study Design

The study design included four nudge-related intervention types and four types of decisions.

The types of interventions were a simple nudge, information provision, and a hybrid nudge following the definitions provided in sections 7.2 and 7.3. The fourth condition was a control condition in which no nudge-related intervention took place.

The four security decisions included the choice of a public WiFi as an example for a simple and frequent decision, the choice to encrypt one's smartphone storage as a simple and infrequent decision, and the choice of a cloud service provider with numerous features as a complex and infrequent decision. Secure password creation was classified as a frequent and complex decision, with numerous password options that among others are influenced by the users' knowledge, experience, time, or the effort associated with memorizing and repeatedly typing the password.

As the focus of this chapter is on password creation, the following example details the different nudge-related interventions in the password creation context. All password creation interventions are based on a password meter developed by Ur *et al.* [313] and the respective code made available open-source. The password meter was slightly adapted for this study.

- *Control*: In the control condition, a plain password entry field was shown along with the instruction to create a password that users had not previously used.
- *Simple Nudge*: In the simple nudge condition, a coloured bar filling with increasing password strength was added to the password field shown in the control condition. The feedback bar aimed to activate the learned connection red-bad/insecure and green-good/secure.
- *Information Provision*: Next to the password entry field, people were provided dynamic password guidance that changed with the user's input. It was based on the password meter designed by Ur *et al.* [313] and the NIST password recommendations [115].
- *Hybrid Nudge*: The hybrid nudge condition combined the simple nudge and information condition. Thus, next to the password entry field, the participants were provided with a coloured feedback bar filling with increasing password strength as well as dynamic password guidance. Figure 20 shows an exemplary screenshot of the hybrid nudge condition, including the simple nudge and the dynamic information.

Manuscript 6 only details these four different conditions, yet, for this research, two additional conditions were analysed. Aiming to analyse whether participants were influenced by the information being static or dynamic, the following variants were added:

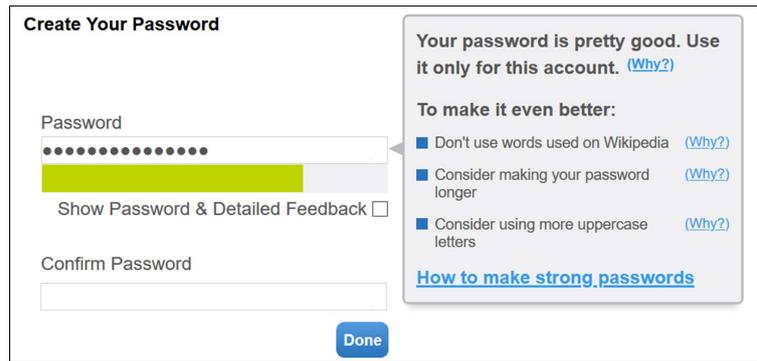


Figure 20: Screenshot of the password creation hybrid nudge condition with dynamic information based on Ur *et al.* [313].

- *Static Information:* Next to the password entry field people were provided with a static text field with password guidance. It was based on the password meter designed by Ur *et al.* [313] and the NIST password recommendations [115]. Variations of the text were furthermore evaluated in a bachelor's thesis conducted by Franziska Koch [170] and co-supervised by Verena Zimmermann.
- *Static Hybrid Nudge:* The static hybrid nudge condition used the same static text as described in the static information condition, combined with the coloured feedback bar described above.

The other three security decisions were designed with the following intervention types:

- *WiFi Choice:* Based on a study by Turland *et al.* [310] participants were asked to select a public WiFi from a list of WiFis with similar names. These were either sorted by signal strength (control condition) or by security (simple nudge). The simple nudge thus was a positioning nudge. The information provision condition consisted of coloured security indicators next to the WiFi name, and the hybrid nudge combined the security indicators with the sorting according to security. In all conditions, users could click on the WiFi name to view the security details.
- *Phone Encryption:* Participants were asked to imagine configuring a new phone which among others included the decision to encrypt the phone. In the control condition users were asked to either click "yes" or "no". The simple nudge condition used a default nudge with "yes" being pre-selected, and the information condition consisted of a short text on encryption and its benefits. The hybrid nudge combined the information text with the default nudge.
- *Cloud Service Choice:* In the control condition, participants were asked to choose one of three fictional cloud service providers based on a description. Each service performed best on one of the multiple described features while all others were held constant. The service provider descriptions were either accompanied by a "most popular" banner above the most secure option as a popularity nudge (simple nudge), an information table highlighting the differences between services (information provision), or both the banner and the information table (hybrid nudge).

Table 2 provides an overview over all conditions and interventions analysed in the study. The decision types were varied within participants and presented in randomized order. For each decision type, users were assigned to one type of intervention, so that they took each decision only once. The sequence of the decision types and intervention types was drawn without replacement to balance sequential effects and to avoid bias on either the security decisions or the kind of intervention. For example, a participant might have been assigned to the following sequence of conditions: 1) WiFi Choice - Simple Nudge, 2) Smartphone Encryption - Control, 3) Password Creation - Hybrid Nudge, and 4) Cloud Service Choice - Information Provision.

Procedure

After agreeing to the informed consent form, participants were asked to imagine being in the situation described on the following pages and to take the related decisions, e.g., having to create a new password for a new account. After completing all four decision scenarios with one type of intervention each, participants were asked to describe the reasons for the choices they made in an open text field.

	INFREQUENT				FREQUENT			
COMPLEX	Choice of Cloud Service				Password Creation			
	Control	Simple Nudge	Information	Hybrid Nudge	Control	Simple Nudge	Information	Hybrid Nudge
	Textual description of services	„Most popular“ banner above secure option	Summary table of differences	Combination of Information and Simple Nudge	Generation of password not previously used	Bar that changes colour and fills with increasing strength	Static/Dynamic Information on what makes a strong password	Combination of Static/Dynamic Information and Simple Nudge
SIMPLE	Encryption of Smart Phone				Choice of Public WiFi			
	Control	Simple Nudge	Information	Hybrid Nudge	Control	Simple Nudge	Information	Hybrid Nudge
	Yes/No Decision	Default option “Yes”	Yes/No Decision + information about encryption	Combination of Information and Simple Nudge	Choice of a network sorted by strength of connection	Choice of a network sorted by security of connection	Choice of a network marked as secure or insecure	Combination of Information and Simple Nudges

Table 2: Overview of the conditions and interventions tested in the study to analyse the single and joint effect of simple nudges and information provision. *Note:* This is an adapted version of a table also included in [354].

These answers were used to analyse whether people were aware of any intervention and to explore additional factors that were deemed important for making the decision.

Afterwards, participants were asked to provide basic demographic information, their technological affinity as well as their security knowledge and attitude using the Affinity for Technology Interaction (ATI) scale [98], the Security Behavior Intentions Scale (SeBIS) [83] and a slightly adapted version of the Human Aspects of Information Security Questionnaire (HAIS-Q) [228]. At the end of the study, participants were asked to provide their created password again to analyse memorability. The study procedure is graphically depicted in Figure 21.

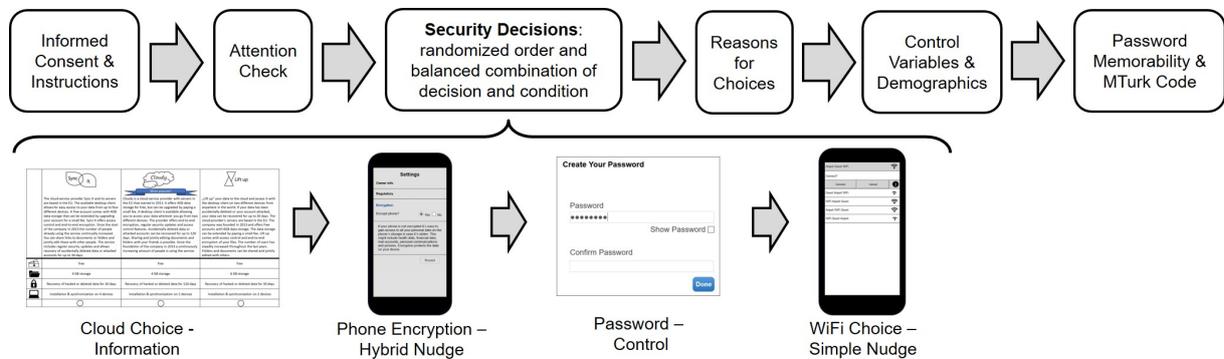


Figure 21: Study procedure of the online study analysing the single and joint effects of nudging and information provision in hybrid nudges. *Note:* Figure also included [354].

Two weeks later, the participants who had previously participated were asked to take part in a follow-up study. The follow-up study again started with an informed consent sheet and a prompt to provide the password created in the main study. Then, participants took the same security decisions as in the main study in a randomized order, including the creation of a new password. The difference was that all participants were assigned to the control conditions of all four conditions. Thus, no nudge intervention or information was provided. After completing the security decisions, participants were again asked to provide reasons for their choices and the newly created password.

After the follow-up study, all participants, including those that had not participated in the follow-up study, received a message with detailed information on nudges in general and the nudges employed in the study.

Sample

The sample of the main study consisted of $N = 450$ people residing in the USA. Of these, 264 identified as male, 180 as female, and six identified as other or did not provide an answer. The age distribution was as follows: 36.7% 18-29 years, 36% 30-39 years, 14.4% 40-49 years, 8.4% 50-59 years, and 3.8% over 60 years. Most participants were employees of various kinds (69.1%), self-employed (16.2%) or reported an IT-related occupation (8.9%). A subset of 330 of the participants again took part in the follow-up study.

The participants were recruited via Amazon Mechanical Turk and compensated with \$2.50 for the main study and \$2 for the follow-up based on a 10\$ per hour rate. The payment exceeded the minimum wage suggestions of most states in the USA.

9.1.3 Results

This section will first describe the results in terms of password creation in detail before briefly summarizing the findings with regard to the other security decisions.

To test hypothesis H_1 in terms of the effectiveness of hybrid nudges, for each security decision the distributions of "secure" and "insecure" decisions in the simple nudge, information and hybrid nudge conditions was compared to that of the control condition without any intervention. To analyse hypothesis H_2 in terms of the educational effects of hybrid nudges and information provision the distributions of "secure" and "insecure" choices in the follow-up study, where all participants were assigned to the control condition, were compared to the results of the respective interventions the participants were assigned to in the main study. In addition, the qualitative data provided as reasons for the participants' choices was categorized using an inductive, open coding approach [199].

Please note that in contrast to the other manuscripts, concrete test results are displayed for the password creation decision in a table format as the related Manuscript 6 only contains a selection of the different conditions analysed in terms of password creation. However, the results in terms of the three other decisions are only summarized briefly as the manuscript already contains all results and as these decisions are not the focus of the synopsis.

Password Creation

As the password strength score by Ur *et al.* [313] and password entropy were measured on an ordinal scale, and password length was not normally distributed, non-parametric tests were conducted.

Participants who stated to have reused a password or used a password manager were excluded from the analysis to avoid bias from results that were not influenced by the interventions. Afterwards, the groups consisted of the following number of participants: control $n = 101$, simple nudge $n = 105$, information static $n = 36$, information dynamic $n = 48$, hybrid nudge static $n = 47$, and hybrid nudge dynamic $n = 54$.

Comparing the results across the six conditions with Kruskal-Wallis tests revealed significant differences in terms of password strength ($H(5) = 42.52$, $p < .001$), length ($H(5) = 39.26$, $p < .001$), and entropy ($H(5) = 45.73$, $p < .001$).

Follow-up Mann-Whitney-U tests revealed significantly higher values in all experimental conditions as compared to the control condition. Furthermore, the dynamic hybrid nudge was more effective in encouraging strong, long and high-entropy passwords as compared to the simple nudge condition and the static information condition. However, the differences between dynamic information provision and the dynamic hybrid nudge were not significant. Likewise, the static and dynamic hybrid nudge did not differ significantly. The test values for the analyses of password strength, password entropy, and password length are shown in Tables 3, 4, and 5, respectively.

After again excluding participants reusing a password or using a password manager in the follow-up study, the sample size in each condition was as follows: control $n = 71$, simple nudge $n = 72$, information static $n = 35$, information dynamic $n = 37$, hybrid nudge static $n = 36$, and hybrid nudge dynamic $n = 44$. A comparison of the results of the follow-up study sorted by the conditions participants were assigned to in the main study using a Kruskal-Wallis test did not reveal any significant differences between conditions.

	<i>PW Strength Main Study</i>								<i>PW Strength Follow-Up Study</i>		
	<i>M</i>	<i>SD</i>	<i>Md</i>	<i>Comparison</i>	<i>Z</i>	<i>df</i>	<i>p</i>	<i>r</i>	<i>M</i>	<i>SD</i>	<i>Md</i>
Control	31.55	25.80	31.69	-	-	-	-	-	36.50	23.36	38.20
Simple Nudge	41.06	23.28	39.33	Control	-2.80	1	.003	.20	37.29	23.60	38.06
Information _s	40.60	22.66	40.42	Control	-2.17	1	.030	.19	36.32	27.49	36.03
Information _d	50.38	25.42	50.96	Control	-3.94	1	<.001	.32	39.38	28.34	39.33
Hybrid Nudge _s	52.89	28.11	53.97	Control	-3.97	1	<.001	.33	35.77	26.83	34.68
Hybrid Nudge _d	58.27	27.72	65.25	Control	-5.29	1	<.001	.44	36.63	25.79	32.35
				Simple Nudge	-3.71	1	<.001	.31			
				Information _s	-3.01	1	.003	.34			
				Information _d	-1.58	1	.058	.17			
				Hybrid Nudge _s	-1.06	1	.288	.11			

Table 3: Descriptive values and Mann-Whitney-U test results of the password strength values, M = Mean, SD = Standard deviation, Md = median, Z = standardized test statistic, df = degrees of freedom, p = level of significance, r = Effect size. *Note:* This is an adapted version of a table also included in [354].

	<i>PW Entropy Main Study</i>								<i>PW Entropy Follow-Up Study</i>		
	<i>M</i>	<i>SD</i>	<i>Md</i>	<i>Comparison</i>	<i>Z</i>	<i>df</i>	<i>p</i>	<i>r</i>	<i>M</i>	<i>SD</i>	<i>Md</i>
Control	54.74	22.82	52.31	-	-	-	-	-	58.42	19.96	56.87
Simple Nudge	67.71	23.88	65.39	Control	-4.13	1	<.001	.29	62.72	20.68	56.87
Information stat	63.17	20.52	58.85	Control	-2.22	1	.027	.19	58.41	21.84	53.59
Information dyn	75.54	29.15	66.73	Control	-4.51	1	<.001	.37	61.87	22.97	56.87
Hybrid Nudge _s	73.75	31.90	71.93	Control	-4.19	1	<.001	.34	59.95	21.84	56.64
Hybrid Nudge _d	78.20	26.33	72.35	Control	-5.44	1	<.001	.45	60.67	23.90	57.86
				Simple Nudge	-2.59	1	.005	.21			
				Information _s	-3.04	1	.002	.34			
				Information _d	-1.27	1	.102	.13			
				Hybrid Nudge _s	-.98	1	.327	.10			

Table 4: Descriptive values and Mann-Whitney-U test results of the password entropy values, M = Mean, SD = Standard deviation, Md = Median, Z = standardized test statistic, df = degrees of freedom, p = level of significance, r = Effect size. *Note:* This is an adapted version of a table also included in [354].

There were neither differences in terms of password strength ($H(5) = .637, p = .986$), nor in terms of entropy ($H(5) = .981, p = .964$), or length ($H(5) = .961, p = .966$). Tables 3, 4, and 5 show the descriptive password strength, entropy, and length values in the main and the follow-up study.

The qualitative analysis of the reasons provided for password creation revealed that security and memorability, often mentioned in combination, were the factors most often considered for password creation. A total of 18 participants mentioned either the provided information or the colour-coded strength bar in relation to their password creation. Yet, no obvious differences across the conditions were discernable.

In terms of memorability, the majority of participants (94.89%) was able to reproduce their created password at the end of the main study. At the beginning of the follow-up study about two weeks later the memorability rate was reduced to 23.71%.

	<i>PW Length Main Study</i>								<i>PW Length Follow-Up Study</i>		
	<i>M</i>	<i>SD</i>	<i>Md</i>	<i>Comparison</i>	<i>Z</i>	<i>df</i>	<i>p</i>	<i>r</i>	<i>M</i>	<i>SD</i>	<i>Md</i>
Control	9.82	3.16	9.00	-	-	-	-	-	10.11	2.93	10.00
Simple Nudge	11.18	3.39	11.00	Control	-3.24	1	.001	.23	10.74	3.07	10.00
Information _s	11.00	2.88	10.00	Control	-2.42	1	.016	.21	10.09	3.41	9.50
Information _d	12.56	4.61	12.00	Control	-4.01	1	<.001	.33	10.62	3.40	10.00
Hybrid Nudge _s	12.45	4.35	12.00	Control	-4.09	1	<.001	.34	13.44	4.26	12
Hybrid Nudge _d	12.89	3.79	12.00	Control	-5.07	1	<.001	.42	10.45	3.40	9.50
				Simple Nudge	-2.93	1	.002	.24			
				Information _s	-2.54	1	.011	.28			
				Information _d	-1.06	1	.145	.11			
				Hybrid Nudge _s	-.69	1	.489	.07			

Table 5: Descriptive values and Mann-Whitney-U test results of the password length values, M = Mean, SD = Standard deviation, Md = Median, Z = standardized test statistic, df = degrees of freedom, p = level of significance, r = Effect size. *Note:* This is an adapted version of a table also included in [354].

WiFi Choice

A Chi^2 test revealed significant differences in the frequency distributions of "secure" and "insecure" choices across all conditions. Follow-up one-sided Chi^2 goodness-of-fit tests revealed that the proportion of secure choices was significantly higher in the simple nudge, information provision, and hybrid nudge condition as compared to the control condition. Furthermore, the hybrid nudge was significantly more effective in encouraging secure WiFi choice options than the simple nudge or information provision on their own.

However, after sorting the follow-up study participants into the intervention groups, they belonged to in the main study, no significant differences across groups were revealed. The exploratory analysis of the participants' reasons for their choices showed that participants in the information and hybrid nudge condition more often referred to privacy and security as a reason for their choice. Many people chose the option on top of the list. However, the position nudge, i.e., the intervention to place the secure option on top of the list, was not directly mentioned by the participants as such.

Phone Encryption

Across all conditions, the number of participants choosing to encrypt and rejecting to encrypt varied significantly. Consecutive one-sided Chi^2 goodness-of-fit tests revealed that all experimental conditions were more effective in encouraging participants to choose "encrypt" compared to the control condition. The hybrid nudge condition was most effective overall and encouraged significantly more "encrypt" choices than the information condition and the simple nudge condition. An analysis of the follow-up study results showed no significant differences between participants that belonged to different conditions in the main study. The qualitative responses provided no indications for large differences between conditions. The default nudge, i.e., the pre-selected "yes"-option was never explicitly mentioned.

Cloud Service Choice

Of the cloud service providers one performed best in terms of security, one provided most data storage, and one allowed for most installations. Comparing the participants' cloud services choices across all conditions with a Chi^2 test revealed significant differences. Like in the other decisions, follow-up tests showed that in all experimental conditions, the secure option was chosen significantly more often. The hybrid nudge condition was also more effective than the information provision condition, but not the simple nudge condition. The comparison of the frequency distributions of the follow-up study, sorted by the interventions the participants were assigned to in the main study, did reveal significant differences. However, follow-up tests did not reveal a clear pattern as each of the services was chosen most often in one of the conditions.

Again, security was more often mentioned as a reason for participants' choices in the information and hybrid nudge condition as compared to the other two. Also, the popularity nudge was explicitly mentioned by a third of the participants that saw the nudge.

9.1.4 Discussion & Implications

This section first interprets the results with regard to the two hypotheses before reflecting on the implications for password creation. The findings and implications should be interpreted in light of this study being conducted in an artificial context via Amazon Mechanical Turk and with a sample that might differ from the general public, for example in terms of age and technological affinity.

Discussion of H_1 : Single and Combined Effects of Nudging and Information Provision

Regarding H_1 , the results revealed that the hybrid nudge condition across all four security decisions was more effective in encouraging secure choices as compared to the control condition. Furthermore, in most cases, the hybrid nudge as a combination of a simple nudge and information provision, was also more effective than either intervention on its own. Exceptions were found in the password creation decision and the cloud service choice. Nevertheless, the results suggest that the hybrid nudge is at least as, and perhaps even more effective, than a mere simple nudge or information provision. Thus, the results overall speak in favour of H_1 .

The findings also have implications in terms of the nudge's transparency: First, following the Type 1 and Type 2 nudge differentiation [123], the colour-coded bar used in the password creation decision and the "most popular" banner used in the cloud service choice should be transparent to the user in that they were visible to the user. Contrary to that, the position nudge used in the WiFi choice and the pre-selected "yes" option should be less salient as interventions. At least the exploratory findings from the qualitative analysis supported these assumptions as the strength bar and the banner were mentioned a couple of times as a reason for the participants' choices and appeared to make people reflect on them. In contrast, the position nudge and the default nudge were not explicitly mentioned.

In terms of the strength bar and the banner, two exemplary quotes illustrate the participants' reasoning:

"I tried to make security box green, so I made an inordinately long password." (Password Creation, Condition Dynamic Hybrid Nudge, Main Study)

"The services were fairly similar, but I choose "Cloudy" as it was the most popular. I figure if it's that popular, then it must be good and reliable. If it was poor quality, then few people would use it." (Cloud Service Choice, Condition Hybrid Nudge, Main Study)

The findings also imply that further increasing the nudge's transparency by adding information on *why* the person is being nudged towards a certain option or *how* a secure choice can be supported does not necessarily decrease the simple nudge's effectiveness. Furthermore, the combination does not seem to compromise the idea of nudging as envisioned by Thaler and Sunstein as Sunstein states: "*there is no opposition between education on the one hand and nudges on the other. Many nudges are educative. Even when they are not, they can complement, and not displace, consumer education*" [298, p.207].

Discussion of H_2 : Educational Effects of Information Provision and Hybrid Nudges

Contrary to the assumptions, H_2 concerning the educational effect of information provision or hybrid nudges cannot be confirmed. When participants in the follow-up study were clustered according to the conditions they were assigned to, no significant differences were found across conditions. One exception is the cloud service choice, but no clear pattern in line with or contrary to the hypothesis was found. This indicates that the effect of the intervention does not necessarily translate to future or related interventions in which the nudge is not present. Considering that security is only a secondary aim for users it may well be that users care about security and are willing to make secure choices if assisted so that it does not significantly impact their primary aim, i.e. their primary task. Yet, without the intervention, the primary task may be the users' focus.

One implication might be to display the intervention every time the decision is taken. This may be useful for decisions that are taken only once or infrequently. However, the effects of repeated exposure to the intervention in frequent decisions are not sufficiently explored yet. In the long term, adverse side effects such as reactance to the intervention, annoyance, or habituation may occur. For frequent decisions, it might thus be beneficial to explore interventions that not only help users make a secure decision the first time, but that facilitate choice for future related decisions by offering suitable options or tools. For example, after nudging users towards and informing them about secure privacy settings, users might have the option to make their selection a default setting for future decisions.

Implications for Password Creation

In terms of password creation, it could be shown that the dynamic hybrid nudge encouraged users to create stronger passwords than the control condition, the simple nudge condition, and the static information condition. Even though the difference between the hybrid information condition and the hybrid nudge was insignificant, the descriptive values suggested that the coloured strength bar was rather helpful than counterproductive. Furthermore, that some people referred to the intervention as a reason for their choice is an indication for the users' awareness of the intervention and thus, its transparency.

The comparison of static and dynamic information separately and in the hybrid nudges indicates that dynamic information might be more suitable in supporting users even though the differences between the two hybrid nudge variants were insignificant. Reasons might include the interactivity as such, but also the differences in information provision resulting from it. While the static variant presented all information and suggestions at once and might thus appear overwhelming, the dynamic information only presented the piece of information deemed most helpful for the currently entered password based on an underlying algorithm by Ur *et al.* [313]. It was further presented as a concrete and actionable suggestion. Thus, the dynamic information might be more suitable to mitigate the factors potentially contributing to the ineffectiveness of information provision including information overload, complexity, and a missing link to the decision [24, 132].

Overall, the results suggest that hybrid nudges are a promising strategy to support users in creating secure passwords without constraining the choice set and while being transparent to the user. However, this study only analysed one particular password nudge, that is a password meter with a colour-coding nudge. From this study, it remains unclear whether the intervention would have been equally successful, or even more successful, using another type of nudge targeting another bias, norm, or heuristic. A comparison with the hybrid nudge compensating stronger passwords with later password expiry in the field study is difficult. The field study was not only conducted in another setting and with a different sample but also comprised different information elements. Thus, the next section details a study to compare hybrid password meters that only differ with regard to the type of nudge deployed in them.

9.2 Designing Hybrid Password Nudges for Secure and Usable Password Creation



Figure 22: Research step 10.

Manuscript 7: Zimmermann, V., Marky, K., and Renaud, K. Hybrid password meters for more secure passwords - A comprehensive study of password meters and nudges. *Behaviour & Information Technology* (submitted).

Based on the conclusions of the study described in the previous section 9.1, this section details an online study on the effects of hybrid password meters based on different types of nudges on password creation. The study is described in detail in Manuscript 7 "Hybrid password meters for more secure passwords - A comprehensive study of password meters and nudges" that is currently under review [352].

This is an ‘Original Manuscript’ of an article submitted to *Behaviour & Information Technology*, published by Taylor & Francis Group, available online: <https://www.tandfonline.com/toc/tbit20/current>. The following section is a summary thereof.

The manuscript first describes the process and results of a systematic literature review on password meters. In line with the study described in section 9.1, the review provides indications for the assumption that hybrid password meters are more helpful in encouraging secure password creation than “plain” password meters.

“Plain” password meters provide the user with some form of feedback on password strength such as a text stating “weak” or “strong”. Until that point the password meter can be described as a form of, more or less neutral, information provision.

In contrast, a hybrid password meter provides the feedback in a form that targets a certain bias or norm, and by so doing aims to “nudge” users towards increasing password strength. An example is provided by password meters making use of fear appeals [318] or social comparisons [84]. The literature review further indicates that beyond an indication of password strength, additional information on how to increase password strength, like in the password meter developed by Ur *et al.* [313] and shown in Figure 20, might be a relevant aspect in supporting users’ password creation. A hybrid password meter according to this definition thus consists of a) information in terms of current password strength and on how to increase password strength, and b) a feedback nudge *encouraging* users to increase the password strength.

The first aim of the study described in Manuscript 7 was to test the assumption that hybrid password meters would be more effective in encouraging stronger passwords than a password meter exclusively using a simple nudge or password information.

H_1 : A password meter combining information and a feedback nudge leads to increased password strength values, increased password length, and increased password entropy than either on its own.

The second aim was to compare different hybrid password meter variants based on different types of nudges with each other. The aspects of interest included potential differences in terms of password strength, password memorability, and the users’ perceptions of the password meter. The latter two aspects were especially important in order to support not only secure but also usable password creation. The resulting research questions were formulated as follows:

How do hybrid password meters based on different feedback nudges, e.g., targeting different biases, heuristics, and norms, impact:

- RQ_1 : password creation, that is password strength, length and entropy?
- RQ_2 : password memorability?
- RQ_3 : users’ perceptions of the password meter and password creation?

The following sections describe the design of the hybrid password meters analysed in the online study, the results of the study, and the implications for supporting secure password creation.

9.2.1 Method

In a between-subjects online study, a slightly adapted version of the original password meter developed by Ur *et al.* [313] was compared to six different hybrid password meter variants based on different types of nudges. In addition, two control conditions that exclusively included a simple nudge or password information served as a comparison. Across all conditions, password strength was calculated based on the same 100-point score [313] for comparability (even though visualized differently for the users in different conditions). Additionally, password length and password entropy were measured. The next section details the different hybrid password meter variants before describing the procedure and sample.

Hybrid Password Meters

Overall, the variants tested in the study consisted of seven hybrid password meters and two control conditions. Small exemplary screenshots of some conditions are shown in Figure 23. The original and control conditions are depicted in the upper row and the variations of the original condition in the lower row.

- *Original - Hybrid Nudge*: Building on the results of the study described in section 9.1 and as a basis for the design of different hybrid password meter variants, the password meter developed by Ur *et al.* [313] (Figure 20) was selected.
- *Control - Simple Nudge*: Similar to the simple nudge condition in the previous study, this control condition only provided password strength feedback in the form of a coloured feedback bar targeted to activate the learned connection red-bad/insecure and green-good/secure based on by Ur *et al.* [313].
- *Control - Information*: Similar to the information condition in the previous study, this control condition exclusively provided dynamic password information based on Ur *et al.* [313].

The six variants of the original hybrid password meter not only used different kinds of nudges but were also designed considering different aspects of the password creation context. The first aspect of password creation is the individual *person* creating the password with their own experiences, knowledge, and aims.

A second aspect is the *password creation process* itself. Like other security-related tasks, password creation is a secondary task that users have to complete to get to their primary task such as sending an email or browsing a social network. Finally, password creation takes place within a wider *social context* of the individual that includes previous experiences and behaviours of others, other people’s advice and suggestions, or social comparisons.

For each of the aspects, two hybrid password meter variants were designed to specifically target the person, the password creation context, or the social context. These only differed from the original by Ur *et al.* [313] in terms of the type of nudge employed and respective adaptations of the wording in the instruction and password information.

The Person:

- *(Positive) Fear Appeal Nudge*: The idea of an effective fear appeal is to create both a high perceived threat (e.g., by displaying the time needed to crack a password) as a form of motivation and high perceived efficacy to counteract the threat (e.g., by providing actionable suggestions for improving password security) [337]. To not induce overly negative feelings in participants based on ethical considerations [245], the fear appeal nudge used in this study was framed positively. The participants were informed about how much the time to crack a password increases with their changes to the password. As a visual component, a picture of a hacker was added that became increasingly frustrated with increasing password strength.
- *Motivation Nudge*: The motivation nudge aimed to encourage users to create secure passwords without inducing fear. The intervention consisted of a visualization of a little runner that came closer to the finish line with increasing password strength. The dynamic visualization was supplemented by motivating statements such as "Take a final leap!"

The Password Creation Context:

- *Compensation Nudge*: Similar to the intervention in the field study described in section 6.2, this nudge aimed to compensate for the additional effort of stronger password creation with a later password expiry date. Very strong passwords with the maximum score of 100 points had no expiry date at all following the NIST [115] suggestion to only change secure passwords if there is an indication for a compromised account. The intervention was visualized with a picture of a little calendar in which a date was marked later and later.
- *Reciprocity Nudge*: This intervention was based on the idea of compensating the increased effort for secure password creation with "giving" something from a technical or provider side, respectively. The nudge highlighted the technical efforts undertaken to ensure a secure environment for password creation and storage in a list. A little "technical" strength bar that was already filled and green was displayed next to the list. Participants were then asked to return the favour by contributing to security and filling their "password creation" strength bar as well.

The Social Context:

- *Descriptive Social Norm Nudge*: A descriptive norm relates to the perceptions of what the members of a social group actually do [56]. It can especially influence behaviour when made salient to the person [56]. This intervention thus visualized the participants’ password strength in relation to that of an "average user".

To increase salience, the strength bar also took the form of a person as a representative of the average user and filled with increasing password strength. The design of the descriptive norm nudge was based on the findings of a master's thesis by Christiane Rosa [257] that has been co-supervised by Verena Zimmermann.

- *Injunctive Social Norm Nudge*: In contrast to a descriptive norm that relates to actual behaviour, an injunctive norm relates to the perceptions of desired or accepted behaviour [56]. The intervention consisted of ten emoticons as a strength bar that "rated" the participant's password strength in terms of its acceptability. The reasoning was that the weak password of one person could also endanger other accounts in the system [145]. The design of the injunctive norm nudge was based on the findings of a master's thesis by Christiane Rosa [257] that has been co-supervised by Verena Zimmermann.

Procedure

The online study was conducted via the online platform Amazon Mechanical Turk. To avoid bias from participants who participated in the previous study described in section 9.1, these were excluded from this study. For a quick overview, the study procedure including screenshots of different conditions is visualized in Figure 23.

After providing their informed consent, participants were asked to create a new, not previously used password for a fictional, important online service. For password creation, the participants were randomly assigned to one of the nine conditions. Afterwards, the participants were asked to rate the password creation process, the created password, and their feelings and perceptions when using the password meter. The items were partially inspired by the ones asked by Ur *et al.* [313] and the questionnaire "AttrakDiff" [127], but adapted and extended for the study.

Afterwards, the participants' technological affinity (ATI, [98]), their security intention with regard to passwords (SeBIS, [83]), and basic demographic information were collected.

Finally, participants were asked to re-enter the password they created and to state how they created and memorized the password, e.g., whether they memorized the password or used a password manager. In a follow-up study two weeks later, the participants were asked to reproduce their password again to check for memorability.

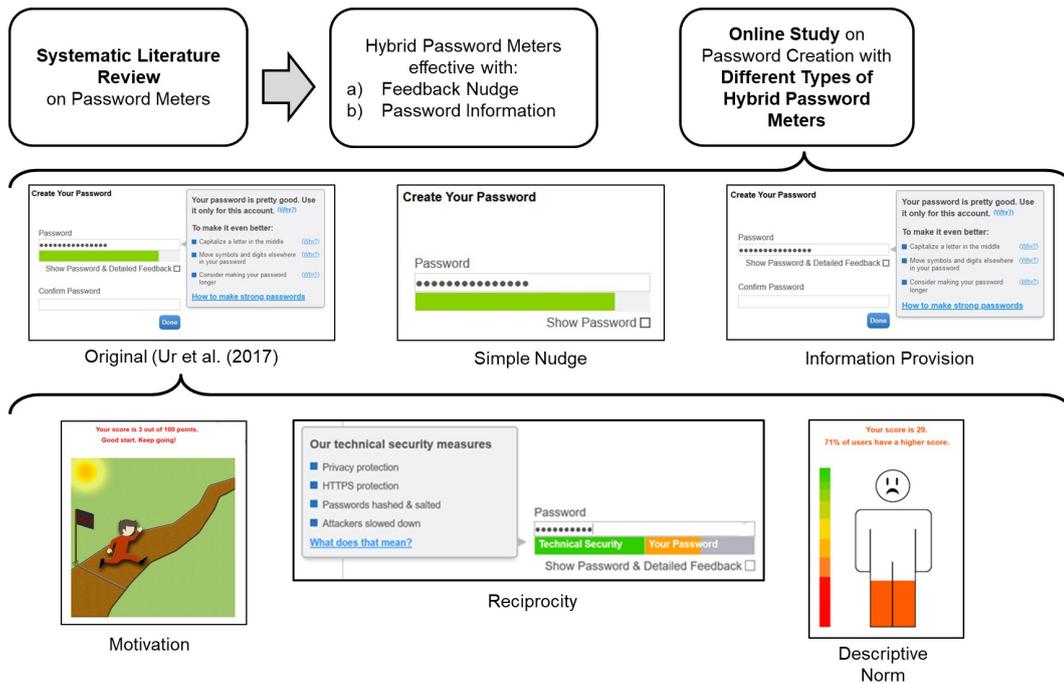


Figure 23: Study procedure of the online study analysing hybrid password meters based on different biases and heuristics as described in Manuscript 7 that has been under review by the journal *Behaviour & Information Technology* at the time of the publication of this dissertation [352]. The exemplary screenshots show varying password scores to illustrate differences in the feedback.

Sample

The sample consisted of $N = 645$ people with residence in the USA. Of these, 379 identified as male, 255 as female, and two as diverse. All participants were aged 18 and older with the following distribution: 6.4% 18-24 years, 48.8% 25-34 years, 26.7% 35-44 years, 10.4% 45-54 years, 4.8% 55-64 years, and 1.6% 65 years and older. The remaining 1.4% did not provide an answer. The majority of people were employed (71.3%), self-employed (17.8%) or college/university students (5.7%). A total of 229 participants reported some kind of IT-related studies or occupation.

The participants were recruited via Amazon Mechanical Turk and compensated with \$2.50 based on a 10\$ per hour rate.

9.2.2 Results

Before the analysis of H_1 and R_1 concerning differences in password strength, length, and entropy differences, all participants that stated to have reused a password or used a password manager to create the password for them were excluded leading to a sample size of $N = 521$. Similar to the previous study described in section 9.1, the ordinal password strength and entropy data, and the non-normally distributed strength data were analysed using non-parametric Kruskal-Wallis tests. Benjamini-Hochberg corrected one-sided Mann-Whitney-U tests were used to follow up on significant findings.

To answer H_1 and RQ_1 , the nine password creation conditions were compared in terms of password strength, length, and entropy. The Kruskal-Wallis tests revealed significant differences across the nine conditions. The highest median values for password strength, length, and entropy were produced in the motivation nudge condition, the lowest in simple nudge condition. A first follow-up comparison concerned the original hybrid nudge and its "components", the simple nudge and information condition. While the original hybrid nudge encouraged longer and higher entropy passwords than the information condition, the differences to the simple nudge were not significant. Further follow-up tests showed that most of the six hybrid nudge variants were more effective in encouraging stronger, longer, and higher entropy passwords than the simple nudge or the information condition on their own.

In particular, the motivation, the fear appeal, the reciprocity, and the descriptive norm nudge were more effective than the information or simple nudge on all three counts, i.e., strength, length, and entropy. The motivation nudge was even more effective than the original hybrid password nudge. Results in terms of the compensation and injunctive norm nudge were mixed and insignificant, respectively.

In terms of RQ_2 concerning memorability, nearly all participants (94.42%) were able to reproduce their passwords at the end of the study. After two weeks, this rate decreased to 34.83% of the people that returned for the follow-up. Exploratory comparisons revealed that the memorability rates were lowest for the simple nudge (21.43%) and information condition (21.21%). In contrast, memorability rates were highest for the original hybrid nudge, the reciprocity nudge, and the compensation nudge (between 44.44 and 50%).

An analysis of the rating items to answer RQ_3 concerning the users' perceptions revealed that the hybrid nudge variants were generally perceived very positive, e.g., as very easy, fun, intuitive and helpful. The simple nudge was described as the least novel and informative, the information condition as least pleasant and motivating. The original hybrid nudge was perceived as the least fun and easy. The created password was rated best in the motivation nudge condition and worst in the simple nudge condition. Furthermore, the participants felt especially appreciated, competent and capable using the fear appeal and compensation nudge, and least competent and assured with the reciprocity nudge.

9.2.3 Discussion & Implications

The results at least partially confirm H_1 in that four conditions of the hybrid password meter variants were more effective in encouraging strong, long, and high entropy passwords than the simple nudge or information condition on their own. Comparing hybrid password nudges amongst each other (R_1), only the motivation nudge was more effective than the original hybrid nudge condition. Furthermore, the exploratory analysis in terms of R_2 shows no indication for a reduced memorability but suggests higher memorability rates for hybrid password meters.

In line with that, also the user perceptions in terms of the hybrid password meter variants (R_3) were more favourable as compared to the simple nudge and information condition. Additionally, the positive perceptions of the (positive) fear appeal nudge indicate that, as intended, no negative feelings were induced in the participants. Furthermore, the at least equally effective motivation nudge indicates that fear appeals might not be a necessary prerequisite for eliciting motivation.

Aside from the differences, it seems that all password meters helped users to align their security perception with technical security measures. The ratings in term of the motivation nudge passwords being "best" and those in the simple nudge condition being "worst" were congruent with the descriptive strength values.

Overall, the findings suggest that hybrid password meters, in general, better support the creation of secure and usable passwords as compared to the single use of information or a simple nudge. Furthermore, they are also perceived as being more supportive by the participants.

However, no clear differences between hybrid password meters targeting either the person, the password creation context, or the social context emerged. Thus, it seems that the *combination* of a) information on password strength and password creation, and b) a nudge encouraging users to increase password strength, is more relevant than the *type of nudge* used within the combination.

Finally, the results should be interpreted against a background similar to that of the previous study in section 9.1. The study was conducted in a somewhat artificial context on Amazon Mechanical Turk for comparability with the previous study and to control for various external influences. Aside from that, the long-term effects of displaying a hybrid password meter frequently require further research.

The next section, the discussion, will take up and reflect on the results and the associated limitations of all studies presented so far in more detail.

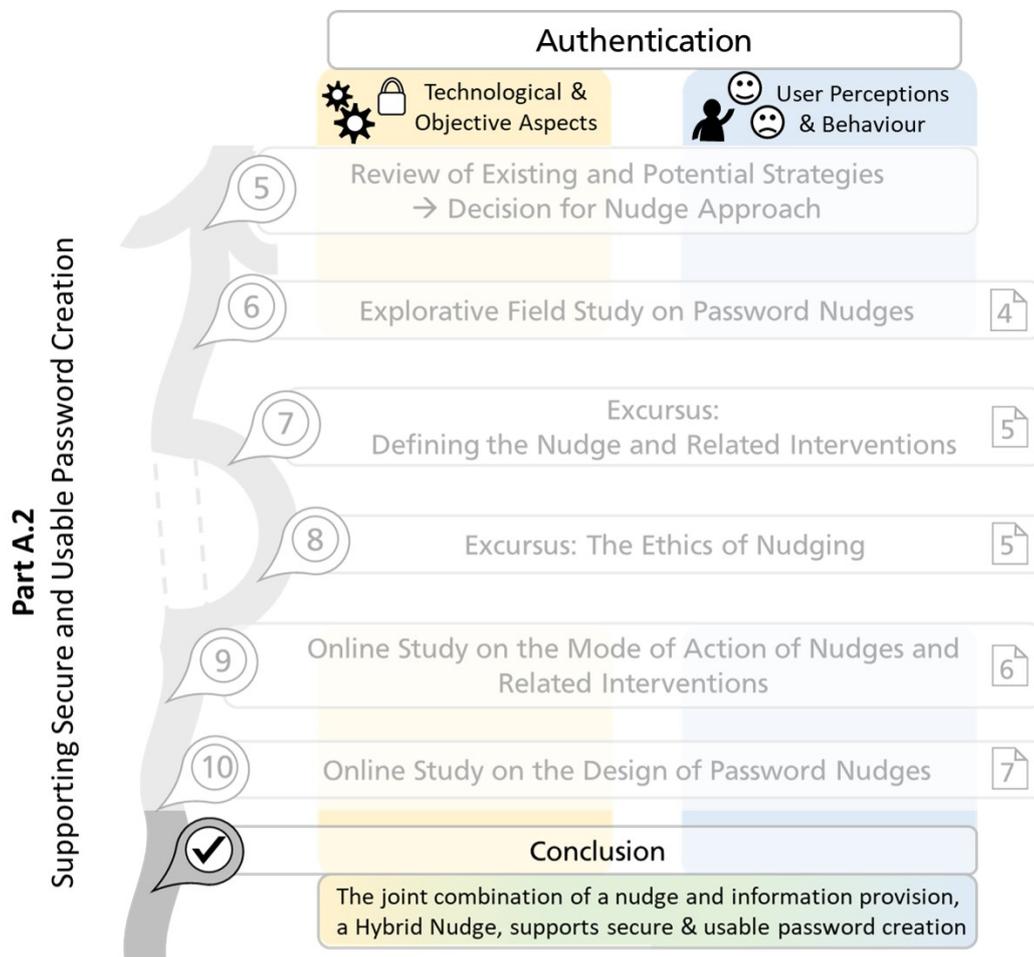


Figure 24: Conclusion

The discussion will start with a brief summary of the main findings across all research steps presented in this dissertation and regarding the two main research questions guiding this research. Afterwards, the results will be interpreted in terms of their contribution and their implications for research and practice. The next section will then reflect on the limitations associated with different methodological and content-related aspects. An outlook on future research on hybrid password meters and beyond the context of password creation will be provided before concluding.

10.1 Summary of Findings

This research was guided by two research aims: first, to explore the user perceptions of different forms of authentication as described in Part A.1, and second, to support secure and usable authentication based on the findings as detailed in Part A.2.

With regard to the first aim, this research started with an analysis and rating of authentication schemes based on the framework developed by Bonneau *et al.* [34]. As a result, the rating led to the selection of ten different authentication schemes out of five different authentication categories (e.g., knowledge-based and biometric schemes). Together with the "challengers" password and fingerprint authentication, the schemes were compared in a laboratory study using interactive mock-ups of each scheme. The findings revealed that user perceptions of the schemes varied significantly, for example, in terms of preference, perceived usability, perceived effort, perceived security, and privacy concerns. Furthermore, while usability-related constructs correlated with the users' preference, security and privacy ratings did not.

However, besides familiarity with a scheme and ease of use, perceived security and privacy concerns were often mentioned as reasons for preferring or disliking a scheme.

Overall, passwords were rated highest in terms of preference and usability-related measures despite the cognitive load passwords pose for users. The biometric fingerprint scheme was rated second in terms of preference, with high usability and security ratings, but also high levels of privacy concerns.

Afterwards, an online study analysed the impact of the type of scheme and the context of use on users' perceptions. A comparison of the knowledge-based password, biometric fingerprint, and token-based PhotoTAN scheme confirmed the assumed single and combined impact of scheme and context on user perceptions. Again, the password scheme was preferred across all contexts of use, even though the fingerprint scheme was perceived as more secure. As in the laboratory study, the fingerprint's high security ratings were accompanied by high privacy concerns.

As described in the interim conclusion, a focus was set on secure and usable password creation in the second part of this research. The reasons included not only the high levels of user preference revealed in the two studies and their pilot studies, but also the recognition of the persistence of passwords. Given the lack of a "silver bullet" in terms of other authentication schemes as an adequate replacement, password authentication will likely be relevant for some more time [133]. Furthermore, research showed that password security not only depends on technical measures but very much on the human factor, that is the users' password creation and password handling. Current password practice, including weak and reused passwords, as well as misconceptions in terms of password security [116, 314, 315], indicates the need for supporting users in creating secure passwords.

The subsequent part pursuing the second aim thus started with a review of strategies for enhancing password security. The aim was to balance technical security requirements with the users' needs. Based on previous work and the human-centred stance taken in this research, it was decided to focus on strategies that support the user without restricting choice or applying strict rules. As a promising, but not yet well-researched approach in terms of password security, the concept of nudging emerged [303].

Three consecutive exploratory field studies trialled several password nudges, small changes to the password creation interface aimed to encourage the creation of stronger passwords without enforcement and by activating automatic cognitive processes [303]. A number of nudges that were supposed to work by priming or by activating social comparisons and norms trialled in the first two studies did not lead to significant improvements. However, a *hybrid nudge*, combining a visual nudge, the compensation of stronger passwords with later expiry dates, and educational elements informing users about password strength in a third study, was effective compared to the results of the second field study.

A reflection on the trialled nudges and methodology led to a detailed examination of the nudge-related literature. This review resulted in a differentiation of the nudge from related interventions such as a code, sludge, or information provision. Based on the field study, the combination of a nudge and information provision was labelled a *hybrid nudge*. Furthermore, guidelines for the ethical deployment of nudges were derived from a discussion of arguments for and against nudging. The transparency of nudges to the user was found to be an important aspect of their ethical deployment. First, nudges themselves can be designed to be more or less visible to the user in terms of informing them about their existence and purpose [123]. Second, the combination of a nudge targeting automatic cognitive processes with information provision targeting reflective reasoning might further contribute to the nudge's transparency.

Therefore, an online study tested the single and joint effects of nudging on different cybersecurity-related decisions, including password creation. It was found that across all decisions, the hybrid nudge was at least as or even more effective in encouraging secure choices as compared to the individual deployment of a nudge or information provision. However, assumed educational effects of information provision on future decisions in which the nudge is absent could not be confirmed. In terms of password creation, in particular, the hybrid nudge was more effective than the exclusive use of a nudge, i.e., a colour-coded strength bar, or the use of static information. Yet, the hybrid nudge was not more effective than dynamic textual information as a variant of information provision.

As the online study only analysed one type of (hybrid) nudge within a password meter, another study was conducted to analyse potential differences in password creation based on nudges targeting different types of biases and heuristics. Another focus was on the usability of hybrid nudges.

Thus memorability and user perceptions were considered as well. The hybrid password nudge analysed in the previous study was compared to six hybrid nudge variations that only differed in terms of the nudges included. The study furthermore included two control conditions, an exclusive nudge and an exclusive information condition.

The results revealed that the hybrid nudges, for the most part, were more effective in encouraging secure passwords as compared to the control conditions. Furthermore, there were no indications for decreased but rather increased memorability rates.

Also, the subjective user perceptions were more favourable with regard to the hybrid nudges. However, with one exception, no significant differences could be found between the different hybrid nudge variants.

The implications and the contribution of these findings will be discussed in the next section.

10.2 Implications & Contribution

The results of this research have implications across the different research steps. This section thus reviews the contribution and implications of the research steps in chronological order as presented in this research.

10.2.1 Rating of Authentication Schemes

The rating of authentication schemes was conducted according to an established framework by Bonneau *et al.* [34] and a refinement proposed by Mayer *et al.* [197]. The rating results have been implemented in ACCESS, a publicly available authentication choice support platform¹ [198, 247]. This research contributed to the platform by extending its database by 40 additional schemes to a total of 85 included schemes. By so doing, researchers and practitioners can profit from the rating results when selecting an authentication scheme for their own purpose. Furthermore, ACCESS includes a discussion module that allows other researchers and practitioners to extend the database further or to suggest changes or updates to already included schemes. Given some of the rating challenges presented in the limitations section 10.3, this option might further enhance the quality and actuality of the rating results.

With regard to the next research step, the rating results of the objective usability, deployability, and security features revealed that despite adding 40 schemes to the database, no easy replacement for the password scheme exists. Given the advantages and disadvantages of different schemes in different aspects, the "best" option depends on the individual prioritization of features, and sometimes also on the concrete implementation of the scheme. Considering user perceptions in the next research step was thus deemed relevant to shine light on the seemingly intractable issue.

10.2.2 Studies on User Perceptions of Authentication Schemes

The laboratory study compared authentication schemes in terms of subjective user perceptions. It contributes to the existing literature in that a large number of schemes across multiple authentication categories was analysed in terms of *actual* as compared to *hypothesized* interaction. So far, subjective user perceptions of authentication schemes have been studied less frequently than technical aspects and often in the form of surveys (e.g., [23, 101, 152]) or focus groups asking users to imagine the interaction with the scheme (e.g., [81]).

In terms of the results, the study contributes by shining light on differences in the users' perceptions and potential reasons for them. For example, the finding that users prefer password authentication despite its downsides in terms of cognitive load is surprising, but reasonable considering the reasons provided by the participants. These included high familiarity with the scheme, ease of use, and speed, especially compared with other partially very complex schemes. Like in other studies, fingerprint authentication scored high in terms of usability and especially security [23, 81, 139]. That it was not preferred over password authentication may be due to the high level of privacy concerns, some participants reporting experience with the scheme only in certain contexts, and about 40% of the participants having no previous experience with biometrics. Thus, the results imply that familiarity or experience with a scheme and the context of use might impact its perception.

¹ available from: <https://access.secuso.org/>

In addition, the combination of the quantitative and qualitative results from this study led to the assumption of a security-privacy trade-off in authentication schemes that may lead some people to reject providing their personal information for biometric schemes even though perceived as very secure. Furthermore, the lack of correlation between security or privacy and preference indicates that the two constructs may also be weighed up with other aspects when users are indicating their preference.

An explanation for this might be that security and privacy are complex constructs that are often invisible to the user. The potential inability to estimate the security and privacy of a scheme might result in giving more weight to other factors such as usability. This has two important implications: First, some kind of security-privacy trade-off is assumed that should be further analysed in future work.

Second, in terms of interface design, security and privacy should be made visible and easily graspable for users. That way, it could be adequately considered in the users' decision process, and potential adverse effects resulting from misconceptions [140] could be reduced. One approach to achieve this is the "security theatre" suggested by Schneier [268] that, correctly used, can contribute to aligning the security perception with the technical security level.

Based on the implications of the laboratory study, an additional online study was conducted to shine light on the assumed security-privacy trade-off, the influence of familiarity, and the context of use on user perceptions. The findings confirmed the previous result that biometric fingerprint authentication was deemed most secure, but less preferred than password authentication, perhaps due to the high level of privacy concerns. This also became visible in the ambiguous distribution of preference ratings in terms of fingerprint authentication. Furthermore, while preference ratings varied across contexts of use, security and privacy perceptions were only related to the scheme. This provides another indication for security and privacy somehow being traded-off with other factors such as usability-related measures that might be easier to include in the decision from a user perspective. Furthermore, even though indications for the impact of familiarity were visible, these could not be confirmed statistically, perhaps due to methodological aspects.

Overall, the implications of the online study confirm and mirror that of the laboratory study: First, user perceptions that differ across schemes and contexts should be considered as a relevant factor influencing acceptance and ultimately security. Second, the potential influence of familiarity on user preference of authentication schemes should be further analysed along with the assumed security-privacy trade-off. Third, the call for including usable security and privacy information in the interface design so that it can be considered in users' decision process needs to be repeated.

10.2.3 Field Studies on Password Nudges

Two field studies trialling several password nudges show that not any nudge works just anywhere but that the consideration of the context and the choice is highly relevant. Aside from the limited variance in the results because of the chosen strength measure, the insignificant findings may well have context-related reasons. Most of the trialled nudges did not provide users with an indication of password strength but just encouraged "stronger" passwords. Even if picked upon by the participants, there might have been uncertainty about whether that aim was accomplished. This might be especially relevant in complex and multi-faceted decisions like password creation where it might not be clear what the "right" choice users are nudged towards actually is.

The certainty provided by the nudge in a third field study might thus have been one of the reasons for its positive outcome as compared to the other nudges. It provided feedback in terms of password strength in a graspable format, that is a later expiry date for a stronger password.

Thus, the findings imply that the context and decision type should be carefully analysed before deploying a nudge. In line with that, Caraban *et al.* [46] and Brown [41] agree that nudges are not "one-size-fits-all" solutions but that the effectiveness of nudges is largely influenced by their fit with the users, their goals, and the aspects of the decision context. Likewise, Lindhout and Reniers [184] suggest to first analyse the situation at hand and the individual behaviour within the situation before selecting or designing a nudge.

The field studies also had implications for the nudge concept that appeared not as clear-cut as initially thought. The implications thereof are described in the next section.

10.2.4 The Concept of Nudging

When reflecting on the nudges trialled in the field study, ambiguity in terms of what actually counts as a nudge arose. For example, priming users with a change of wording from "password" to "secret" to encourage stronger passwords seemed to comply with the original definition provided by Thaler and Sunstein [303]. This includes "*any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives*" [303, p.6].

Furthermore, nudges usually make use of automatic cognitive processes [43, 122, 129] associated with System 1 according to Dual Process Theories [88, 156, 285, 290].

Yet, what about making users reflect on their password strength as tested in another nudge variation? Does reflection, associated with System 2 information processing, still count as a nudge? And what about compensation in the form of later expiry?

Can the compensation be considered a "significant" incentive exceeding the definition? This uncertainty also became visible in the nudge debate and led to some variations of the original definition as, e.g., proposed by Hansen [122] or Calo [43]. Some researchers even questioned the novelty of the concept when compared to information provision [122] and stated that many existing interventions could retrospectively be labelled as nudges [192].

As an implication of this ambiguity, in this research, an extensive literature review was conducted. It led to an adaption of the definition of the nudge and its differentiation from related interventions such as information provision. Even though probably not perfectly clear-cut across all cases either, the differentiation supported the analysis of the single and joint effects of nudging and information provision without mixing the two beforehand. The definition also allowed for a differentiated analysis of the ethical considerations in terms of nudging. The contribution of this research step does not only lie in a summary of arguments for and against nudging, but also in transferring these arguments to ethical principles guiding psychological research inspired by a similar procedure used by McMillan, Morrison and Chalmers [202]. These not only contributed to the design of the nudges used in the following research steps, especially regarding transparency, but may also provide a valuable aid for other researchers. They may use the differentiation of the concepts and the discussion of their ethical considerations to evaluate interventions they intend to use. They may furthermore benefit from the ethical checklist providing first aid for conducting nudge-related research. Sharing the insights with the research community might also encourage feedback from other researchers and contribute to further improvements in terms of the nudge's classification or ethical guidelines.

10.2.5 Online Studies on Hybrid Password Nudges

Both online studies provided indications that the hybrid nudge as a combination of a nudge and information provision is at least as, and sometimes even more effective than either intervention on its own. Even though based on a selection, this could be shown across different cybersecurity decisions in general and for password creation in particular. Furthermore, the hybrid nudges were preferred over exclusive single nudge or information interventions from a user perspective and did not seem to impact memorability of the created passwords negatively. The mainly insignificant differences between the hybrid nudge variants in the second online study indicate that the combination of a carefully selected nudge and information provision as such might be more important than its individual nudge components.

Taken together, hybrid password nudges appear to be a promising approach for supporting secure as well as usable password creation. However, across all implications, it has to be acknowledged that the studies so far analysed individual password creation in a somewhat artificial online context even though research suggests that the results might be comparable to actual password creation [92, 171, 200]. The field studies exploring the application of (hybrid) nudges "in the wild" are relevant in this regard but took place at an early stage of this research so that the results are not directly comparable. Further research is thus required to evaluate the generalisability to multiple real-life accounts. This and other limitations are discussed in section 10.3. Directions for future research to address the discussed aspects are explored in section 10.4.1.

Within the broader authentication context of this research, the promising results in terms of hybrid nudges indicate that their use might also be suitable in other authentication schemes. As suggested by Senarath, Arachchilage, and Gupta [277], hybrid nudges could also be applied for encouraging stronger answers in personal questions as a fallback mechanism.

Furthermore, the literature review on password meters also revealed indications for their applicability to graphical schemes such as Android Unlock Patterns [288, 294]. Future work should thus consider the use of hybrid nudges for encouraging secure and usable authentication across different schemes.

Furthermore, the findings have implications regarding the nudge and its transparency: First, the findings imply that information provision does not necessarily diminish the nudge's effectiveness.

In line with that, also Kroese, Marchiori, and de Ridder [173] found that the combination of a nudge and information provision did not decrease the effectiveness of the intervention when they combined a nudge encouraging healthy food choices with a sign informing about this intention.

Second, the inclusion of information appears favourable in terms of ethical considerations. As information provision is supposed to target System 2 information processing, it may enhance users' awareness of the intervention. This may be especially true in combination with transparent Type 2 nudges. Initial support for this assumption is provided by some participants in the first online study specifically mentioning the nudge intervention as a reason for their decision.

Yet, the information provided in the two online studies provided reasons for encouraging the secure decision and support to decide securely but did not explicitly inform about the nudge. Thus, especially in the case of the invisible Type 1 nudges, users might still have been unaware of this intervention component. Future studies should thus trial more explicit information with regard to the nudge and potentially move towards more transparent Type 2 nudges.

That the hoped for educational and long-term effects of the included information could not be confirmed, has several implications: On the one hand, it suggests that the intervention should always be present to exert its influence. On the other hand, repeated exposure to the intervention in the case of relatively frequent decisions such as password creation may lead to habituation or annoyance. Thus, future research should first analyse the long-term effects of repeated exposure to (hybrid) nudges and second, explore ways to decrease user effort in the long-term. One possibility would be to provide users with the option to make storing the created password in a password manager a default setting after creating a long and secure password with the hybrid nudge for the first time. This way, the cognitive load and the potential number of password resets may be reduced in the long-term and also across multiple accounts. This is also discussed in section 10.4.1, along with other directions for future research.

10.3 Reflection & Limitations

The reflection on this research can be broadly divided into methodological and content-related considerations. Thus, this section will first reflect on the different methodologies and study designs before discussing authentication- and password-related aspects.

10.3.1 Methodological Considerations

Overall, this research included literature-based approaches and empirical studies involving users. These will be detailed in the following.

Literature-based approaches

The literature-based approaches in this research included the analysis and rating of authentication schemes in Manuscript 1, a discussion and reflection on the ethical deployment of nudging in Manuscript 5, a systematic literature review on password meters in Manuscript 7, and a problematization approach used in Manuscript 8.

Rating of Authentication Schemes. The analysis and rating of authentication schemes in Manuscript 1 formed the basis for the selection of authentication schemes for the next research steps. However, it was not without challenges and limitations: First of all, the underlying literature analysis does not claim completeness. As the focus of the analysis was on user-centred authentication, especially technical authentication approaches might not be included in the rating. Furthermore, the rating process as such posed several challenges. Based on their different maturity levels and spread, some schemes were only described in a single paper and as a concept. Others were described across multiple papers with detailed, but also varying descriptions. In the first case, some assumptions for the rating had to be logically derived from the concepts. In the second case, one description out of many had to be selected.

Next, the authentication landscape is dynamically evolving. It is thus possible that some schemes advanced since the rating, while others might become outdated. Moreover, some rating criteria were not clear-cut. For example, the sub-feature "negligible-cost-per-user" may depend on whether the costs per user or stationary access point are considered. These challenges were addressed by having a team of researchers rate and discuss the authentication schemes and applying consistent criteria in ambiguous cases.

Examination of the Nudge Concept. The relevance of the extended examination of the nudge-related literature in Manuscript 5 emerged from the reflection on the insignificant results in the field studies on password nudges (see section 6.2). It became clear that a deeper understanding of the nudge concept and its implications was necessary to differentiate the nudge from related interventions and to design nudges in line with ethical considerations. The literature review included various perspectives and disciplines, such as economics, philosophy, psychology, and medicine. Still, there are limitations to this approach. First, the reviewed arguments for and against nudging, and thus the resulting implications, might not be exhaustive. Second, the ethical principles are formulated in a rather generic way to support understandability and applicability.

Thus, they may not account for the complexity of certain situations or special characteristics of people. For example, on the one hand, deviations from the guidelines, e.g., in that people are only informed about the nudge after the intervention, may be justified in some cases. On the other hand, people with certain disabilities or mental issues might require special care and perhaps a stricter application of the guidelines.

Systematic Literature Review. In contrast to the two examples mentioned before, the systematic literature review on password meters in Manuscript 7 exhaustively included all publications across a number of relevant journals and conferences that included certain search terms and fulfilled certain criteria. Besides, a forward and backward search of all included publications was conducted. Despite these measures, it may be that relevant articles using other terms or published at other venues were not identified. The publications were used to compare the study designs, password meters and their results in terms of the question "what makes effective password meters?". This provided valuable insights into the factors that might positively affect password creation. However, due to differences between studies and contexts, and also sometimes short or partially ambiguous descriptions, no "evidence" could be derived from the comparison. Thus, to analyse the derived assumptions further, an additional study controlling for differences was necessary, as described in Manuscript 7.

Problematization Approach. Manuscript 8 that is described in more detail in the outlook section 10.4, made use of a problematization approach. Introduced by Bacchi [18, 19], the approach is based on the idea that in any discipline measures and solutions are developed based on what is considered to be "the problem". Identifying the "problem" and questioning the underlying assumptions can lead to new perspectives, and thus may open the room for alternative measures and solutions. In this research, the approach has been used to identify what is considered to be the problem in the cybersecurity area. To do so, a number of industry reports, policy documents, and hacker reports have been analysed. Even though the selection of publications has been made based on relevance, it remains a selection. The inclusion of additional publications may have led to a different set of "problematizations" or differences in the weighting of different aspects. Furthermore, even though two researchers were involved, the analysis includes a subjective element that may have influenced the results.

Empirical User Studies

This dissertation included three general types of empirical user studies: a) a laboratory study as described in Manuscript 2, b) online studies conducted via online survey platforms in Manuscript 3, 6, and 7, and c) field studies as described in Manuscript 4. The studies often combined quantitative data in the form of ratings or decisions with qualitative data such as open answers or short follow-up interviews. Each of these study types has certain advantages and limitations. According to McGrath [201], selecting a study design is subject to consideration as it is impossible to maximise generalisability, precision, and realism at the same time.

Laboratory Study. The laboratory study in Manuscript 2 ensured the same setup and devices for all participants. The development of interactive HTML-Mock-Ups to simulate the selected authentication schemes also allowed to control for differences in the design, brand, or maturity levels of the schemes. It furthermore allowed for protecting the participants' privacy as no personal information had to be stored or shared with a third party. In that, precision and control for external influences were high.

A certain degree of realism was created by having interactive mock-ups that responded to the participants' input and by the inclusion of realistic email tasks for which participants authenticated. Yet, the sample consisted of a relatively small and homogeneous group of psychology (in IT) students. The majority of the $N = 41$ participants was female with most being in their early twenties. Thus, the results of the study only have limited generalisability.

Online Studies. Contrary to that, the samples were larger and more heterogeneous in the online studies conducted via the online survey platforms Clickworker or Amazon Mechanical Turk. The participants were more evenly distributed across gender and age and had various occupations. This might contribute to the generalisability of the findings.

Nevertheless, research showed that people who are registered with survey platforms such as Amazon Mechanical Turk might be biased in terms of age, education, or experience with certain types of studies [226]. It also has to be noted that the participants were financially compensated for their participation.

In terms of precision, the online study set-up allowed to control for some variables, such as for the stimuli used, their order, the type of device used, and partially for the participants' attention via attention check items. Still, the environment was less controlled compared to the laboratory study.

It was not possible, for example, to check whether the participants completed the study on their own or whether they were distracted by other tasks or people. The realism of the study might be similar to that of a laboratory study in that the tasks were somewhat artificial. For example, two of the password nudge studies (Manuscript 6 and 7) asked participants to role-play a certain scenario. This might impact the external validity of the findings even though previous research has shown passwords that were created in these kinds of role-playing tasks to be rather representative for actual password creation for real-life accounts [92, 171, 200].

Field Studies. The field studies score highest in terms of realism. The students' real university account with their actual and relevant passwords was used to test the effectiveness of different password nudges. Furthermore, the study did not consist of one-time interventions, but the interventions were integrated in the system for an academic year. For ethical reasons, the only "deduction" in terms of realism was that users were informed about the study beforehand and could freely decide to participate or to reject. However, the high level of realism was accompanied by some downsides. Based on ethical considerations and system restrictions, it was not possible to collect certain demographic or control variables or to have participants rate the interventions. This would have been highly beneficial for understanding why some interventions were not successful. Furthermore, the lack of an adequate control group in the third study led to a provisional but not ideal comparison with the previous study's results. The generalisability of the study was probably higher than that of the laboratory study due to the much larger sample size and the increased realism of the task. However, it was also limited because only students of one university with a major in computer science were studied.

Finally, reflecting on the insignificant findings of the field studies, it might have been beneficial to first test password nudges and related interventions in a more controlled setting before transferring the promising interventions to realistic field settings. Still, taking the opportunity at the time it arose also had an important learning effect. The insignificant findings led to an in-depth reflection and an extended examination of the literature that encouraged a new line of research, as described in Manuscripts 5, 6, and 7. It furthermore resulted in improvements to the methodological setup and a more clear-cut definition of the nudge.

Besides the study designs in general, the samples and the applied password strength measures require further consideration.

Samples. The two studies conducted to address the first research aim in terms of user perceptions of authentication schemes were conducted with German samples. In contrast, the studies to address the second research aim involved English-speaking samples. The change happened for two reasons: First, studies concerning password nudges and especially password meters have often been conducted with English-speaking samples. For example, the password meter by Ur *et al.* [313] that part of this research builds on, was developed and tested with an English-speaking sample. For the purpose of comparability with previous findings, and because the password meter was based on English dictionaries, an English-speaking sample was selected. Second, a research cooperation with an English-speaking professor has been established during the second phase of this research. The use of English as scientific language facilitated the exchange about the study designs, stimuli and results, and their inclusion in joint publications.

Password Strength Measures. The measures to analyse password strength first included a 5-point score metric based on the zxcvbn password strength estimator [331]. After realizing that the score considerably constrained the variance of the results as described in section 6.2, other measures were sought.

As different measures are currently used to estimate password strength as visible in the literature review in Manuscript 7, it was decided to include multiple measures. First, the length of a password is an important contributor to its strength [115, 163, 171, 237]. Second, password entropy, as introduced by Shannon [278], is a measure for the password's complexity or unpredictability. Third, the 100-point password strength score by Ur *et al.* [313] was used for comparability with other studies and is based on a number of heuristics and calculated from the number of guesses an attacker would need to crack the password. Even though the measures partially influence one another, e.g., length contributes to entropy as well as the 100-point score, the measures were applied in parallel to be able to detect and discuss variations in the results.

10.3.2 Content-related Considerations

The content-related considerations concern the interdisciplinary nature and the human-centred focus of this research as well as aspects of password creation.

Interdisciplinary Research. Due to the interdisciplinary nature of this research, it included not only psychological constructs, methodologies, and research questions, but also approaches from other areas such as the rating framework for authentication schemes developed by Bonneau *et al.* [34] or the problematization approach by Bacchi [18, 19]. Furthermore, with regard to the computer science aspect of this research, an intense preoccupation with the technological foundations of authentication and password security took place. However, this research was primarily conducted from a psychological rather than a computer science perspective. While technical security was always considered, for example, in the rating framework and in terms of password strength, the focus of this research was on human-centred security and human perceptions of security. Instead of aiming for "maximum" security from a technological perspective, the aim was to understand user perceptions of security, to support users in increasing password strength, and to align perceptions of security with technological measures. However, a better understanding of the human factor in security may also contribute to a better understanding of the complex and multi-faceted construct of security overall.

Password Creation. The second part of this research focused on supporting secure and usable password creation with the help of (hybrid) nudges. This human-centred approach did not constrain the user in their choice of passwords. No requirements were enforced. While it could be shown that this approach fostered secure password creation overall, it would also allow for very short or very weak passwords. Still, some real-life accounts holding sensitive information, or organisational accounts with extensive access rights, might require a certain minimum password strength. Furthermore, in some cases, there might be legal requirements in terms of authentication as, e.g., detailed in the European Payment Services Directive (PSD2) [62] for banking services. Thus, even though the idea that all users voluntarily increase their password strength above a certain level when adequately supported is desirable, some real-life accounts might need to enforce a minimum length or minimum strength level. However, this does not exclude the use of hybrid password nudges. They might provide feedback to the user in terms of the minimum requirements before trying to encourage even stronger security levels beyond that point.

Furthermore, this research so far focused on individual password creation as it was necessary first to understand the nudge concept and the mechanisms underlying nudging. However, as already acknowledged in the introduction and as found in the online study in Manuscript 3, the context is a relevant aspect for understanding users' choices and perceptions. Users authenticate for many accounts with varying sensitivity. Even though this research showed that hybrid password nudges or hybrid password meters might be beneficial for supporting secure password creation in individual accounts, this effect might not transfer to the many real-life accounts to the same extent. While creating and memorizing a few strong passwords might be possible, it is unfeasible on a large scale. One limitation of this research is that the effects of the interventions across accounts or in the long-term have not yet been researched. Therefore, the next section provides some promising ideas for addressing the scalability and memorability aspect, e.g., by integrating hybrid nudges into tools that work across multiple accounts.

10.4 Outlook & Future Work

This section first provides an outlook in terms of future work on password creation and hybrid password nudges across multiple accounts. Second, the potential future application of the findings beyond the context of password creation to a human-centred cybersecurity approach labelled "Cybersecurity, Differently" is presented.

10.4.1 Considering the Password Creation Context

As already acknowledged in the introduction, users have to manage numerous accounts and associated passwords. Still, security or password advice is often presented without any contextual acknowledgement of the significance and quantity of accounts [293]. So far, also this research focused on individual passwords as a starting point.

Hybrid password nudges or password meters may well support creating secure passwords for an individual account without compromising memorability. However, the interventions did not yet address the memorability issue per se, nor did they consider password creation and reuse across accounts. In this regard, related work points towards at least three paths that might work well in combination with hybrid password meters:

The first is password managers that aim to lessen the cognitive effort for memorizing passwords by storing them securely in a password "vault". As discussed in section 5.2.2, password managers require users only to memorize one strong master password that protects access to all the others. Hybrid password meters could be implemented to guide users towards a very secure master password. Furthermore, a hybrid password meter might also be implemented within the password manager for cases in which the password manager does not offer an automated password generator, or in which users prefer generating passwords themselves over system-generated random passwords. This is especially relevant given the findings of Pearman *et al.* [230] and Lyastani *et al.* [186] who showed that using a password manager does not increase password strength per se. They showed that the influence on password strength depends on the users' strategies and their reliance on technical password-generating tools provided in the password manager.

The other way round, a hybrid password meter could also be used to "nudge" users towards making use of a password manager in the long-term. Considering the uncertain and perhaps adverse effects of frequent exposure to hybrid password meters as discussed in sections 9.1.4 and 10.2, a hybrid password meter could include a link to a password manager. After creating a strong password with a hybrid password meter for the first time, the user could choose to make storing the created password in a password manager a default setting in order to reduce cognitive load. The user might further have the option to completely switch from self-created to computer-generated and stored passwords in the future. This would be in line with the idea discussed in section 9.1.4 to not only help users to make a secure decision the first time but to facilitate choice for future related decisions by offering suitable options or tools.

The second path concerns password feedback mechanisms that consider password strength and similarity across different user accounts. One such promising example is provided by Kim *et al.* [167]: In a dashboard, the users' passwords are rated against each other while considering password similarity and the sensitivity of certain accounts. The same password used for the banking account and a newsletter would be visually grouped to make users aware of the similarity. It would also be rated differently in terms of strength, i.e., as weaker for the more sensitive banking account. It might be helpful to integrate a hybrid password meter into the dashboard to support strong password creation when users decide to change a password to increase their overall strength score.

The third path includes portfolio approaches that acknowledge that completely ruling out weak passwords and reuse are unrealistic, given the number of accounts that users manage [96]. Instead, Florencio, Herley, and van Oorshot [96] suggest to consider realistic attack scenarios and to find suitable compromises in terms of risks and effort. For example, the authors present some principles, including the grouping of accounts for password reuse. In line with that, Zhang-Kennedy, Chiasson, and van Oorshot [345] reviewed and updated wide-spread password advice to facilitate realistic human management of passwords. For example, they suggest to strategically reuse passwords to match the account value or to keep written passwords hidden instead of completely arguing against writing passwords down. In this context, hybrid password nudges or hybrid password meters could be used to support the creation of passwords for very sensitive accounts. However, they might not be used when creating passwords for non-sensitive accounts.

These considerations are mirrored by the concept of "Equitable Security" [241]. It suggests to not deploy nudges for just everything, but to consider increased costs and effort.

10.4.2 Cybersecurity, Differently

This section first describes the authentication-related insights that inspired a new viewpoint on cybersecurity. Second, the resulting "Cybersecurity, Differently" approach described in Manuscript 8 is summarized.

From Password Creation to Cybersecurity

In as early as 1987 Paans and Herschberg acknowledged "*It seems sensible to consider the user, for a change, as an adult partner and to accept that he has certain responsibilities of his own when using the system. One of those responsibilities definitely is the control of his own password.*" [224, p.409].

Even though studies on the generalisability to real-life accounts are yet to be conducted, the findings of this research indicate that users might well be able to live up to that responsibility. Although the hybrid password nudges did not constrain the users in their password choices, they encouraged stronger password creation voluntarily.

For this potential to unfold, it was found important to consider several aspects: Users likely create weak passwords or reuse passwords not because they do not want to protect their data but to cope with the cognitive load posed by multiple passwords [94, 293], misconceptions in terms of technical security measures [275, 314, 315], and security being a secondary rather than a primary aim [332]. Understanding the user's aims, the context of the task, and the implications of deployed measures thus seems highly relevant to enable users to contribute to success. In line with that, this authentication-related research revealed the importance of considering, e.g., potential security-privacy trade-offs (Manuscript 2), the sensitivity of the account type (Manuscript 3), the increased effort associated with stronger passwords (Manuscript 4), the ethical implications of nudging (Manuscript 5), or the relevance of feedback for aligning perceptions with technical security measures (Manuscript 6 and 7).

Comparisons with more constraining measures such as password policies further indicate that these do not necessarily lead to stronger passwords than the use of password meters [171, 175, 313], even when the meter enforces no minimum requirements [338]. Furthermore, like all strategies, password policies are accompanied by their own disadvantages, including the creation of predictable password patterns [237] or usability issues [142].

Thus, overall, it seems that providing users with flexibility and suitable support is a promising approach for enabling users to be "*the first line of defense*" [128, p.20]. The next section describes how these insights may be of relevance beyond the context of authentication and password creation.

Summary of "Cybersecurity, Differently"

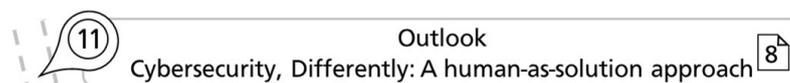


Figure 25: Research Step 11.

Manuscript 8: Zimmermann, V., and Renaud, K. Moving from a 'human-as-problem' to a 'human-as-solution' cybersecurity mindset. *International Journal of Human-Computer Studies* 131 (2019), 169–187. doi:10.1016/j.ijhcs.2019.05.005.

This section is a brief summary of Manuscript 8 "*Moving from a 'human-as-problem' to a 'human-as-solution' cybersecurity mindset*" by Zimmermann and Renaud [353] that details the findings of applying a "problematization approach" [18, 19] to the area of cybersecurity. In the following, the problematization approach is briefly described before summarizing its outcome and the main guidelines derived from questioning underlying assumptions.

According to Bacchi [18, 19], in different areas, the assumption about what the "problem" is, informs the measures and solutions to address the problem. For example, if a crime was viewed to be caused by poverty, measures might be directed at changing the circumstances. However, if a crime was assumed to be caused by a person's behaviour, measures might be developed to constrain or change the behaviour.

A mismatch might occur if the assumptions are incomplete, outdated, or not based on evidence. It is thus important to reveal and question the underlying assumptions. This can be done via "problematizing" using the "What's the problem represented to be" (WPR) approach [18] that has been originally applied to feminist theory and critical policy studies.

The research aimed to analyse what the "problem" in cybersecurity is considered to be, to question the assumptions, and, if applicable, to suggest alternative viewpoints to the problem. To do so, a slightly adapted version of the original approach by Bacchi [18] was conducted. The six process steps are graphically depicted in Figure 26 and summarized below.

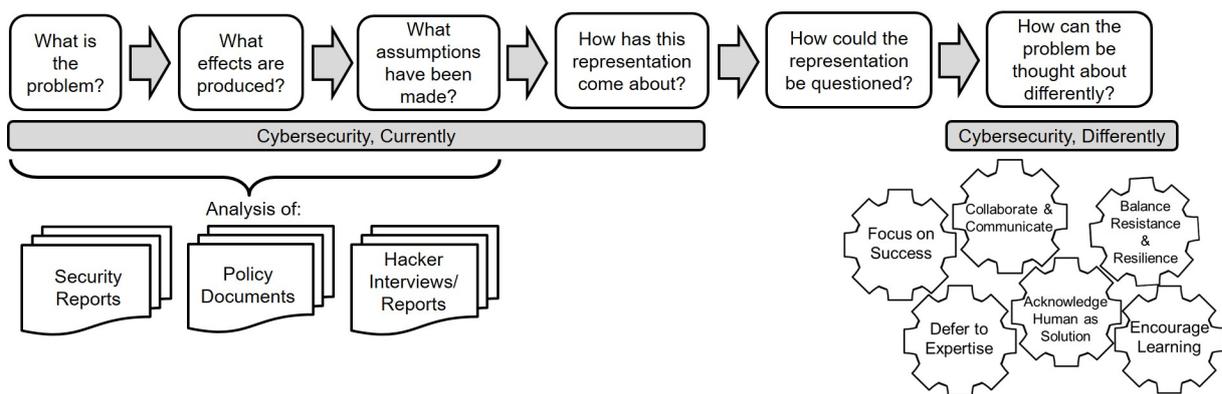


Figure 26: Graphical depiction of the problematization approach and the "Cybersecurity, Differently" principles derived thereof, adapted from [353].

What is the problem? To uncover the "problem" in the cybersecurity area, a selection of governmental policy documents, security reports from industry and research, as well as some interviews with and reports about hackers have been analysed using an open coding approach.

A code was assigned whenever a) a specific cybersecurity problem, b) a measure, or c) a strategic aim was mentioned. For example, the notion "lack of clarity around who exactly is responsible" in the Cisco security report [58, p.24] was coded as a cybersecurity problem. The notion "it is vital we place a strong focus on securing our information systems and building the skills" in the New Zealand strategy document [213, p.2] was coded as strategic aim.

After coding all documents, the codes were clustered into 18 "problem" categories by two researchers. These could be further summarized into eight problems concerning the individual and ten problems concerning the societal level. On the individual level, these, for example, included a lack of cybersecurity awareness, knowledge and skills (Problem 1), or people not following security best practices (Problem 4). On the societal level, a lack of global communication and collaboration (Problem 12), or the inability to defend against and respond to cybersecurity threats (Problem 13) were identified.

Overall, the process revealed that the human actor in various roles, either explicitly or implicitly, has often been considered a "problem" in cybersecurity. This concerns, for example, developers who create and maintain security technologies, policymakers who develop security standards, and end-users who use security technologies for various purposes. Thus, this first research step revealed a "human-as-problem" viewpoint.

What effects are produced? Like in the previous step, the codings from the document analysis were used to reveal the measures designed to deal with the "problem", i.e., the human actor.

The analysis revealed that to identify the problem or responsible person, and to prevent future adverse events, often root cause analyses are conducted in the aftermath of an adverse event. They rely on the assumption that adverse events can be traced back to single components within the socio-technical system.

The measures to deal with the "human-as-problem" assumption could be clustered into three major groups: (1) exclude the human actor from the socio-technical system, e.g., by use of automation, (2) educate and train the human actor, e.g., with regular training or use of manuals, and (3) control and constrain the users, e.g., by using security policies.

What assumptions have been made? For each of the measures supposed to deal with the "human-as-problem" perspective, the underlying assumptions were derived. For example, the measure to educate and train people is based on the assumption that human actors lack sufficient awareness, knowledge, and the ability to act securely.

All assumptions taken together form the current viewpoint on the cybersecurity problem, labelled "Cybersecurity, Currently". It is based on the assumption that socio-technical systems are decomposable so that problems can be traced back to individual components such as the human actor via root cause analyses. The current viewpoint is characterized by a resistance stance with the aim to prevent errors by applying the excluding, educating, and constraining measures described above.

How has this representation come about? Analysing the historical roots can be helpful for better understanding the revealed assumptions. However, an extensive exploration of the historical development was outside the scope of the analysis.

How could this representation be questioned? Again, each of the revealed assumptions was analysed in terms of how it could be questioned. For example, the assumption that a problem in today's complex and interconnected socio-technical systems can be traced back to an individual component appears increasingly difficult or mere impossible. Regarding the cybersecurity domain, for example, the World Economic Forum [339, p.7] acknowledges that the "*greater interdependence among different infrastructure networks is increasing the scope for systemic failures – whether from cyberattacks, software glitches, natural disasters or other causes – to cascade across networks and affect society in unanticipated ways*". Thus, systemic failures appear to emerge from the interactions between the system components and external influences rather than being caused by a single component.

The "human-as-problem" assumption could also be questioned with instances in which human actors considerably contribute to security, as illustrated by the following quote from the Microsoft report: "*An employee that spots and reports a suspicious email could head off an extensive phishing campaign. And employees that note unexpected latency in systems can set off investigations that uncover lurking threat actors.*" [128, p.20].

How can the problem be thought about differently? So far, the analysis relied on the document analysis and the categories derived from the coding approach. However, to question the underlying assumptions of "Cybersecurity, Currently" and to derive an alternative viewpoint, relevant insights from related disciplines were considered. For example, in the field of management, Hart [125] describes how controlling the human may create a situation in which the human *becomes* the problem. In contrast, allowing people to take responsibility can increase their agency for acting securely.

The related field of safety science also offers valuable insights: Similar to the security area, in safety science, human actors have long been considered a problem to control, and problems have often been labelled "human error" [71, 131]. Yet, with increasing system complexity, safety scientists acknowledged that problems could no longer be easily traced back to individual system components such as individual human actors [73, 233]. At the same time, some researchers analysed the nearly error-free functioning of complex socio-technical systems such as nuclear aircraft carriers [256]. They found that the human actors within the systems contributed to that success. The systems, among others, were characterized by high flexibility, deference to expertise, and a focus on safety. Threats were actively looked for and addressed, and errors were considered a possibility to learn. Based on these insights and later approaches (e.g., [70, 137, 138]), a new viewpoint on safety emerged. "Safety, Differently" no longer considered the human "*a problem to control*" [70, p.13] but "*a solution to harness*" [70, p.235].

Based on the adaption of relevant principles from "Safety, Differently" and related approaches to a new area of application, the alternative viewpoint on security, "Cybersecurity, Differently", was developed. "Cybersecurity, Differently" comprises the following principles:

- *Principle A: Acknowledge the Human Actor's Ability to be Part of the Solution.* This principle does not suggest that humans do not make mistakes, but that they have the ability to contribute to the solution when supported rather than constrained. Error and success are two sides of the same coin, and the label can only be provided in hindsight [137].

-
- *Principle B: Balance Resistance & Resilience.* In contrast to the resistance stance taken by "Cybersecurity, Currently", resilience describes the capability to flexibly adapt to and recover from unanticipated events [138]. In order to build resilience, systems should anticipate unexpected events, monitor the past and current situation, flexibly respond to emerging situations, and learn from negative as well as positive outcomes [138].
 - *Principle C: Communicate & Collaborate.* Communication and collaboration are relevant in two regards: first, the communication and collaboration in human-technology "teams" that should complement rather than replace each other, and second, communication between humans. Transparent communication within and outside of teams, or even beyond organizations and states, as called for by PwC [48], may contribute to increased resilience.
 - *Principle D: Defer to Expertise.* Deference of expertise suggests that regardless of any hierarchy the person with the highest level of expertise for a certain task should be involved in the decision process [70]. Besides security experts, this also includes end-users that have a high level of expertise with regard to their profession, tasks, and aims.
 - *Principle E: Encourage Learning.* To learn from and prevent future adverse events, it should be focused on *how* it happened instead of looking for someone to blame [69]. To establish a learning culture, organizations could, for example, establish risk-free and anonymous reporting [69].
 - *Principle F: Focus on Success.* Besides the things that go wrong, it should also be focused on normal operation and success, which form the vast majority of events [137]. Often, the same factors that contribute to failure also contribute to success [72, 138]. Considering both sides can thus provide valuable insights.

These principles provide an initial step towards appreciating and fostering the human actor's potential to contribute to success in the area of cybersecurity. They acknowledge the complexity and interconnectivity of today's socio-technical systems in which problems emerge from the interaction of processes and components. Considering the human factor in cybersecurity by applying these principles, opens the path to a more human-centred approach in which the human actor is considered an equal partner in cybersecurity, instead of a problem to control.

Nevertheless, it needs to be acknowledged that the "human-as-problem" outcome of the analysis, and thus also the outcome of the subsequent steps, was based on a small selection of documents. The "Cybersecurity, Differently" principles are thus preliminary. Furthermore, the approach is based on the assumption that the general human actor aims to do a good job rather than to compromise security actively. Still, malicious behaviour cannot be excluded. Thus, a balance between providing flexibility to the large majority of well-intended human actors and measures to detect malicious actions should be sought.

The aim of this research primarily is to spark interest in the vision of "Cybersecurity, Differently" and to encourage future work towards this human-centred approach. For example, future research with regard to the principles and their practical deployment would be highly valuable. The development and evaluation of concrete measures may help to make the vision of "Cybersecurity, Differently" more graspable and to contribute to a change in perspective.

10.5 Conclusion

This line of interdisciplinary research considered the area of authentication from a psychological and human-centred perspective. Thereby, it targeted two main aims: first, to explore the user perceptions of different forms of authentication, and second, to support secure and usable authentication. To address these aims, this research made use of different methodological approaches including qualitative and quantitative data as well as elements from a number of disciplines, such as psychology, computer science, behavioural economics, ethics, and policy studies.

Regarding the first research aim, the combination of a literature review, a comparative rating of authentication schemes, and two types of user studies led to acknowledging the relevance of password authentication already believed dead by some (e.g., [20, 189]). An extension and actualization of an authentication scheme rating conducted by Bonneau *et al.* [34] revealed that replacing the password with a scheme outranking it in all regards still is an intractable issue. Furthermore, even though users are aware of the downsides of passwords, they seem to prefer passwords because of their year-long experience with them and ease of use, at least when compared to other existing schemes.

For example, password authentication was rated less complex than some cognitive schemes and was not accompanied by the privacy concerns often mentioned in relation to biometric schemes.

The consideration of these findings and the fact that password security is largely dependent on the users' password creation and handling led to an exploration of measures to support secure and usable password creation as a next step. From a review of existing strategies, Thaler and Sunstein's concept of nudging [303] emerged as a promising approach for encouraging stronger password choices without constraining the user. The results of three consecutive field studies, a review of the nudge concept and its ethical implications, and two online studies revealed hybrid nudges as a promising strategy for supporting users in creating secure passwords.

Hybrid nudges are essentially a combination of a nudge and information provision. The results indicate that the combination can potentially leverage the power of the nudge, while increasing the intervention's transparency by informing users about the reasons for encouraging secure password creation and by providing feedback on goal attainment. Even though the results form relevant steps towards the second research aim, they still warrant future work to evaluate the hybrid nudge's effectiveness in real-life settings, across multiple accounts, and in the long-term. Next steps would thus include the transfer of hybrid nudges to actual accounts and their integration with concepts that consider the multitude of user accounts with varying sensitivity.

Overall, this line of research provided relevant insights regarding user perceptions of authentication schemes as such and across different contexts of use. The analysis of the nudge concept and its ethical implications may provide an aid for future nudge-related research and application given the debate surrounding the ethical deployment of nudges. Furthermore, the concept of hybrid nudges contributes to empirical research on digital nudging that appears promising not only for future password-related research, but also for the wider field of cybersecurity. Finally, the finding that supporting instead of constraining users can lead to successful results provides potential for an even more human-centred approach in cybersecurity as envisioned in the "Cybersecurity, Differently" mindset.

Authors: Verena Zimmermann, Nina Gerber, Peter Mayer, Marius Kleboth, Alexandra von Preuschen, Konstantin Schmidt

Status: published

The article has originally been published as: Zimmermann, V., Gerber, N., Mayer, P., Kleboth, M., von Preuschen, A., and Schmidt, K. Keep on rating – on the systematic rating and comparison of authentication schemes. *Information & Computer Security* 27, 5 (2019), 621–635. doi:10.1108/ICS-01-2019-002.

© 2019 Emerald Publishing. All rights reserved.

Note: This article is an extended version of Zimmermann, V., Gerber, N., Kleboth, M., von Preuschen, A., Schmidt, K., and Mayer, P. The Quest to Replace Passwords Revisited – Rating Authentication Schemes. In *Proceedings of the 12th International Symposium on Human Aspects of Information Security & Assurance (HAISA 2018)*, Dundee, UK (Plymouth, UK, 2018), Plymouth University, 38-48.

Supplementary Material: The complete results and a description of the rated authentication schemes and rating features can be accessed with the following link: http://www.arbing.psychologie.tu-darmstadt.de/home/forschung_4/forschungsergebnisse_fai.de.jsp

Journal Metrics:

- Computing Research & Education (CORE) Ranking: C (Journal renamed from Information Management & Computer Security) ²
- ClarivateTM Index: Emerging Sources Citation Index ³
- Google H5 Index: not applicable
- Scopus[®] CiteScore 2019: 2.2 ⁴
- Thomson ReutersTM Impact Factor 2019: 1.515 ⁵
- Metrics accessed: 27th August 2020

² <http://portal.core.edu.au/jnl-ranks/?search=information+and+computer+security&by=all&source=CORE2020&sort=atitle&page=1>

³ <https://mjl.clarivate.com/search-results>

⁴ <https://www.scopus.com/sourceid/21100421900>

⁵ <https://www.scimagojr.com/journalsearch.php?q=21100421900&tip=sid&clean=0>

Authors: Verena Zimmermann, Nina Gerber

Status: published

The article has originally been published as: Zimmermann, V., and Gerber, N. The password is dead, long live the password—a laboratory study on user perceptions of authentication schemes. *International Journal of Human-Computer Studies* 133 (2020), 26–44. doi:10.1016/j.ijhcs.2019.08.006.

© 2019 Elsevier Ltd. All rights reserved.

Supplementary Material: The supplementary material including screenshots of all mock-ups used in the study can be accessed using the following link: <https://doi.org/10.1016/j.ijhcs.2019.08.006>.

Journal Metrics:

- Computing Research & Education (CORE) Ranking: A ⁶
- ClarivateTM Index: Science Citation Index (SCI), Social Science Citation Index (SSCI) ⁷
- Google H5 Index: 43 (Median: 70) ⁸
- Scopus[®] CiteScore 2019: 5.8 ⁹
- Thomson ReutersTM Impact Factor 2019: 4.572 ¹⁰
- Metrics accessed: 27th August 2020

⁶ <http://portal.core.edu.au/jnl-ranks/?search=international+journal+of+human-computer+studies&by=all&source=CORE2020&sort=atitle&page=1>

⁷ <https://mjl.clarivate.com/search-results>

⁸ https://scholar.google.de/citations?hl=de&view_op=search_venues&vq=international+journal+of+human-computer+studies&btnG=

⁹ <https://www.scopus.com/sourceid/12960>

¹⁰ <https://www.scimagojr.com/journalsearch.php?q=12960&tip=sid&clean=0>

Authors: Verena Zimmermann, Paul Gerber, Alina Stöver

Status: Submission in preparation.

Supplementary Material: The supplementary material includes the visual description of the registration and authentication procedure of the analysed authentication schemes, the scenario descriptions, a list of items and scales used, a list of the demographics collected, a summary of the selected reasons for authentication preference sorted by context of use, and a table summarizing the differences in reasons for authentication preferences.

Authors: Karen Renaud, Verena Zimmermann

Status: published

The article has originally been published as: Renaud, K., and Zimmermann, V. Nudging folks towards stronger password choices: providing certainty is the key. *Behavioural Public Policy* 3, 2 (2019), 228–258. doi:10.1017/bpp.2018.3.

Renaud, K., and Zimmermann, V. Nudging folks towards stronger password choices: providing certainty is the key - CORRIGENDUM. *Behavioural Public Policy* 3, 1 (2019), 127–127. doi:10.1017/bpp.2018.31.

© 2018/2019 Cambridge University Press. All rights reserved.

Journal Metrics:

- Computing Research & Education (CORE) Ranking: not applicable
- ClarivateTM Index: relatively new journal, not indexed yet
- Google H5 Index: relatively new journal, not indexed yet
- Scopus[®] CiteScore 2019: relatively new journal, no CiteScore yet
- Thomson ReutersTM Impact Factor 2019: relatively new journal, no impact factor yet
- Metrics accessed: 27th August 2020

Authors: Karen Renaud, Verena Zimmermann

Status: published

The article has originally been published as: Renaud, Renaud, K., and Zimmermann, V. Ethical guidelines for nudging in information security & privacy. *International Journal of Human-Computer Studies* 120 (2018), 22–35. doi:10.1016/j.ijhcs.2018.05.011.

© 2018 Elsevier Ltd. All rights reserved.

Journal Metrics:

- Computing Research & Education (CORE) Ranking: A ⁶
- ClarivateTM Index: Science Citation Index (SCI), Social Science Citation Index (SSCI) ⁷
- Google H5 Index: 43 (Median: 70) ⁸
- Scopus[®] CiteScore 2019: 5.8 ⁹
- Thomson ReutersTM Impact Factor 2019: 4.572 ¹⁰
- Metrics accessed: 27th August 2020

Authors: Verena Zimmermann, Karen Renaud

Status: published

The article has originally been published as: Zimmermann, V., and Renaud, K. The Nudge Puzzle: Matching Nudge Interventions to Cybersecurity Decisions. *ACM Transactions on Computer-Human Interaction* 28, 1 (2021), 7:1-7:45. doi:10.1145/3429888.

© 2021 Copyright held by the owner/author(s). Publication rights licensed to ACM.

Supplementary Material: The supplementary material for this article contains the scenario descriptions, images of the HTML mock-ups used in the study, a list of the items and scales, and the code books for the qualitative analysis (including category descriptions and exemplary quotes). It can be accessed via: <https://dl.acm.org/doi/10.1145/3429888>.

Journal Metrics:

- Computing Research & Education (CORE) Ranking: A* ¹¹
- ClarivateTM Index: Science Citation Index Expanded (SCIE) ¹²
- Google H5 Index: 34 (Median: 53) ¹³
- Scopus[®] CiteScore 2019: 6.4 ¹⁴
- Thomson ReutersTM Impact Factor 2019: 5.884 ¹⁵
- Metrics accessed: 27th August 2020

¹¹ <http://portal.core.edu.au/jnl-ranks/?search=transactions+on+computer+human+interaction&by=all&source=CORE2020&sort=atitle&page=1>

¹² <https://mjl.clarivate.com/search-results>

¹³ https://scholar.google.de/citations?hl=de&view_op=search_venues&vq=transactions+on+computer+human+interaction&btnG=

¹⁴ <https://www.scopus.com/sourceid/26199>

¹⁵ <https://www.scimagojr.com/journalsearch.php?q=26199&tip=sid&clean=0>

Authors: Verena Zimmermann, Karola Marky, Karen Renaud

Status: Under review by the journal Behaviour & Information Technology, published by Taylor & Francis Group, available online: <https://www.tandfonline.com/toc/tbit20/current>.

Journal Metrics:

- Computing Research & Education (CORE) Ranking: B ¹⁶
- ClarivateTM Index: Science Citation Index Expanded (SCIE), Social Sciences Citation Index (SSCI) ¹⁷
- Google H5 Index: 36 (Median: 48) ¹⁸
- Scopus[®] CiteScore 2019: 3.5 ¹⁹
- Thomson ReutersTM Impact Factor 2019: 2.697 ²⁰
- Metrics accessed: 27th August 2020

¹⁶ <http://portal.core.edu.au/jnl-ranks/?search=behaviour+and+information+technology&by=all&source=CORE2020&sort=atitle&page=1>

¹⁷ <https://mjl.clarivate.com/search-results>

¹⁸ https://scholar.google.de/citations?hl=de&view_op=search_venues&vq=behaviour+%26+information+technology&btnG=

¹⁹ <https://www.scopus.com/sourceid/12125>

²⁰ <https://www.scimagojr.com/journalsearch.php?q=12125&tip=sid&clean=0>

Authors: Verena Zimmermann, Karen Renaud

Status: published

The article has originally been published as: Zimmermann, V., and Renaud, K. Moving from a ‘human-as-problem’ to a ‘human-as-solution’ cybersecurity mindset. *International Journal of Human-Computer Studies* 131 (2019), 169–187. doi:10.1016/j.ijhcs.2019.05.005.

© 2019 Elsevier Ltd. All rights reserved.

Supplementary Material: Audio slides summarizing the article as well as short method article describing the applied problematization approach can be retrieved from the supplementary material files from: <https://doi.org/10.1016/j.ijhcs.2019.05.005>.

Journal Metrics:

- Computing Research & Education (CORE) Ranking: A ⁶
- ClarivateTM Index: Science Citation Index (SCI), Social Science Citation Index (SSCI) ⁷
- Google H5 Index: 43 (Median: 70) ⁸
- Scopus[®] CiteScore 2019: 5.8 ⁹
- Thomson ReutersTM Impact Factor 2019: 4.572 ¹⁰
- Metrics accessed: 27th August 2020

Bibliography

- [1] Abo-Zahhad, M., Ahmed, S. M., and Abbas, S. N. A new multi-level approach to EEG based human authentication using eye blinking. *Pattern Recognition Letters* 82 (2016), 216–225. doi:10.1016/j.patrec.2015.07.034.
- [2] Acquisti, A., Adjerid, I., Balebako, R., Brandimarte, L., Cranor, L. F., Komanduri, S., Leon, P. G., Sadeh, N., Schaub, F., Sleeper, M., Wang, Y., and Wilson, S. Nudges for privacy and security: Understanding and assisting users' choices online. *ACM Computing Surveys (CSUR)* 50, 3 (2017), 1–41.
- [3] ACT Government - Transport Canberra and City Services. Evaluation of Effectiveness of the 'Smiley Face' Sign Trial. Website, 2020. Online available from www.cityservices.act.gov.au/__data/assets/.../Evaluation-Report-Smiley-Face-Signs.pdf; last accessed 14 September 2020.
- [4] Adams, A., and Sasse, M. A. Users are not the enemy. *Communications of the ACM* 42, 12 (1999), 40–46. doi:10.1145/322796.322806.
- [5] Adams, A., Sasse, M. A., and Lunt, P. Making passwords secure and usable. In *People and Computers XII - Proceedings of HCI '97*, H. Thimbleby, B. O'Conaill, and P. J. Thomas, Eds. Springer, London, UK, 1997, pp. 1–19. doi:10.1007/978-1-4471-3601-9_1.
- [6] Agrafioti, F., Gao, J., and Hatzinakos, D. Heart Biometrics: Theory, Methods and Applications. In *Biometrics*, J. Yang, Ed. InTechOpen, Shanghai, China, 2011, pp. 199–216. doi:10.5772/18113.
- [7] Alkaldi, N., Renaud, K., and Mackenzie, L. Encouraging password manager adoption by meeting adopter self-determination needs. In *Proceedings of the 52nd Hawaii International Conference on System Sciences (HICSS 2019)*, Grand Wailea, HI, USA (Honolulu, HI, USA, 2019), University of Hawai'i at Manoa, pp. 4824–4833. isbn:9780998133126.
- [8] Allcott, H., and Kessler, J. B. The welfare effects of nudges: A case study of energy use social comparisons. *American Economic Journal: Applied Economics* 11, 1 (2019), 236–276. doi:10.1257/app.20170328.
- [9] Almuhiemedi, H., Schaub, F., Sadeh, N., Adjerid, I., Acquisti, A., Gluck, J., Cranor, L. F., and Agarwal, Y. Your location has been shared 5,398 times! A field study on mobile app privacy nudging. In *Proceedings of the 33rd ACM Conference on Human Factors in Computing Systems (CHI '15)*, Seoul, South Korea (New York, NY, USA, 2015), ACM, pp. 787–796. doi:10.1145/2702123.2702210.
- [10] Alsaadi, I. M. Physiological Biometric Authentication Systems, Advantages, Disadvantages And Future Development: A Review. *International Journal of Scientific & Technology Research* 4, 12 (2015), 285–289.
- [11] Altinkemer, K., and Wang, T. Cost and benefit analysis of authentication systems. *Decision Support Systems* 51, 3 (2011), 394–404. doi:10.1016/j.dss.2011.01.005.
- [12] American Psychological Association. Ethical Principles of Psychologists and Code of Conduct. Website, 2016. Online available from <http://www.apa.org/ethics/code/index.aspx>; last accessed 14 September 2020.
- [13] Andics, A., McQueen, J. M., Petersson, K. M., Gál, V., Rudas, G., and Vidnyánszky, Z. Neural mechanisms for voice recognition. *Neuroimage* 52, 4 (2010), 1528–1540. doi:10.1016/j.neuroimage.2010.05.048.
- [14] Askarin, M. M., Wong, K., and Phan, R. C.-W. Planting attack on latent fingerprints. *IET Biometrics* 7, 5 (2017), 396–404. doi:10.1049/iet-bmt.2016.0113.
- [15] Associated Press. Fake Speed Bumps Create Optical Illusion, Driver Confusion. Website, June 2008. Online available from <http://www.foxnews.com/story/2008/06/27/fake-speed-bumps-create-optical-illusion-driver-confusion.html>; last accessed 14 September 2020.
- [16] Aurigemma, S., Mattson, T., and Leonard, L. So much promise, so little use: What is stopping home end-users from using password manager applications? In *Proceedings of the 50th Hawaii*

International Conference on System Sciences (HICSS 2017), Waikoloa Village, HI, USA (Atlanta, GA, USA, 2017), Association for Information Systems, pp. 4061–4070. isbn:978-0-9981331-0-2.

- [17] Ayyagari, R., Lim, J., and Hoxha, O. Why Do Not We Use Password Managers? A Study on the Intention to Use Password Managers. *Contemporary Management Research* 15, 4 (2019), 227–245. doi:10.7903/cmr.19394.
- [18] Bacchi, C. *Analysing Policy: What's the problem represented to be?* Pearson Higher Education Australia, 2009. isbn:978-0733985751.
- [19] Bacchi, C., et al. Why study problematizations? Making politics visible. *Open Journal of Political Science* 2, 01 (2012), 1–8. doi:10.4236/ojps.2012.21001.
- [20] Bachmann, M. Passwords are dead: alternative authentication methods. In *Proceedings of the 2014 IEEE Joint Intelligence and Security Informatics Conference (ISI 2014)*, The Hague, Netherlands (New York, NY, USA, 2014), IEEE, pp. 322–322. doi:10.1109/JISIC.2014.67.
- [21] Bateson, M., Callow, L., Holmes, J. R., Roche, M. L. R., and Nettle, D. Do images of ‘watching eyes’ induce behaviour that is more pro-social or more normative? A field experiment on littering. *PLoS One* 8, 12 (2013), e82055. doi:10.1371/journal.pone.0082055.
- [22] Behavioural Insights Team. Behavioural Insights Team annual update 2010–11, 2011. Online available from <https://www.gov.uk/government/publications/behavioural-insights-team-annual-update>; last accessed 13 September 2020.
- [23] Ben-Asher, N., Kirschnick, N., Sieger, H., Meyer, J., Ben-Oved, A., and Möller, S. On the need for different security methods on mobile phones. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (Mobile HCI '11)*, Stockholm, Sweden (New York, NY, USA, 2011), ACM, pp. 465–473. doi:10.1145/2037373.2037442.
- [24] Ben-Shahar, O., and Schneider, C. The Failure of Mandated Disclosure. *University of Pennsylvania Law Review* 159, 3 (2011), 647–749.
- [25] Benjamini, Y., and Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 57, 1 (1995), 289–300. doi:10.2307/2346101.
- [26] Beynon-Davies, P. Personal identity management and electronic government: The case of the national identity card in the UK. *Journal of Enterprise Information Management* 20, 3 (2007), 244–270. doi:10.1108/17410390710740727.
- [27] Bhagavatula, R., Ur, B., Iacovino, K., Kywe, S. M., Cranor, L. F., and Savvides, M. Biometric authentication on iphone and android: Usability, perceptions, and influences on adoption. In *Proceedings of the Workshop on Usable Security (USEC'15)*, San Diego, CA, USA (Reston, VA, USA, 2015), Internet Society. doi:10.14722/usec.2015.23003.
- [28] Bhattacharyya, D., Ranjan, R., Alisherov, F., and Choi, M. Biometric Authentication: A Review. *International Journal of u-and e-Service, Science and Technology* 2, 3 (2009), 13–27.
- [29] Blamey, A., Mutrie, N., and Tom, A. Health promotion by encouraged use of stairs. *British Medical Journal* 311, 7000 (1995), 289–290. doi:10.1136/bmj.311.7000.289.
- [30] Blonder, G. E. Graphical password, Sept. 24 1996. US Patent 5,559,961.
- [31] Blumenthal-Barby, and Burroughs, H. Seeking Better Health Care Outcomes: The Ethics of Using the “Nudge”. *The American Journal of Bioethics* 12, 2 (2012), 1–10. doi:10.1080/15265161.2011.634481.
- [32] Blumenthal-Barby, J., and Naik, A. D. In defense of nudge–autonomy compatibility. *The American Journal of Bioethics* 15, 10 (2015), 45–47. doi:10.1080/15265161.2015.1074304.
- [33] Bonneau, J., Bursztein, E., Caron, I., Jackson, R., and Williamson, M. Secrets, Lies, and Account Recovery: Lessons from the Use of Personal Knowledge Questions at Google. In *Proceedings of the 24th International Conference on World Wide Web (WWW '15)*, Florence, Italy (New York, NY, USA, 2015), ACM, pp. 141–150. doi:10.1145/2736277.2741691.
- [34] Bonneau, J., Herley, C., Van Oorschot, P. C., and Stajano, F. The quest to replace passwords: A framework for comparative evaluation of web authentication schemes. In *Proceedings of the 33rd*

-
- IEEE Symposium on Security and Privacy (S&P'12), San Francisco, CA, USA* (New York, NY, USA, 2012), IEEE, pp. 553–567. doi:10.1109/SP.2012.44.
- [35] Borgen, H., Bours, P., and Wolthusen, S. D. Visible-spectrum biometric retina recognition. In *Proceedings of the 2008 International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP 2008), Harbin, China* (New York, NY, USA, 2008), IEEE, pp. 1056–1062. doi:10.1109/IIH-MSP.2008.345.
- [36] Bošnjak, L., and Brumen, B. Rejecting the death of passwords: Advice for the future. *Computer Science and Information Systems* 16, 1 (2019), 313–332. doi:10.2298/CSIS180328016B.
- [37] Brewer, M. B. Ingroup identification and intergroup conflict: When does ingroup love become outgroup hate? In *Rutgers series on self and social identity; Vol. 3. Social identity, intergroup conflict, and conflict reduction*, R. D. Ashmore, L. Jussim, and D. Wilder, Eds. Oxford University Press, Oxford, UK, 2001, pp. 17–41.
- [38] Brooke, J. SUS: a 'quick and dirty' usability scale. In *Usability evaluation in industry*, P. W. Jordan, B. Thomas, I. L. McClelland, and B. Weerdmeester, Eds. Taylor & Francis, 1996, pp. 189–194. doi:10.1201/9781498710411-35.
- [39] Brooks, T. Should we nudge informed consent? *The American Journal of Bioethics* 13, 6 (2013), 22–23. doi:10.1080/15265161.2013.781710.
- [40] Brostoff, S., and Sasse, M. A. Are Passfaces More Usable Than Passwords? A Field Trial Investigation. In *People and computers XIV—usability or else!* Springer, London, United Kingdom, 2000, pp. 405–424. doi:10.1007/978-1-4471-0515-2_27.
- [41] Brown, P. A nudge in the right direction? Towards a sociological engagement with libertarian paternalism. *Social Policy and Society* 11, 3 (2012), 305–317. doi: 10.1017/S1474746412000061.
- [42] Buchoux, A., and Clarke, N. L. Deployment of keystroke analysis on a smartphone. In *Proceedings of the 6th Australian Information Security Management Conference, Perth, Australia* (2008), Research Online, pp. 1–8. doi:10.4225/75/57b55a56b876a.
- [43] Calo, R. Code, Nudge, or Notice? *Iowa Law Review* 99, 2 (2014), 773–802.
- [44] Calzolari, G., and Nardotto, M. Nudging with information: a randomized field experiment on reminders and feedback. *CEPR Discussion Paper No. DP8571* (2011).
- [45] Cao, K., and Jain, A. K. Hacking mobile phones using 2D printed fingerprints. *Michigan State University, Technical Report MSU-CSE-16-2* (2016).
- [46] Caraban, A., Karapanos, E., Gonçalves, D., and Campos, P. 23 Ways to Nudge: A Review of Technology-Mediated Nudging in Human-Computer Interaction. In *Proceedings of the 37th ACM Conference on Human Factors in Computing Systems (CHI '19), Glasgow, UK* (New York, NY, USA, 2019), ACM, pp. 1–15. doi:10.1145/3290605.3300733.
- [47] Castano, E., Yzerbyt, V., Paladino, M.-P., and Sacchi, S. I Belong, therefore, I Exist: Ingroup Identification, Ingroup Entitativity, and Ingroup Bias. *Personality and Social Psychology Bulletin* 28, 2 (2002), 135–143. doi:10.1177/0146167202282001.
- [48] Castelli, C., Gabriel, B., Yates, J., and Booth, P. Strengthening digital society against cyber shocks — Key findings from The Global State of Information Security Survey 2018, 2018. Online available from <https://www.pwc.com/us/en/services/consulting/cybersecurity/library/information-security-survey/strengthening-digital-society-against-cyber-shocks.html>, last accessed 14 September 2020.
- [49] Cherapau, I., Muslukhov, I., Asanka, N., and Beznosov, K. On the Impact of Touch ID on iPhone Passcodes. In *Proceedings of the 11th Symposium On Usable Privacy and Security (SOUPS '15), Ottawa, Canada* (Berkeley, CA, USA, 2015), USENIX Association, pp. 257–276.
- [50] Cheswick, W. Rethinking Passwords. *Communications of the ACM* 56, 2 (2013), 40–44. doi:10.1145/2408776.2408790.
- [51] Chiasson, S., Forget, A., Biddle, R., and Van Oorschot, P. C. Influencing users towards better passwords: persuasive cued click-points. In *Proceedings of the 22nd British HCI Group Annual*

- [52] Chiasson, S., Stobert, E., Forget, A., Biddle, R., and Van Oorschot, P. C. Persuasive Cued Click-Points: Design, Implementation, and Evaluation of a Knowledge-Based Authentication Mechanism. *IEEE Transactions on Dependable and Secure Computing* 9, 2 (2011), 222–235. doi:10.1109/TDSC.2011.55.
- [53] Chiasson, S., and Van Oorschot, P. C. Quantifying the security advantage of password expiration policies. *Designs, Codes and Cryptography* 77, 2-3 (2015), 401–408. doi:10.1007/s10623-015-0071-9.
- [54] Chiasson, S., Van Oorschot, P. C., and Biddle, R. Graphical password authentication using cued click points. In *Proceedings of the 12th European Symposium on Research in Computer Security (ESORICS 2007), Dresden, Germany* (Berlin/Heidelberg, Germany, 2007), J. Biskup and J. López, Eds., Springer, pp. 359–374. doi:10.1007/978-3-540-74835-9_24.
- [55] Choe, E. K., Jung, J., Lee, B., and Fisher, K. Nudging People Away from Privacy-Invasive Mobile Apps through Visual Framing. In *Proceedings of the IFIP Conference on Human-Computer Interaction (INTERACT 2013), Cape Town, South Africa* (Berlin/Heidelberg, Germany, 2013), P. Kotzé, G. Marsden, G. Lindgaard, J. Wesson, and M. Winckler, Eds., Springer, pp. 74–91. doi:10.1007/978-3-642-40477-1_5.
- [56] Cialdini, R. B., Reno, R. R., and Kallgren, C. A. A focus theory of normative conduct: recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology* 58, 6 (1990), 1015–1026. doi:10.1037/0022-3514.58.6.1015.
- [57] Cialdini, R. B., and Trost, M. R. Social influence: Social norms, conformity and compliance. In *The handbook of social psychology*, D. T. Gilbert, S. T. Fiske, and G. Lindzey, Eds. McGraw-Hill, 1998, pp. 151–192. isbn:0470137487.
- [58] Cisco. Cisco 2018 Annual Cybersecurity Report, 2018. Online available from https://www.cisco.com/c/m/en_au/products/security/offers/annual-cybersecurity-report-2018.html, last accessed 14 September 2020.
- [59] Clark, R. L., Maki, J. A., and Morrill, M. S. Can simple informational nudges increase employee participation in a 401 (k) plan? *Southern Economic Journal* 80, 3 (2014), 677–701. doi:0.4284/0038-4038-2012.199.
- [60] Corbató, F. J. On building systems that will fail. *Communications of the ACM* 34, 9 (1991), 72–81.
- [61] Corbató, F. J., Merwin-Daggett, M., and Daley, R. C. An experimental time-sharing system. In *Proceedings of the Spring Joint Computer conference, San Francisco, California* (New York, NY, USA, 1962), ACM, pp. 335–344.
- [62] Council of European Union. Directive (EU) 2015/2366 of the European Parliament and of the Council of 25 November 2015 on payment services in the internal market. Website, 2015. Online available from <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32015L2366>; last accessed 14 September 2020.
- [63] Coventry, L., De Angeli, A., and Johnson, G. Honest it’s me! Self service verification. In *Proceedings of the Workshop on Human-Computer Interaction and Security Systems, part of CHI '03, Fort Lauderdale, FL, USA* (New York, NY, USA, 2003), ACM, pp. 1–4.
- [64] Cranor, L. Federal Trade Commission: Time to rethink mandatory password changes. Website, 2016. Online available from <https://www.ftc.gov/news-events/blogs/techftc/2016/03/time-rethink-mandatory-password-changes>; last accessed 10 September 2020.
- [65] Dahlbäck, N., Jönsson, A., and Ahrenberg, L. Wizard of Oz studies—why and how. *Knowledge-Based Systems* 6, 4 (1993), 258–266. doi:10.1016/0950-7051(93)90017-N.
- [66] Davis, F. D. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS quarterly* 13, 3 (1989), 319–340. doi:10.2307/249008.
- [67] De Angeli, A., Coutts, M., Coventry, L., Johnson, G. I., Cameron, D., and Fischer, M. H. VIP: a visual approach to user authentication. In *Proceedings of the Working Conference on Ad-*

vanced Visual Interfaces (AVI '02), Trento, Italy (New York, NY, USA, 2002), ACM, pp. 316–323. doi:10.1145/1556262.1556312.

- [68] De Luca, A., Hang, A., Von Zezschwitz, E., and Hussmann, H. I Feel Like I'm Taking Selfies All Day!: Towards Understanding Biometric Authentication on Smartphones. In *Proceedings of the 33rd ACM Conference on Human Factors in Computing Systems (CHI '15), Seoul, South Korea* (New York, NY, USA, 2015), ACM, pp. 1411–1414. doi:10.1145/2702123.2702141.
- [69] Dekker, S. *Just Culture: Balancing Safety and Accountability*. CRC Press, Boca Raton, FL, USA, 2012. isbn: 978-1409440611.
- [70] Dekker, S. *Safety Differently: Human Factors for a New Era*. CRC Press, Boca Raton, FL, USA, 2014. isbn:9781482241990.
- [71] Dekker, S. *The Field Guide to Understanding 'Human Error'*. CRC Press, Boca Raton, FL, USA, 2014. isbn:978-1472439055.
- [72] Dekker, S. Why do things go right? Website, September 2018. Online available from <http://www.safetydifferently.com/why-do-things-go-right/>; last accessed 14 September 2020.
- [73] Dekker, S., Cilliers, P., and Hofmeyr, J.-H. The complexity of failure: Implications of complexity theory for safety investigations. *Safety Science* 49, 6 (2011), 939–945. doi:10.1016/j.ssci.2011.01.008.
- [74] Department of Health Education and Welfare. The Belmont Report - Ethical Principles and Guidelines for the Protection of Human Subjects of Research. Website, 1979. Online available from <https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/#xrespect>; last accessed 14 September 2020.
- [75] Dhamija, R., Perrig, A., et al. Deja Vu-A User Study: Using Images for Authentication. In *Proceedings of the 9th USENIX Security Symposium (USENIX Security '00), Denver, CO, USA* (Berkeley, CA, USA, 2000), vol. 9, USENIX Association, pp. 1–4.
- [76] Dharavath, K., Talukdar, F., and Laskar, R. Study on biometric authentication systems, challenges and future trends: A review. In *Proceedings of the 2013 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC 2013), Madurai, India* (New York, NY, USA, 2013), IEEE, pp. 1–7. doi:10.1109/ICCIC.2013.6724278.
- [77] Dictionary, M.-W. Merriam-Webster. Website, 2020. Online available from <https://www.merriam-webster.com/>; last accessed 14 September 2020.
- [78] Dirik, A. E., Memon, N., and Birget, J.-C. Modeling user choice in the PassPoints graphical password scheme. In *Proceedings of the 3rd Symposium on Usable Privacy and Security (SOUPS '07), Pittsburgh, PA, USA* (New York, NY, USA, 2007), ACM, pp. 20–28. doi:10.1145/1280680.1280684.
- [79] DiSilvestro, R. What does not budge for any nudge? *American Journal of Bioethics* 12, 2 (2012), 14–15. doi:10.1080/15265161.2011.634956.
- [80] Dolan, P., Hallsworth, M., Halpern, D., King, D., and Vlaev, I. MINDSPACE: influencing behaviour for public policy, 2010. Online available from <https://www.instituteforgovernment.org.uk/publications>; last accessed 13 September 2020.
- [81] Dörflinger, T., Voth, A., Krämer, J., and Fromm, R. “My smartphone is a safe!” The user's point of view regarding novel authentication methods and gradual security levels on smartphones. In *Proceedings of the 2010 International Conference on Security and Cryptography (SECRYPT), Athens, Greece* (New York, NY, USA, 2010), IEEE, pp. 1–10. isbn:978-989-8425-18-8.
- [82] Drimer, S., Murdoch, S. J., and Anderson, R. Optimised to fail: Card readers for online banking. In *Proceedings of the International Conference on Financial Cryptography and Data Security, Accra Beach, Barbados* (Berlin/Heidelberg, Germany, 2009), R. Dingledine and P. Golle, Eds., Springer, pp. 184–200. doi:10.1007/978-3-642-03549-4_11.
- [83] Egelman, S., and Peer, E. Scaling the Security Wall: Developing a Security Behavior Intentions Scale (SeBIS). In *Proceedings of the 33rd ACM Conference on Human Factors in Computing Systems (CHI '15), Seoul, South Korea* (New York, NY, USA, 2015), ACM, pp. 2873–2882. doi:10.1145/2702123.2702249.

-
- [84] Egelman, S., Sotirakopoulos, A., Muslukhov, I., Beznosov, K., and Herley, C. Does my password go up to eleven? The impact of password meters on password selection. In *Proceedings of the 31st ACM Conference on Human Factors in Computing Systems (CHI '13)*, Paris, France (New York, NY, USA, 2013), ACM, pp. 2379–2388. doi:10.1145/2470654.2481329.
- [85] Eliasson, C., Fiedler, M., and Jørstad, I. A Criteria-Based Evaluation Framework for Authentication Schemes in IMS. In *Proceedings of the International Conference on Availability, Reliability and Security (ARES 2009)*, Fukuoka, Japan (New York, NY, USA, 2009), IEEE, pp. 865–869. doi:10.1109/ARES.2009.166.
- [86] European Federation of Psychologists' Association. Meta-Code of Ethics. Website, 2005. Online available from https://www.bdp-verband.de/binaries/content/assets/beruf/efpa_metacode_en.pdf; last accessed 14 September 2020.
- [87] Evans, J. S. B. Logic and human reasoning: An assessment of the deduction paradigm. *Psychological Bulletin* 128, 6 (2002), 978–996. doi:10.1037/0033-2909.128.6.978.
- [88] Evans, J. S. B. In two minds: dual-process accounts of reasoning. *Trends in Cognitive Sciences* 7, 10 (2003), 454–459. doi:10.1016/j.tics.2003.08.012.
- [89] Evans, J. S. B., Barston, J. L., and Pollard, P. On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition* 11, 3 (1983), 295–306. doi:10.3758/BF03196976.
- [90] Evans, J. S. B., and Over, D. E. *Rationality and Reasoning*. Psychology Press, 1996.
- [91] Evans, J. S. B., and Stanovich, K. E. Dual-process theories of higher cognition: Advancing the debate. *Perspectives on psychological science* 8, 3 (2013), 223–241. doi:10.1177/1745691612460685.
- [92] Fahl, S., Harbach, M., Acar, Y., and Smith, M. On the ecological validity of a password study. In *Proceedings of the 9th Symposium on Usable Privacy and Security (SOUPS '13)*, Newcastle, UK (New York, NY, USA, 2013), ACM, pp. 1–13. doi:10.1145/2501604.2501617.
- [93] Fano, R. M., and Corbató, F. J. Time-sharing on computers. *Scientific American* 215, 3 (1966), 128–143.
- [94] Florencio, D., and Herley, C. A large-scale study of web password habits. In *Proceedings of the 16th International Conference on World Wide Web (WWW '07)*, Banff, Canada (New York, NY, USA, 2007), ACM, pp. 657–666. doi:10.1145/1242572.1242661.
- [95] Florêncio, D., Herley, C., and Van Oorschot, P. C. An administrator's guide to internet password research. In *Proceedings of the 28th Large Installation System Administration Conference (LISA '14)*, Seattle, WA, USA (Berkeley, CA, USA, 2014), USENIX Association, pp. 35–42. isbn:978-1-931971-17-1.
- [96] Florêncio, D., Herley, C., and van Oorschot, P. C. Password Portfolios and the Finite-Effort User: Sustainably Managing Large Numbers of Accounts. In *Proceedings of the 23rd USENIX Security Symposium (USENIX Security 14)*, San Diego, CA, USA (Berkeley, CA, USA, 2014), USENIX Association, pp. 575–590.
- [97] Forget, A., Chiasson, S., and Biddle, R. User-centred authentication feature framework. *Information & Computer Security* 23, 5 (2015), 497–515. doi:10.1108/ICS-08-2014-0058.
- [98] Franke, T., Attig, C., and Wessel, D. A personal resource for technology interaction: development and validation of the affinity for technology interaction (ATI) scale. *International Journal of Human-Computer Interaction* 35, 6 (2019), 456–467. doi:10.1080/10447318.2018.1456150.
- [99] Furnell, S. Assessing password guidance and enforcement on leading websites. *Computer Fraud & Security* 2011, 12 (2011), 10–18. doi:10.1016/S1361-3723(11)70123-3.
- [100] Furnell, S., Alotaibi, F., and Esmael, R. Aligning Security Practice with Policy: Guiding and Nudging towards Better Behavior. In *Proceedings of the 52nd Hawaii International Conference on System Sciences (HICSS '19)*, Grand Wailea, HI, USA (Honolulu, HI, USA, 2019), University of Hawai'i at Manoa, pp. 5618–5627. doi:10.24251/HICSS.2019.676.
- [101] Furnell, S., and Evangelatos, K. Public awareness and perceptions of biometrics. *Computer Fraud & Security* 2007, 1 (2007), 8–13. doi:10.1016/S1361-3723(07)70006-4.

-
- [102] Furnell, S., Khern-am nuai, W., Esmael, R., Yang, W., and Li, N. Enhancing security behaviour by supporting the user. *Computers & Security* 75 (2018), 1–9. doi:10.1016/j.cose.2018.01.016.
- [103] Furnell, S. M., Dowland, P., Illingworth, H., and Reynolds, P. L. Authentication and Supervision: A Survey of User Attitudes. *Computers & Security* 19, 6 (2000), 529–539. doi:10.1016/S0167-4048(00)06027-2.
- [104] Gemalto. GrIDSure. Website, 2006 - 2019. Online available from <https://www3.thalesgroup.com/sas/grid-tokens.html>; last accessed 01 September 2020.
- [105] Gerber, N., and Zimmermann, V. Security vs. privacy? User preferences regarding text passwords and biometric authentication. In *Proceedings of the Mensch und Computer 2017-Workshopband, Regensburg, Germany* (Bonn, Germany, 2017), M. Burghardt, R. Wimmer, C. Wolff, and C. Womser-Hacker, Eds., Gesellschaft für Informatik e.V., pp. 279–287. doi: 10.18420/muc2017-ws05-0405.
- [106] Gibson, M., Renaud, K., Conrad, M., and Maple, C. Musipass: authenticating me softly with my song. In *Proceedings of the 2009 Workshop on New Security Paradigms Workshop (NSPW '09)*, Oxford, UK (New York, NY, USA, 2009), ACM, pp. 85–100. doi:10.1145/1719030.1719043.
- [107] Gigerenzer, G., and Gaissmaier, W. Heuristic decision making. *Annual Review of Psychology* 62 (2011), 451–482. doi:10.1146/annurev-psych-120709-145346.
- [108] Gilbert, D. T. Thinking lightly about others: Automatic components of the social inference process. In *Unintended thought*, J. S. Uleman and J. A. Bargh, Eds. The Guilford Press, New York, NY, USA, 1989, pp. 189–211. isbn:9780898623796.
- [109] Goel, V., Buchel, C., Frith, C., and Dolan, R. J. Dissociation of mechanisms underlying syllogistic reasoning. *Neuroimage* 12, 5 (2000), 504–514. doi:10.1006/nimg.2000.0636.
- [110] Goel, V., and Dolan, R. J. Explaining modulation of reasoning by belief. *Cognition* 87, 1 (2003), B11–B22. doi:10.1016/S0010-0277(02)00185-3.
- [111] Golla, M., and Dürmuth, M. On the accuracy of password strength meters. In *Proceedings of the 2018 ACM Conference on Computer and Communications Security (CCS '18)*, Toronto, Canada (New York, NY, USA, 2018), ACM, pp. 1567–1582. doi:10.1145/3243734.3243769.
- [112] Golla, M., Hahn, B., zu Selhausen, K. M., Hosseini, H., and Dürmuth, M. Bars, Badges, and High Scores: On the Impact of Password Strength Visualizations. In *Proceedings of Who Are You?! Adventures in Authentication Workshop (WAY 2018)*, Baltimore, MD, USA (Berkeley, CA, USA, 2018), USENIX Association, pp. 1–7.
- [113] Golle, P., and Wagner, D. Cryptanalysis of a Cognitive Authentication Scheme (Extended Abstract). In *Proceedings of the 28th IEEE Symposium on Security and Privacy (S&P'07)*, Oakland, CA, USA (New York, NY, USA, 2007), IEEE, pp. 66–70. doi:10.1109/SP.2007.13.
- [114] Google Inc. Touch Gesture Actions From A Device's Lock Screen. Patent, 2011. US Patent No. 20110283241.
- [115] Grassi, P. A., Perlner, R. A., Newton, E. M., Regenscheid, A. R., Burr, W. E., Richer, J. P., Lefkovitz, N. B., Danker, J. M., and Theofanos, M. F. Special Publication (NIST SP) - 800-63B: Digital identity guidelines: Authentication and lifecycle management. Tech. rep., National Institute of Standards and Technology, 2017. doi:10.6028/NIST.SP.800-63b.
- [116] Grawemeyer, B., and Johnson, H. Using and managing multiple passwords: A week to a view. *Interacting with Computers* 23, 3 (2011), 256–267. doi:10.1016/j.intcom.2011.03.007.
- [117] Groff, B. K. Optical illusion speed bump and method of using the same, May 16 2006. US Patent 7,044,679.
- [118] Guel, M. D. A Framework for Choosing Your Next Generation Authentication/Authorization System. *Information Security Technical Report* 7, 1 (2002), 63–78. doi:10.1016/S1363-4127(02)00107-3.
- [119] Habib, H., Naeini, P. E., Devlin, S., Oates, M., Swoopes, C., Bauer, L., Christin, N., and Cranor, L. F. User Behaviors and Attitudes Under Password Expiration Policies. In *Proceedings of the 14th Symposium on Usable Privacy and Security (SOUPS '18)*, Baltimore, MD, USA (Berkeley, CA, USA, 2018), USENIX Association, pp. 13–30.

-
- [120] Haller, N., Metz, C., Nesser, P., and Straw, M. RFC 2289: A one-time password system. *Network Working Group Request for Comments* (1998). doi:10.17487/RFC2289.
- [121] Han, J., and Bhanu, B. Individual recognition using gait energy image. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28, 2 (2006), 316–322. doi:10.1109/TPAMI.2006.38.
- [122] Hansen, P. G. The definition of nudge and libertarian paternalism: Does the hand fit the glove? *European Journal of Risk Regulation* 7, 1 (2016), 155–174. doi:10.1017/S1867299X00005468.
- [123] Hansen, P. G., and Jespersen, A. M. Nudge and the Manipulation of Choice: A Framework for the Responsible Use of the Nudge Approach to Behaviour Change in Public Policy. *European Journal of Risk Regulation* 4, 1 (2013), 3–28. doi:https://www.jstor.org/stable/24323381.
- [124] Harris, J. Time to make up your mind: why choosing is difficult. *British Journal of Learning Disabilities* 31, 1 (2003), 3–8. doi:10.1046/j.1468-3156.2003.00181.x.
- [125] Hart, W. *Verdraaide Organisaties*. Vakmedianet, 2013. isbn: 978-9013105735.
- [126] Haselton, M. G., Nettle, D., and Andrews, P. W. The evolution of cognitive bias. In *The Handbook of Evolutionary Psychology*, D. M. Buss, Ed. Wiley Online Library, 2015, pp. 1–20. doi:10.1002/9780470939376.ch25.
- [127] Hassenzahl, M., Burmester, M., and Koller, F. AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In *Proceedings of the Mensch & Computer 2003, Stuttgart, Germany* (Wiesbaden, Germany, 2003), S. G. and Z. J., Eds., Vieweg+Teubner Verlag, pp. 187–196. doi:10.1007/978-3-322-80058-9_19.
- [128] Hatekar, Abhijeet and others. Microsoft Security Intelligence Report - Volume 23. Website, 2018. Online available from <https://www.microsoft.com/en-us/security/business/security-intelligence-report>, last accessed 14 September 2020.
- [129] Hausman, D. M., and Welch, B. Debate: To Nudge or Not to Nudge. *The Journal of Political Philosophy* 18, 1 (2010), 123–136. doi: 10.1111/j.1467-9760.2009.00351.x.
- [130] Hayashi, E., Dhamija, R., Christin, N., and Perrig, A. Use your illusion: secure authentication usable anywhere. In *Proceedings of the 4th Symposium on Usable Privacy and Security (SOUPS '08), Pittsburgh, PA, USA* (New York, NY, USA, 2008), ACM, pp. 35–45. doi:10.1145/1408664.1408670.
- [131] Heinrich, H. W. *Industrial Accident Prevention. A Scientific Approach*, fourth ed. McGraw-Hill Book Company, Inc., New York & London, 1959.
- [132] Herley, C. So Long, and No Thanks for the Externalities: The Rational Rejection of Security Advice by Users. In *Proceedings of the 2009 New Security Paradigms Workshop (NSPW '09), Oxford, UK* (New York, NY, USA, 2009), ACM, pp. 133–144.
- [133] Herley, C., and Van Oorschot, P. A Research Agenda Acknowledging the Persistence of Passwords. *IEEE Security & Privacy* 10, 1 (2011), 28–36. doi:10.1109/MSP.2011.150.
- [134] Hermans, D., Houwer, J. D., and Eelen, P. The affective priming effect: Automatic activation of evaluative information in memory. *Cognition & Emotion* 8, 6 (1994), 515–533. doi:10.1080/02699939408408957.
- [135] Hermans, J., and Peeters, R. Realizing Pico: Finally No More Passwords! *IACR Cryptology ePrint Archive 2014* (2014), 1–19.
- [136] Hirano, M., Takeuchi, M., Tomoda, T., and Nakano, K.-I. Keyless entry system with radio card transponder (automobiles). *IEEE Transactions on Industrial Electronics* 35, 2 (1988), 208–216. doi:10.1109/41.192651.
- [137] Hollnagel, E. *Safety-I and Safety-II: The Past and Future of Safety Management*. CRC Press, Boca Raton, FL, USA, 2014. isbn: 978-1472423054.
- [138] Hollnagel, E., Woods, D., and Leveson, N. *Resilience Engineering: Concepts and Precepts*. Ashgate Publishing, Aldershot, UK, 2006. isbn: 978-0754646419.
- [139] Holz, C., and Bentley, F. R. On-demand biometrics: Fast cross-device authentication. In *Proceedings of the 34th ACM Conference on Human Factors in Computing Systems (CHI '16), San Jose, CA, USA* (New York, NY, USA, 2016), ACM, pp. 3761–3766. doi:10.1145/2858036.2858139.

-
- [140] Huang, D.-L., Rau, P.-L. P., Salvendy, G., Gao, F., and Zhou, J. Factors affecting perception of information security and their impacts on IT adoption and security practices. *International Journal of Human-Computer Studies* 69, 12 (2011), 870–883. doi:10.1016/j.ijhcs.2011.07.007.
- [141] Huh, J. H., Kim, H., Bobba, R. B., Bashir, M. N., and Beznosov, K. On the memorability of system-generated pins: Can chunking help? In *Proceedings of the 11th Symposium On Usable Privacy and Security (SOUPS '15), Ottawa, Canada* (Berkeley, CA, USA, 2015), USENIX Association, pp. 197–209. isbn:9781931971249.
- [142] Inglesant, P. G., and Sasse, M. A. The true cost of unusable password policies: password use in the wild. In *Proceedings of the 28th ACM Conference on Human Factors in Computing Systems (CHI '10), Atlanta, GA, USA* (New York, NY, USA, 2010), ACM, pp. 383–392. doi:10.1145/1753326.1753384.
- [143] Irakleous, I., Furnell, S. M., Dowland, P. S., and Papadaki, M. An experimental comparison of secret-based user authentication technologies. *Information Management & Computer Security* 10, 3 (2002), 100–108. doi:10.1108/09685220210431854.
- [144] Jain, A. K., Bolle, R., and Pankanti, S., Eds. *Biometrics: Personal Identification in Networked Society*, vol. 479. Springer US, 2006. doi:10.1007/978-0-387-32659-7.
- [145] Jain, A. K., Ross, A., and Prabhakar, S. An introduction to biometric recognition. *IEEE Transactions on circuits and systems for video technology* 14, 1 (2004). doi:10.1007/978-0-387-32659-7.
- [146] Jakobsen, M., and Serritzlew, S. Effects on knowledge of nudging citizens with information. *International Journal of Public Administration* 39, 6 (2016), 449–458. doi:10.1080/01900692.2015.1020550.
- [147] Jakobsson, M., Yang, L., and Wetzel, S. Quantifying the security of preference-based authentication. In *Proceedings of the 4th ACM workshop on Digital Identity Management (DIM '08), Alexandria, VA, USA* (New York, NY, USA, 2008), ACM, pp. 61–70. doi:10.1145/1456424.1456435.
- [148] Jansen, W., Gavrilu, S. I., Korolev, V., Ayers, R. P., and Swanstrom, R. Picture password: a visual login technique for mobile devices. Tech. rep., National Institute of Standards and Technology, Gaithersburg, MD, USA), 2003. doi:10.6028/NIST.IR.7030.
- [149] Jeffreys, A. J., Wilson, V., and Thein, S. L. Individual-specific ‘fingerprints’ of human DNA. *Nature* 316, 6023 (1985), 76–79. doi:10.1038/316076a0.
- [150] Jermyn, I., Mayer, A., Monroe, F., Reiter, M. K., and Rubin, A. D. The design and analysis of graphical passwords. In *Proceedings of the 8th USENIX Security Symposium (USENIX Security '99) (Washington, D.C., USA)*. USENIX Association, Berkeley, CA, USA, 1999.
- [151] John, P., Cotterill, S., Richardson, L., Moseley, A., Smith, G., Stoker, G., Wales, C., Liu, H., and Nomura, H. *Nudge, nudge, think, think: Experimenting with ways to change civic behaviour*. Bloomsbury, London, UK, 2013. isbn:9781780935553.
- [152] Jones, L. A., Antón, A. I., and Earp, J. B. Towards understanding user perceptions of authentication technologies. In *Proceedings of the 2007 ACM Workshop on Privacy in Electronic Society (WPES '07), Alexandria, VA, USA* (New York, NY, USA, 2007), ACM, pp. 91–98. doi:10.1145/1314333.1314352.
- [153] Just, M. Designing and evaluating challenge-question systems. *IEEE Security & Privacy* 2, 5 (2004), 32–39. doi:10.1109/MSP.2004.80.
- [154] Just, M., and Aspinall, D. Personal choice and challenge questions: a security and usability assessment. In *Proceedings of the 5th Symposium on Usable Privacy and Security (SOUPS '09), Mountain View, CA, USA* (New York, NY, USA, 2009), ACM, p. 8. doi:10.1145/1572532.1572543.
- [155] Kahneman, D. Maps of bounded rationality: A perspective on intuitive judgment and choice (Nobel prize lecture). Website, 2002. Online available from <https://www.nobelprize.org/prizes/economic-sciences/2002/kahneman/lecture/>; last accessed 14 September 2020.
- [156] Kahneman, D. *Thinking, Fast and Slow*. Farrar, Straus and Giroux, 2013. isbn:0374533555.
- [157] Kaleta, J. P., Lee, J. S., and Yoo, S. Nudging with construal level theory to improve online password use and intended password choice: A security-usability tradeoff perspective. *Information Technology & People* 32, 4 (2019), 993–1020. doi:10.1108/ITP-01-2018-0001.

-
- [158] Kankane, S., DiRusso, C., and Buckley, C. Can we nudge users toward better password management?: An initial study. In *Proceedings of the 36th ACM Conference on Human Factors in Computing Systems - Extended Abstracts (CHI EA '18), Montreal, Canada* (New York, NY, USA, 2018), ACM, pp. 1–6. doi:10.1145/3170427.3188689.
- [159] Karnan, M., Akila, M., and Krishnaraj, N. Biometric personal authentication using keystroke dynamics: A review. *Applied soft computing* 11, 2 (2011), 1565–1573. doi:10.1016/j.asoc.2010.08.003.
- [160] Katsini, C., Belk, M., Fidas, C., Avouris, N., and Samaras, G. Security and Usability in Knowledge-based User Authentication: A Review. In *Proceedings of the 20th Pan-Hellenic Conference on Informatics (PCI '16), Patras, Greece* (New York, NY, USA, 2016), ACM, pp. 1–6. doi:10.1145/3003733.3003764.
- [161] Keith, M., Shao, B., and Steinbart, P. A Behavioral Analysis of Passphrase Design and Effectiveness. *Journal of the Association for Information Systems* 10, 2 (2009), 63–89. doi:10.17705/1jais.00184.
- [162] Keith, M., Shao, B., and Steinbart, P. J. The usability of passphrases for authentication: An empirical field study. *International journal of Human-Computer Studies* 65, 1 (2007), 17–28. doi:10.1016/j.ijhcs.2006.08.005.
- [163] Kelley, P. G., Komanduri, S., Mazurek, M. L., Shay, R., Vidas, T., Bauer, L., Christin, N., Cranor, L. F., and Lopez, J. Guess again (and again and again): Measuring password strength by simulating password-cracking algorithms. In *Proceedings of the 33rd IEEE Symposium on Security and Privacy (S&P'12), San Francisco, CA, USA* (New York, NY, USA, 2012), IEEE, pp. 523–537. doi:10.1109/SP.2012.38.
- [164] Kelly, D., and Morar, N. Nudging and the ecological and social roots of human agency. *The American Journal of Bioethics* 16, 11 (2016), 15–17. doi:10.1080/15265161.2016.1222018.
- [165] Kerr, J., Eves, F., and Carroll, D. Encouraging Stair Use: Stair-Riser Banners Are Better Than Posters. *American Journal of Public Health* 91, 8 (2001), 1192–1193. doi:10.1016/j.yjpm.2007.11.009.
- [166] Kim, J. Y. Efficiency of paid authentication methods for mobile devices. *Wireless Personal Communications* 93, 2 (2017), 543–551. doi:10.1007/s11277-016-3286-9.
- [167] Kim, T. H.-J., Stuart, H. C., Hsiao, H.-C., Lin, Y.-H., Zhang, L., Dabbish, L., and Kiesler, S. YourPassword: applying feedback loops to improve security behavior of managing multiple passwords. In *Proceedings of the 9th ACM Symposium on Information, Computer and Communications Security (ASIA CCS '14), Kyoto, Japan* (New York, NY, USA, 2014), ACM, pp. 513–518. doi:10.1145/2590296.2590345.
- [168] Kingston Technology Corporation. Ironkey - Mobile Security Solutions. Website, 2019. Online available from <https://www.ironkey.com/en-US/solutions/>; last accessed 06 September 2020.
- [169] Kinnunen, T., Sedlak, F., and Bednarik, R. Towards task-independent person authentication using eye movement signals. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications (ETRA '10), Austin, TX, USA* (New York, NY, USA, 2010), ACM, pp. 187–190. doi:10.1145/1743666.1743712.
- [170] Koch, F. Secure and memorable passwords? - Communicating the NIST 2017 guidelines for creating secure passwords. Master's thesis, Technische Universität Darmstadt, Darmstadt, Germany, 2017. Master's Thesis.
- [171] Komanduri, S., Shay, R., Kelley, P. G., Mazurek, M. L., Bauer, L., Christin, N., Cranor, L. F., and Egelman, S. Of passwords and people: measuring the effect of password-composition policies. In *Proceedings of the 29th ACM Conference on Human Factors in Computing Systems (CHI '11), Vancouver, Canada* (New York, NY, USA, 2011), ACM, pp. 2595–2604. doi:10.1145/1978942.1979321.
- [172] Krause, A. A. Systems for authenticating the use of transaction cards having a magnetic stripe, July 18 2000. US Patent 6,089,451.
- [173] Kroese, F. M., Marchiori, D. R., and de Ridder, D. T. Nudging healthy food choices: a field experiment at the train station. *Journal of Public Health* 38, 2 (2016), e133–e137. doi:10.1093/pubmed/fdv096.

-
- [174] Kuber, R., and Yu, W. Feasibility study of tactile-based authentication. *International Journal of Human-Computer Studies* 68, 3 (2010), 158–181. doi:10.1016/j.ijhcs.2009.11.001.
- [175] Kulkarni, D. A novel web-based approach for balancing usability and security requirements of text passwords. *International Journal of Network Security & its Applications* 2, 3 (2010), 1–16. doi:10.5121/ijnsa.2010.2301.
- [176] Kumar, A., Garg, S., and Hanmandlu, M. Biometric authentication using finger nail plates. *Expert Systems with Applications* 41, 2 (2014), 373–386. doi:10.1016/j.eswa.2013.07.057.
- [177] Kumar, A., and Prathyusha, K. V. Personal Authentication Using Hand Vein Triangulation and Knuckle Shape. *IEEE Transactions on Image processing* 18, 9 (2009), 2127–2136. doi:10.1109/TIP.2009.2023153.
- [178] Kumar, A., and Zhang, D. Hand-Geometry Recognition Using Entropy-Based Discretization. *IEEE Transactions on Information Forensics and Security* 2, 2 (2007), 181–187. doi:10.1109/TIFS.2007.896915.
- [179] Kumar, N. User Authentication using Musical Password. *International Journal of Computer Applications* 59, 9 (2012), 1–4.
- [180] Kuo, C., Romanosky, S., and Cranor, L. F. Human selection of mnemonic phrase-based passwords. In *Proceedings of the 2nd Symposium on Usable Privacy and Security (SOUPS '06), Pittsburgh, PA, USA* (New York, NY, USA, 2006), ACM, pp. 67–78. doi:10.1145/1143120.1143129.
- [181] Lester, R. Electronic recognition door lock, May 22 1973. US Patent 3,733,861.
- [182] Li, Z., He, W., Akhawe, D., and Song, D. The emperor’s new password manager: Security analysis of web-based password managers. In *Proceedings of the 23rd USENIX Security Symposium (USENIX Security '14), San Diego, CA, USA* (Berkeley, CA, USA, 2014), USENIX Association, pp. 465–479.
- [183] Lin, Y., Osman, M., and Ashcroft, R. Nudge: concept, effectiveness, and ethics. *Basic and Applied Social Psychology* 39, 6 (2017), 293–306. doi:10.1080/01973533.2017.1356304.
- [184] Lindhout, P., and Reniers, G. What about nudges in the process industry? Exploring a new safety management tool. *Journal of Loss Prevention in the Process Industries* 50, Part A (2017), 243–256. doi:10.1016/j.jlp.2017.10.006.
- [185] Loeber, S., Vollstädt-Klein, S., Wilden, S., Schneider, S., Rockenbach, C., Dinter, C., von der Goltz, C., Hermann, D., Wagner, M., Winterer, G., and Kiefer, F. The effect of pictorial warnings on cigarette packages on attentional bias of smokers. *Pharmacology Biochemistry and Behavior* 98, 2 (2011), 292–298. doi:10.1016/j.pbb.2011.01.010.
- [186] Lyastani, S. G., Schilling, M., Fahl, S., Backes, M., and Bugiel, S. Better managed than memorized? Studying the Impact of Managers on Password Strength and Reuse. In *Proceedings of the 27th USENIX Security Symposium (USENIX Security 18), Baltimore, MD, USA* (Berkeley, CA, USA, 2018), USENIX Association, pp. 203–220.
- [187] Ma, Y., and Feng, J. Evaluating Usability of Three Authentication Methods in Web-Based Application. In *Proceedings of the 9th International Conference on Software Engineering Research, Management and Applications (SERA 2011), Baltimore, MD, USA* (New York, NY, USA, 2011), IEEE, pp. 81–88. doi:10.1109/SERA.2011.18.
- [188] Maetz, Y., Onno, S., and Heen, O. Recall-A-Story, a story-telling graphical password system [Poster]. In *Proceedings of the 5th Symposium On Usable Privacy and Security (SOUPS '09), Mountain View, CA, USA* (New York, NY, USA, 2009), ACM. doi:0.1145/1572532.1572566.
- [189] Mahnken, S. Today’s authentication options: the need for adaptive multifactor authentication. *Biometric Technology Today 2014*, 7 (2014), 8–10. doi:10.1016/S0969-4765(14)70126-2.
- [190] Malhotra, N. K., Kim, S. S., and Agarwal, J. Internet users’ information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information Systems Research* 15, 4 (2004), 336–355. doi:10.2307/23015787.
- [191] Mannan, M., and Van Oorschot, P. C. Leveraging personal devices for stronger password authentication from untrusted computers. *Journal of Computer Security* 19, 4 (2011), 703–750. doi:10.3233/JCS-2010-0412.

-
- [192] Marchiori, D. R., Adriaanse, M. A., and De Ridder, D. T. Unresolved questions in nudging research: Putting the psychology back in nudging. *Social and Personality Psychology Compass* 11, 1 (2017), e12297. doi:10.1111/spc3.12297.
- [193] Mare, S., Baker, M., and Gummeson, J. A Study of Authentication in Daily Life. In *Proceedings of the 12th Symposium on Usable Privacy and Security (SOUPS '16), Denver, CO, USA* (Berkeley, CA, USA, 2016), USENIX Association, pp. 189–206.
- [194] Marky, K., Schmitz, M., Zimmermann, V., Herbers, M., Kunze, K., and Mühlhäuser, M. 3D-Auth: Two-Factor Authentication with Personalized 3D-Printed Items. In *Proceedings of the 38th ACM Conference on Human Factors in Computing Systems (CHI '20), Honolulu, HI, USA* (New York, NY, USA, 2020), ACM, pp. 1–12. doi:10.1145/3313831.3376189.
- [195] Martinez-Diaz, M., Fierrez, J., Galbally, J., and Ortega-Garcia, J. Towards mobile authentication using dynamic signature verification: useful features and performance evaluation. In *Proceedings of the 19th International Conference on Pattern Recognition (ICPR 2008), Tampa, FL, USA* (New York, NY, USA, 2008), IEEE, pp. 1–5. doi:10.1109/ICPR.2008.4761849.
- [196] Matsumoto, T., Matsumoto, H., Yamada, K., and Hoshino, S. Impact of artificial "gummy" fingers on fingerprint systems. In *Proceedings of Electronic Imaging - Optical Security and Counterfeit Deterrence Techniques IV, San Jose, CA, USA* (2002), vol. 4677, International Society for Optics and Photonics, pp. 275–289. doi:10.1117/12.462719.
- [197] Mayer, P., Neumann, S., Storck, D., and Volkamer, M. Supporting Decision Makers in Choosing Suitable Authentication Schemes. In *Proceedings of the 10th International Symposium on Human Aspects of Information Security & Assurance (HAISA 2016), Frankfurt, Germany* (Plymouth, United Kingdom, 2016), N. L. Clarke and S. Furnell, Eds., Plymouth University, pp. 67–77.
- [198] Mayer, P., Stumpf, P., Weber, T., and Volkamer, M. ACCESSv2: a collaborative authentication research and decision support platform. In *Proceedings of Who Are You?! Adventures in Authentication Workshop (WAY 2018), Baltimore, MD, USA* (Berkeley, CA, USA, 2018), USENIX Association.
- [199] Mayring, P. *Qualitative content analysis: theoretical foundation, basic procedures and software solution*. SSOAR Open Access Repository, Klagenfurth, Austria, 2014.
- [200] Mazurek, M. L., Komanduri, S., Vidas, T., Bauer, L., Christin, N., Cranor, L. F., Kelley, P. G., Shay, R., and Ur, B. Measuring password guessability for an entire university. In *Proceedings of the 2013 ACM Conference on Computer and Communications Security (CCS '13), Berlin, Germany* (New York, NY, USA, 2013), ACM, pp. 173–186. doi:10.1145/2508859.2516726.
- [201] McGrath, J. E. Methodology matters: Doing research in the behavioral and social sciences. In *Readings in Human-Computer Interaction*, R. Baecker, J. Grudin, W. A. S. Buxton, and S. Greenberg, Eds. Elsevier, 1995, pp. 152–169. doi:10.1016/B978-0-08-051574-8.50019-4.
- [202] McMillan, D., Morrison, A., and Chalmers, M. Categorised Ethical Guidelines for Large Scale Mobile HCI. In *Proceedings of the 31st ACM Conference on Human Factors in Computing Systems (CHI '13), Paris, France* (New York, NY, USA, 2013), ACM, pp. 1853–1862.
- [203] Menkus, B. Understanding the use of passwords. *Computers & Security* 7, 2 (1988), 132–136. doi:10.1016/0167-4048(88)90325-2.
- [204] Michalek, G., Meran, G., Schwarze, R., and Yildiz, Ö. Nudging as a new "soft" policy tool: An assessment of the definitional scope of nudges, practical implementation possibilities and their effectiveness. Tech. rep., Economics Discussion Papers, 2016.
- [205] Mishu, T. I., and Rahman, M. M. Vulnerabilities of Fingerprint Authentication Systems and Their Securities. *International Journal of Computer Science and Information Security (IJCSIS)* 16, 3 (2018), 99–104.
- [206] Mitchell, G. Libertarian paternalism is an oxymoron. *Northwestern University Law Review* 99, 3 (2004), 1245–1278.
- [207] Monroe, F., and Rubin, A. Authentication via keystroke dynamics. In *Proceedings of the 4th ACM Conference on Computer and Communications Security (CCS '97), Zürich, Switzerland* (New York, NY, USA, 1997), ACM, pp. 48–56. doi:10.1145/266420.266434.

-
- [208] Moody, J. Public Perceptions of Biometric Devices: The Effect of Misinformation on Acceptance and Use. *Issues in Informing Science and Information Technology* 1 (2004), 0753–0761. doi:10.28945/775.
- [209] Mulliner, C., Borgaonkar, R., Stewin, P., and Seifert, J.-P. SMS-based one-time passwords: attacks and defense. In *Proceedings of the International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment (DIMVA 2013), Berlin, Germany* (Berlin/Heidelberg, Germany, 2013), K. Rieck, P. Stewin, and J.-P. Seifert, Eds., Springer, pp. 150–159. doi:10.1007/978-3-642-39235-1_9.
- [210] Munyan, D. Coming clean on hygiene. *Biometric Technology Today* 13, 4 (2005), 7–8. doi:10.1016/S0969-4765(05)70288-5.
- [211] Murray, H., and Malone, D. Evaluating password advice. In *Proceedings of the 28th Irish Signals and Systems Conference (ISSC 2017), Killarney, Ireland* (New York, NY, USA, 2017), IEEE, pp. 1–6. doi:10.1109/ISSC.2017.7983609.
- [212] Murray, P. R. Who will nudge the nudgers. *Regulation* 40 (2017), 55.
- [213] New Zealand Department of the Prime Minister and Cabinet. National Cyber Policy Office Proactive Release. Website, 2018. Online available from <https://dpmc.govt.nz/publications/new-zealands-cyber-security-strategy>, last accessed 14 September.
- [214] Nicholson, J., Vlachokyriakos, V., Coventry, L., Briggs, P., and Olivier, P. Simple nudges for better password creation. In *Proceedings of the 32nd International BCS Human Computer Interaction Conference (HCI '18), Belfast, UK* (Swindon, United Kingdom, 2018), BCS Learning & Development Ltd., pp. 1–12. doi:10.14236/ewic/HCI2018.46.
- [215] Nielsen, G., Vedel, M., and Jensen, C. D. Improving usability of passphrase authentication. In *Proceedings of the 12th International Conference on Privacy, Security and Trust (PST 2014), Toronto, Canada* (New York, NY, USA, 2014), IEEE, pp. 189–198. doi:10.1109/PST.2014.6890939.
- [216] Nys, T. R., and Engelen, B. Judging nudging: Answering the manipulation objection. *Political Studies* 65, 1 (2017), 199–214. doi:10.1177/0032321716629487.
- [217] Obar, J. A., and Oeldorf-Hirsch, A. The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society* 23, 1 (2020), 128–147. doi:10.1080/1369118X.2018.1486870.
- [218] O’Gorman, L. Comparing passwords, tokens, and biometrics for user authentication. *Proceedings of the IEEE* 91, 12 (2003), 2021–2040. doi:10.1109/JPROC.2003.819611.
- [219] Ölander, F., and Thøgersen, J. Informing Versus Nudging in Environmental Policy. *Journal of Consumer Policy* 37, 3 (2014), 341–356. doi:10.1007/s10603-014-9256-2.
- [220] OneSpan. Cronto Visual Transaction Signing. Website, 2019. Online available from <https://www.onespan.com/products/transaction-signing/cronto>; last accessed 01 September 2020.
- [221] Oracle. Passlogix. Website, 2019. Online available from <https://www.oracle.com/corporate/acquisitions/passlogix/>; last accessed 03 September 2020.
- [222] Osman, M. An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin & Review* 11, 6 (2004), 988–1010. doi:10.3758/BF03196730.
- [223] Osman, M. Nudge: How far have we come? *Economia. History, Methodology, Philosophy* 6, 4 (2016), 557–570.
- [224] Paans, R., and Herschberg, I. S. Computer security: The long road ahead. *Computers & Security* 6, 5 (1987), 403–416. doi:10.1016/0167-4048(87)90013-7.
- [225] Palmer, A. J. Approach for selecting the most suitable Automated Personal Identification Mechanism (ASMSA). *computers & Security* 29, 7 (2010), 785–806. doi:10.1016/j.cose.2010.03.002.
- [226] Paolacci, G., and Chandler, J. Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science* 23, 3 (2014), 184–188. doi:10.1177/0963721414531598.
- [227] Parno, B., Kuo, C., and Perrig, A. Phoolproof phishing prevention. In *Proceedings of the International Conference on Financial Cryptography and Data Security, Anguilla, British West Isles* (Berlin/Heidelberg, Germany, 2006), G. Di Crescenzo and A. Rubin, Eds., Springer, pp. 1–19. doi:10.1007/11889663_1.

-
- [228] Parsons, K., Calic, D., Pattinson, M., Butavicius, M., McCormac, A., and Zwaans, T. The Human Aspects of Information Security Questionnaire (HAIS-Q): Two further validation studies. *Computers & Security* 66 (2017), 40–51. doi:10.1016/j.cose.2017.01.004.
- [229] Passfaces Corporation. Passfaces - Graphical password technology. Website, 2005 - 2019. Online available from <http://www.passfaces.com/index.htm>; last accessed 03 September 2020.
- [230] Pearman, S., Thomas, J., Naeni, P. E., Habib, H., Bauer, L., Christin, N., Cranor, L. F., Egelman, S., and Forget, A. Let’s Go in for a Closer Look: Observing Passwords in Their Natural Habitat. In *Proceedings of the 2017 ACM Conference on Computer and Communications Security (CCS ’17)*, Dallas, TX, USA (New York, NY, USA, 2017), ACM, pp. 295–310. doi:10.1145/3133956.3133973.
- [231] Pearman, S., Zhang, S. A., Bauer, L., Christin, N., and Cranor, L. F. Why people (don’t) use password managers effectively. In *Proceedings of the 15th Symposium on Usable Privacy and Security (SOUPS ’19)*, Santa Clara, CA, USA (Berkeley, CA, USA, 2019), USENIX Association. isbn:978-1-939133-05-2.
- [232] Peer, E., Egelman, S., Harbach, M., Malkin, N., Mathur, A., and Friks, A. Nudge me right: Personalizing online security nudges to people’s decision-making styles. *Computers in Human Behavior* 109, 106347 (2020). doi:10.1016/j.chb.2020.106347.
- [233] Perrow, C. Normal Accident at Three Mile Island. *Society* 18, 5 (1981), 17–26. doi:10.1007/BF02701322.
- [234] Petrovska-Delacrétaz, D., Chollet, G., and Dorizzi, B. *Guide to Biometric Reference Systems and Performance Evaluation*. Springer, 2009. doi:10.1007/978-1-84800-292-0.
- [235] Porter, S. N. A password extension for improved human factors. *Computers & Security* 1, 1 (1982), 54–56. doi:10.1016/0167-4048(82)90025-6.
- [236] Prakash, R., Kumar, S., Kumar, C., and Mishra, K. Musical password based biometric authentication. In *Proceedings of the 2016 International Conference on Computing, Communication and Automation (ICCCA 2016)*, Greater Noida, India (New York, NY, USA, 2016), IEEE, pp. 1016–1019. doi:10.1109/CCAA.2016.7813865.
- [237] Proctor, R. W., Lien, M.-C., Vu, K.-P. L., Schultz, E. E., and Salvendy, G. Improving computer security for authentication of users: Influence of proactive password restrictions. *Behavior Research Methods, Instruments, & Computers* 34, 2 (2002), 163–169. doi:10.3758/BF03195438.
- [238] Rabkin, A. Personal knowledge questions for fallback authentication: Security questions in the era of Facebook. In *Proceedings of the 4th Symposium on Usable Privacy and Security (SOUPS ’08)*, Pittsburgh, PA, USA (New York, NY, USA, 2008), ACM, pp. 13–23. doi:10.1145/1408664.1408667.
- [239] Raja, F., Hawkey, K., Hsu, S., Wang, K.-L. C., and Beznosov, K. A brick wall, a locked door, and a bandit: a physical security metaphor for firewall warnings. In *Proceedings of the 7th Symposium on Usable Privacy and Security (SOUPS ’11)*, Pittsburgh, PA, USA (New York, NY, USA, 2011), ACM, pp. 1–20. doi:10.1145/2078827.2078829.
- [240] Ratha, N. K., and Govindaraju, V. *Advances in Biometrics: Sensors, Algorithms and Systems*. Springer, London, UK, 2007. doi:10.1007/978-1-84628-921-7.
- [241] Redmiles, E. M., Dickerson, J. P., Gummadi, K. P., and Mazurek, M. L. Equitable Security: Optimizing Distribution of Nudges and Resources. In *Proceedings of the 2018 ACM Conference on Computer and Communications Security (CCS ’18)*, Toronto, Canada (New York, NY, USA, 2018), ACM, pp. 2270–2272. doi:10.1145/3243734.3278507.
- [242] Ren, Z., Meng, J., Yuan, J., and Zhang, Z. Robust hand gesture recognition with kinect sensor. In *Proceedings of the 19th ACM International Conference on Multimedia (MM ’11)*, Scottsdale, AZ, USA (New York, NY, USA, 2011), ACM, pp. 759–760. doi:10.1145/2072298.2072443.
- [243] Renaud, K. Evaluating authentication mechanisms. In *Security and Usability: Designing Secure Systems That People Can Use*, L. F. Cranor and S. Garfinkel, Eds. O’Reilly Media, Sebastopol, CA, USA, 2005, pp. 103–128. isbn:9780596008277.
- [244] Renaud, K. Guidelines for designing graphical authentication mechanism interfaces. *International Journal of Information and Computer Security* 3, 1 (2009), 60–85. doi:10.1504/IJICS.2009.026621.
-

-
- [245] Renaud, K., and Dupuis, M. Cyber Security Fear Appeals: Unexpectedly Complicated. In *Proceedings of the 2019 New Security Paradigms Workshop (NSPW '19), San Carlos, Costa Rica* (New York, NY, USA, 2019), ACM, pp. 42–56. doi:10.1145/3368860.3368864.
- [246] Renaud, K., and Just, M. Pictures or questions?: examining user responses to association-based authentication. In *Proceedings of the 24th BCS Interaction Specialist Group Conference (BCS '10), Dundee, UK* (Swindon, UK, 2010), BCS Learning & Development Ltd., pp. 98–107. doi:10.14236/ewic/HCI2010.14.
- [247] Renaud, K., Volkamer, M., and Maguire, J. ACCESS: describing and contrasting. In *Proceedings of the 2nd International Conference on Human Aspects of Information Security, Privacy, and Trust (HAS 2014), Heraklion, Greece* (Cham, Switzerland, 2014), T. Tryfonas and I. Askoxylakis, Eds., Springer, pp. 183–194. doi:10.1007/978-3-319-07620-1_17.
- [248] Renaud, K., and Zimmermann, V. Enriched nudges lead to stronger password replacements... but implement mindfully. In *Proceedings of the 2017 Information Security for South Africa (ISSA), Johannesburg, South Africa* (New York, NY, USA, 2017), IEEE, pp. 1–9. doi:10.1109/ISSA.2017.8251779.
- [249] Renaud, K., and Zimmermann, V. Ethical guidelines for nudging in information security & privacy. *International Journal of Human-Computer Studies* 120 (2018), 22–35. doi:10.1016/j.ijhcs.2018.05.011.
- [250] Renaud, K., and Zimmermann, V. Guidelines for ethical nudging in password authentication. *SAIEE Africa Research Journal* 109, 2 (2018), 102–118. doi:10.23919/SAIEE.2018.8531951.
- [251] Renaud, K., and Zimmermann, V. Encouraging password manager use. *Network Security 2019* (2019). doi: 10.1016/S1353-4858(19)30075-3.
- [252] Renaud, K., and Zimmermann, V. Nudging folks towards stronger password choices: providing certainty is the key. *Behavioural Public Policy* 3, 2 (2019), 228–258. doi:10.1017/bpp.2018.3.
- [253] Renaud, K., Zimmermann, V., Maguire, J., and Draper, S. Lessons learned from evaluating eight password nudges in the wild. In *Proceedings of the LASER Workshop: Learning from Authoritative Security Experiment Results (LASER 2017), Arlington, VA, USA* (Berkeley, CA, USA, 2017), USENIX Association, pp. 25–37. isbn: 978-1-931971-41-6.
- [254] Riley, C., Benyon, D., Johnson, G. I., and Buckner, K. Security in Context: Investigating the Impact of context on Attitudes towards Biometric Technology. In *Proceedings of the 24th BCS Interaction Specialist Group Conference (BCS '10) (Dundee, UK)* (Swindon, UK, 2010), BCS Learning & Development Ltd., pp. 108–116.
- [255] Riley, C., Buckner, K., Johnson, G., and Benyon, D. Culture & biometrics: regional differences in the perception of biometric authentication technologies. *AI & society* 24, 3 (2009), 295–306. doi:10.1007/s00146-009-0218-1.
- [256] Rochlin, G. I., La Porte, T. R., and Roberts, K. H. The self-designing high-reliability organization: Aircraft carrier flight operations at sea. *Naval War College Review* 40, 4 (1987), 76–92.
- [257] Rosa, C. The Influence of Normative Feedback on Password Creation by End-Users. Master's thesis, Technische Universität Darmstadt, Darmstadt, Germany, 2019. Master's Thesis.
- [258] Rosenthal, R., and Jacobson, L. *Pygmalion In The Classroom—Teacher Expectation and Pupils' Intellectual Development* Holt. Holt, Rinehart and Winston Inc., New York, NY, USA, 1968.
- [259] Rozin, P., Scott, S. E., Dingley, M., Urbanek, J. K., Jiang, H., and Kaltenbach, M. Nudge to nobesity I: Minor changes in accessibility decrease food intake. *Judgment and Decision Making* 6, 4 (2011), 323–332.
- [260] Samant, P., and Agarwal, R. Machine learning techniques for medical diagnosis of diabetes using iris images. *Computer Methods and Programs in Biomedicine* 157 (2018), 121–128.
- [261] Sasse, M. A. Usability and trust in information systems. In *Trust and Crime in Information Societies*, R. Mansell and B. Collins, Eds. Edward Elgar, Cheltenham, UK, 2005, pp. 319–348.
- [262] Scarfone, K., and Souppaya, M. Special Publication (NIST SP) 800-118: Guide to Enterprise Password Management (Draft). Tech. rep., National Institute of Standards and Technology, 2009. This draft has been retired (April 01, 2016).

-
- [263] Schacter, D. L. *The Seven Sins of Memory: How the Mind Forgets and Remembers*. Houghton Mifflin Harcourt, Boston/New York, USA, 2002.
- [264] Schechter, S., Brush, A. B., and Egelman, S. It's No Secret. Measuring the Security and Reliability of Authentication via "Secret" Questions. In *Proceedings of the 30th IEEE Symposium on Security and Privacy (S&P '09), Oakland, CA, USA* (New York, NY, USA, 2009), IEEE, pp. 375–390. doi:10.1109/SP.2009.11.
- [265] Schlömer, T., Poppinga, B., Henze, N., and Boll, S. Gesture recognition with a Wii controller. In *Proceedings of the 2nd International Conference on Tangible and Embedded Interaction (TEI '08), Bonn, Germany* (New York, NY, USA, 2008), ACM, pp. 11–14. doi:10.1145/1347390.1347395.
- [266] Schmidt, R. F., Lang, F., and Heckmann, M. *Physiologie des Menschen: mit Pathophysiologie*. Springer, Berlin/Heidelberg, Germany, 2011. doi:10.1007/978-3-642-01651-6.
- [267] Schneier, B. *Beyond Fear: Thinking Sensibly About Security in an Uncertain World*. Springer Science & Business Media, New York, NY, USA, 2006. doi:10.1007/b97547.
- [268] Schneier, B. The psychology of security. In *Proceedings of the International Conference on Cryptology in Africa (Africacrypt 2008), Casablanca, Morocco* (Berlin/Heidelberg, Germany, 2008), Springer, pp. 50–79. doi:10.1007/978-3-540-68164-9_5.
- [269] Schneier, B. *Secrets and Lies: Digital Security in a Networked World*. John Wiley & Sons, 2011.
- [270] Schneier, B. Choosing secure passwords. Website, 2014. Online available from https://www.schneier.com/blog/archives/2014/03/choosing_secure_1.html; last accessed 08 September 2020.
- [271] Schubert, C. On the ethics of public nudging: Autonomy and agency. *Joint Discussion Paper Series in Economics* (2015).
- [272] Scofield, G. R. And as for the nudgees? *The American Journal of Bioethics* 13, 6 (2013), 25–27. doi:10.1080/15265161.2013.781705.
- [273] Seeing Machines. faceLAB. Website, 2019. Online available from <https://www.seeingmachines.com/>; last accessed 01 September 2020.
- [274] Segreti, S. M., Melicher, W., Komanduri, S., Melicher, D., Shay, R., Ur, B., Bauer, L., Christin, N., Cranor, L. F., and Mazurek, M. L. Diversify to survive: making passwords stronger with adaptive policies. In *Proceedings of the 13th Symposium on Usable Privacy and Security (SOUPS '17), Santa Clara, CA, USA* (Berkeley, CA, USA, 2017), USENIX Association, pp. 1–12.
- [275] Seitz, T., and Hussmann, H. PASDJO: quantifying password strength perceptions with an online game. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction (OzCHI '17), Brisbane, Australia* (New York, NY, USA, 2017), ACM, pp. 117–125. doi:10.1145/3152771.3152784.
- [276] Seitz, T., von Zezschwitz, E., Meitner, S., and Hussmann, H. Influencing Self-Selected Passwords Through Suggestions and the Decoy Effect. In *Proceedings of the 1st European Workshop on Usable Security (EuroUSEC '16), Darmstadt, Germany* (New York, NY, USA, 2016), IEEE, pp. 1–7. doi:10.14722/EUROUSEC.2016.23002.
- [277] Senarath, A., Arachchilage, N. A. G., and Gupta, B. Security strength indicator in fallback authentication: Nudging users for better answers in secret questions. *International Journal for Infonomics* 9, 4 (2017). doi:10.20533/IJI.1742.4712.2016.0150.
- [278] Shannon, C. E. A mathematical theory of communication. *The Bell System Technical Journal* 27, 3 (1948), 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x.
- [279] Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., and Ekelid, M. Speech recognition with primarily temporal cues. *Science* 270, 5234 (1995), 303–304. doi:10.1126/science.270.5234.303.
- [280] Shay, R., Kelley, P. G., Komanduri, S., Mazurek, M. L., Ur, B., Vidas, T., Bauer, L., Christin, N., and Cranor, L. F. Correct horse battery staple: Exploring the usability of system-assigned passphrases. In *Proceedings of the 8th Symposium on Usable Privacy and Security (SOUPS '12), Washington DC, USA* (New York, NY, USA, 2012), ACM, pp. 1–20. doi:10.1145/2335356.2335366.

-
- [281] Shay, R., Komanduri, S., Kelley, P. G., Leon, P. G., Mazurek, M. L., Bauer, L., Christin, N., and Cranor, L. F. Encountering stronger password requirements: user attitudes and behaviors. In *Proceedings of the 6th Symposium on Usable Privacy and Security (SOUPS '10)*, Redmond, WA, USA (2010), pp. 1–20. doi:10.1145/1837110.1837113.
- [282] Siddique, K., Akhtar, Z., and Kim, Y. Biometrics vs passwords: a modern version of the tortoise and the hare. *Computer Fraud & Security 2017*, 1 (2017), 13–17. doi:10.1016/S1361-3723(17)30007-6.
- [283] Singh, Y. N., and Singh, S. K. Evaluation of Electrocardiogram for Biometric Authentication. *Journal of Information Security* 3, 1 (2011), 39–48. doi:10.4236/jis.2012.31005.
- [284] Singh, Y. N., and Singh, S. K. A taxonomy of biometric system vulnerabilities and defences. *International Journal of Biometrics* 5, 2 (2013), 137–159. doi:10.1504/IJBM.2013.052964.
- [285] Sloman, S. A. The empirical case for two systems of reasoning. *Psychological Bulletin* 119, 1 (1996), 3–22. doi:10.1037/0033-2909.119.1.3.
- [286] Smith, H. J., Milberg, S. J., and Burke, S. J. Information Privacy: Measuring Individuals' Concerns About Organizational Practices. *MIS Quarterly* 20, 2 (1996), 167–196.
- [287] Sobrado, L., and Birget, J.-C. Graphical passwords. *The Rutgers Scholar - An Electronic Bulletin for Undergraduate Research* 5 (2003), 12–18.
- [288] Song, Y., Cho, G., Oh, S., Kim, H., and Huh, J. H. On the Effectiveness of Pattern Lock Strength Meters: Measuring the Strength of Real World Pattern Locks. In *Proceedings of the 33rd ACM Conference on Human Factors in Computing Systems (CHI '15)*, Seoul, South Korea (New York, NY, USA, 2015), ACM, pp. 2343–2352. doi:10.1145/2702123.2702365.
- [289] Stajano, F. Pico: No more passwords! In *Proceedings of the 19th International Workshop on Security Protocols (Security Protocols 2011)*, Cambridge, UK (Berlin/Heidelberg, Germany, 2011), B. Christianson, B. Crispo, J. Malcolm, and F. Stajano, Eds., Springer, pp. 49–81. doi:10.1007/978-3-642-25867-1_6.
- [290] Stanovich, K. E., and West, R. F. Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences* 23, 5 (2000), 645–665. doi:10.1017/S0140525X00003435.
- [291] Starnberger, G., Frohofer, L., and Göschka, K. M. QR-TAN: Secure mobile transaction authentication. In *Proceedings of the International Conference on Availability, Reliability and Security (ARES 2009)*, Fukuoka, Japan (New York, NY, USA, 2009), IEEE, pp. 578–583. doi:10.1109/ARES.2009.96.
- [292] Statista. Cybersecurity & Cloud 2018. Website, 2018. Online available from <https://de.statista.com/statistik/studie/id/58204/dokument/cybersecurity-und-cloud/>; last accessed 10 September 2020.
- [293] Stobert, E., and Biddle, R. The password life cycle: user behaviour in managing passwords. In *Proceedings of the 10th Symposium On Usable Privacy and Security (SOUPS '14)*, Menio Park, CA, USA (Berkeley, CA, USA, 2014), USENIX Association, pp. 243–255.
- [294] Sun, C., Wang, Y., and Zheng, J. Dissecting pattern unlock: The effect of pattern strength meter on pattern selection. *Journal of Information Security and Applications* 19, 4–5 (2014), 308–320. doi:10.1016/j.jisa.2014.10.009.
- [295] Sun, J., and Ahluwalia, P. How users respond to authentication methods: A study of security readiness. In *Proceedings of the Americas Conference on Information Systems (AMCIS 2008)*, Toronto, Canada (Atlanta, GA, USA, 2008), Association for Information Systems, pp. 1–10.
- [296] Sun, J., Ahluwalia, P., and Koong, K. S. The more secure the better? A study of information security readiness. *Industrial Management & Data Systems* 111, 4 (2011), 570–588. doi:10.1108/02635571111133551.
- [297] Sunstein, C. R. *Why Nudge?: The Politics of Libertarian Paternalism*. Yale University Press, New Haven, CT, USA, 2014.
- [298] Sunstein, C. R. Nudges do not undermine human agency. *Journal of Consumer Policy* 38, 3 (2015), 207–210. doi:10.1007/s10603-015-9289-1.

-
- [299] Sunstein, C. R. Nudges that fail. *Behavioural Public Policy* 1, 1 (2017), 4–25. doi:10.1017/bpp.2016.3.
- [300] Suo, X., Zhu, Y., and Owen, G. S. Graphical passwords: A survey. In *Proceedings of the 21st Computer Security Applications Conference (ACSAC'05), Tucson, AZ, USA* (New York, NY, USA, 2005), IEEE, pp. 1–10. doi:10.1109/CSAC.2005.27.
- [301] Tao, H., and Adams, C. Pass-go: A proposal to improve the usability of graphical passwords. *International Journal of Network Security* 7, 2 (2008), 273–292.
- [302] Thaler, R. H. Nudge, not sludge. *Science* 361, 6401 (2018), 431–431.
- [303] Thaler, R. H., and Sunstein, C. R. *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press, New Haven & London, 2008.
- [304] The Behavioural Insights Team. Understanding People, Better Outcomes: Behavioral Insights in NSW, 2014. Online available from <https://www.dpc.nsw.gov.au/programs-and-services/behavioural-insights/blog/2014/06/04/understanding-people-better-outcomes-behavioural-insights-in-nsw/>; last accessed 13 September 2020.
- [305] The British Psychological Society. Code of Human Research Ethics (2nd edition). Website, 2014. Online available from <https://www.bps.org.uk/news-and-policy/bps-code-human-research-ethics-2nd-edition-2014>; last accessed 14 September 2020.
- [306] Thorpe, J., and van Oorschot, P. C. Human-Seeded Attacks and Exploiting Hot-Spots in Graphical Passwords. In *Proceedings of the 16th USENIX Security Symposium (USENIX Security '07), Boston, MA, USA* (Berkeley, CA, USA, 2007), USENIX Association, pp. 1–8.
- [307] Thorpe, J., Van Oorschot, P. C., and Somayaji, A. Pass-thoughts: authenticating with our minds. In *Proceedings of the 2005 Workshop on New Security Paradigms (NSPW '05), Lake Arrowhead, CA, USA* (New York, NY, USA, 2005), ACM, pp. 45–56. doi:10.1145/1146269.1146282.
- [308] Toledano, D. T., Fernández Pozo, R., Hernández Trapote, Á., and Hernández Gómez, L. Usability evaluation of multi-modal biometric verification systems. *Interacting with Computers* 18, 5 (2006), 1101–1122. doi:10.1016/j.intcom.2006.01.004.
- [309] Turk, M. A., and Pentland, A. P. Face Recognition Using Eigenfaces. In *Proceedings of the 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 1991), Maui, HI, USA* (New York, NY, USA, 1991), IEEE, pp. 586–591. doi:10.1109/CVPR.1991.139758.
- [310] Turland, J., Coventry, L., Jeske, D., Briggs, P., and van Moorsel, A. Nudging towards security: Developing an application for wireless network selection for android phones. In *Proceedings of the 2015 British HCI conference (BCS-HCI 2015), Lincoln, UK* (New York, NY, USA, 2015), ACM, pp. 193–201. doi:10.1145/2783446.2783588.
- [311] Tversky, A., and Kahneman, D. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology* 5, 2 (1973), 207–232. doi:10.1016/0010-0285(73)90033-9.
- [312] Uellenbeck, S., Dürmuth, M., Wolf, C., and Holz, T. Quantifying the security of graphical passwords: the case of android unlock patterns. In *Proceedings of the 20th ACM Conference on Computer and Communications Security (CCS '13), Berlin, Germany* (New York, NY, USA, 2013), ACM, pp. 161–172. doi:10.1145/2508859.2516700.
- [313] Ur, B., Alfieri, F., Aung, M., Bauer, L., Christin, N., Colnago, J., Cranor, L. F., Dixon, H., Emami Naeini, P., Habib, H., Johnson, N., and Melicher, W. Design and evaluation of a data-driven password meter. In *Proceedings of the 35th ACM Conference on Human Factors in Computing Systems (CHI '17), Denver, CO, USA* (New York, NY, USA, 2017), ACM, pp. 3775–3786. doi:10.1145/3025453.3026050.
- [314] Ur, B., Bees, J., Segreti, S. M., Bauer, L., Christin, N., and Cranor, L. F. Do Users' Perceptions of Password Security Match Reality? In *Proceedings of the 34th ACM Conference on Human Factors in Computing Systems (CHI '16), San Jose, CA, USA* (New York, NY, USA, 2016), ACM, pp. 3748–3760. doi:10.1145/2858036.2858546.

-
- [315] Ur, B., Noma, F., Bees, J., Segreti, S. M., Shay, R., Bauer, L., Christin, N., and Cranor, L. F. "I Added '!' at the End to Make It Secure": Observing Password Creation in the Lab. In *Proceedings of the 11th Symposium On Usable Privacy and Security (SOUPS '15)*, Ottawa, Canada (2015), pp. 123–140.
- [316] Ur, B., Segreti, S. M., Bauer, L., Christin, N., Cranor, L. F., Komanduri, S., Kurilova, D., Mazurek, M. L., Melicher, W., and Shay, R. Measuring real-world accuracies and biases in modeling password guessability. In *Proceedings of the 24th USENIX Security Symposium (USENIX Security '15)*, Washington, D.C., USA (Berkeley, CA, USA, 2015), USENIX Association, pp. 463–481. isbn:9781931971232.
- [317] Van Acker, S., Hausknecht, D., Joosen, W., and Sabelfeld, A. Password Meters and Generators on the Web: From Large-Scale Empirical Study to Getting It Right. In *Proceedings of the 5th ACM Conference on Data and Application Security and Privacy (CODAPSY '15)*, San Antonio, TX, USA (New York, NY, USA, 2015), ACM, pp. 253–262. doi:10.1145/2699026.2699118.
- [318] Vance, A., Eargle, D., Ouimet, K., and Straub, D. Enhancing Password Security through Interactive Fear Appeals: A Web-Based Field Experiment. In *Proceedings of the 46th Hawaii International Conference on System Sciences (HICSS 2013)*, Wailea, HI, USA (New York, NY, USA, 2013), IEEE, pp. 2988–2997. doi: 10.1109/HICSS.2013.196.
- [319] Velásquez, I., Caro, A., and Rodríguez, A. Authentication schemes and methods: A systematic literature review. *Information and Software Technology* 94 (2018), 30–37. doi:10.1016/j.infsof.2017.09.012.
- [320] von Zezschwitz, E., Eiband, M., Buschek, D., Oberhuber, S., De Luca, A., Alt, F., and Hussmann, H. On quantifying the effective password space of grid-based unlock gestures. In *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia (MUM '16)*, Rovaniemi, Finland (New York, NY, USA, 2016), ACM, pp. 201–212. doi:10.1145/3012709.3012729.
- [321] Wang, D., Gu, Q., Cheng, H., and Wang, P. The request for better measurement: A comparative evaluation of two-factor authentication schemes. In *Proceedings of the 11th ACM Asia Conference on Computer and Communications Security (ASIA CCS '16)*, Xi'an, China (New York, NY, USA, 2016), ACM, pp. 475–486. doi:10.1145/2897845.2897916.
- [322] Wang, Y., Leon, P. G., Acquisti, A., Cranor, L. F., Forget, A., and Sadeh, N. A field trial of privacy nudges for facebook. In *Proceedings of the 32nd ACM Conference on Human Factors in Computing Systems (CHI '14)*, Toronto, Canada (New York, NY, USA, 2014), ACM, pp. 2367–2376. doi:10.1145/2556288.2557413.
- [323] Wash, R., Rader, E., Berman, R., and Wellmer, Z. Understanding password choices: How frequently entered passwords are re-used across websites. In *Proceedings of the 12th Symposium on Usable Privacy and Security (SOUPS '16)*, Denver, CO, USA (Berkeley, CA, USA, 2016), USENIX Association, pp. 175–188. isbn:978-1-931971-31-7.
- [324] Wayman, J. L., Jain, A. K., Maltoni, D., and Maio, D. *Biometric Systems: Technology, Design and Performance Evaluation*. Springer, London, UK, 2005.
- [325] Weaver, A. C. Biometric authentication. *Computer* 39, 2 (2006), 96–97. doi:10.1109/MC.2006.47.
- [326] Wei, M., Golla, M., and Ur, B. The Password Doesn't Fall Far: How Service Influences Password Choice. In *Proceedings of Who Are You?! Adventures in Authentication Workshop (WAY 2018)*, Baltimore, MD, USA (Berkeley, CA, USA, 2018), USENIX Association.
- [327] Weinmann, M., Schneider, C., and Vom Brocke, J. Digital nudging. *Business & Information Systems Engineering* 58, 6 (2016), 433–436. doi:10.1007/s12599-016-0453-1.
- [328] Weinshall, D. Cognitive authentication schemes safe against spyware. In *Proceedings of the 27th IEEE Symposium on Security and Privacy (S&P'06)*, Oakland, CA, USA (New York, NY, USA, 2006), IEEE, pp. 1–6. doi:10.1109/SP.2006.10.
- [329] Weir, M., Aggarwal, S., Collins, M., and Stern, H. Testing metrics for password creation policies by attacking large sets of revealed passwords. In *Proceedings of the 17th ACM Conference on Computer and Communications Security (CCS '10)*, Chicago, IL, USA (New York, NY, USA, 2010), ACM, pp. 162–175. doi:10.1145/1866307.1866327.

-
- [330] Weiss, R., and De Luca, A. PassShapes: utilizing stroke based authentication to increase password memorability. In *Proceedings of the 5th Nordic Conference on Human-Computer Interaction (NordiCHI '08), Lund, Sweden* (New York, NY, USA, 2008), ACM, pp. 383–392. doi:10.1145/1463160.1463202.
- [331] Wheeler, D. L. zxcvbn: Low-budget password strength estimation. In *Proceedings of the 25th USENIX Security Symposium (USENIX Security '16), Austin, TX, USA* (Berkeley, CA, USA, 2016), USENIX Association, pp. 157–173. isbn:978-1-931971-32-4.
- [332] Whitten, A., and Tygar, J. D. Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0. In *Proceedings of the 8th USENIX Security Symposium (USENIX Security '99), Washington, D.C., USA* (1999), vol. 8, pp. 169–184.
- [333] Wiedenbeck, S., Waters, J., Birget, J.-C., Brodskiy, A., and Memon, N. PassPoints: Design and longitudinal evaluation of a graphical password system. *International Journal of Human-Computer Studies* 63, 1–2 (2005), 102–127. isbn:9780399122237.
- [334] Wildes, R. P. Iris Recognition: An Emerging Biometric Technology. *Proceedings of the IEEE* 85, 9 (1997), 1348–1363. doi:10.1109/5.628669.
- [335] Wilds, M., and Chambers, S. Bar code authentication, Jan. 21 2010. US Patent App. 12/306,460.
- [336] Wilkinson, T. M. Nudging and Manipulation. *Political Studies* 61, 2 (2013), 341–355. doi:10.1111/j.1467-9248.2012.00974.x.
- [337] Witte, K. Putting the fear back into fear appeals: The extended parallel process model. *Communication Monographs* 59, 4 (1992), 329–349.
- [338] Woo, S. S., and Mirkovic, J. GuidedPass: Helping Users to Create Strong and Memorable Passwords. In *Proceedings of the 21st International Symposium on Research in Attacks, Intrusions, and Defenses (RAID 2018), Heraklion, Greece* (Cham, Switzerland, 2018), M. Bailey, T. Holz, M. Stamatogiannakis, and S. Ioannidis, Eds., Springer, pp. 250–270. doi:10.1007/978-3-030-00470-5_12.
- [339] World Economic Forum. The Global Risks Report 2017 12th Edition. Website, 2017. Online available from http://www3.weforum.org/docs/GRR17_Report_web.pdf; last accessed 14 September 2020.
- [340] Yan, J., Blackwell, A., Anderson, R., and Grant, A. Password memorability and security: Empirical results. *IEEE Security & Privacy* 2, 5 (2004), 25–31. doi:10.1109/MSP.2004.81.
- [341] Yan, P., and Bowyer, K. W. Biometric Recognition Using 3D Ear Shape. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29, 8 (2007), 1297–1308. doi:10.1109/TPAMI.2007.1067.
- [342] Yıldırım, M., and Mackie, I. Encouraging users to improve password security and memorability. *International Journal of Information Security* 18, 6 (2019), 741–759. doi:10.1007/s10207-019-00429-y.
- [343] Yubico. The YubiKey. Website, 2019. Online available from <https://www.yubico.com/products/yubikey-hardware/>; last accessed 06 September 2020.
- [344] Zhang, Y., Monrose, F., and Reiter, M. K. The security of modern password expiration: An algorithmic framework and empirical analysis. In *Proceedings of the 17th ACM Conference on Computer and Communications Security (CCS 2010), Chicago, IL, USA* (New York, NY, USA, 2010), ACM, pp. 176–186. doi:10.1145/1866307.1866328.
- [345] Zhang-Kennedy, L., Chiasson, S., and van Oorschot, P. Revisiting password rules: facilitating human management of passwords. In *Proceedings of the 2016 APWG Symposium on Electronic Crime Research (eCrime), Toronto, Canada* (New York, NY, USA, 2016), IEEE, pp. 1–10. doi:10.1109/ECRIME.2016.7487945.
- [346] Zimmermann, V., and Gerber, N. “If It Wasn't Secure, They Would Not Use It in the Movies” - Security Perceptions and User Acceptance of Authentication Technologies. In *Proceedings of the 5th International Conference on Human Aspects of Information Security, Privacy, and Trust (HAS 2017), Vancouver, Canada* (Cham, Switzerland, 2017), T. Tryfanos, Ed., Springer, pp. 265–283. doi:10.1007/978-3-319-58460-7_18.
- [347] Zimmermann, V., and Gerber, N. The password is dead, long live the password – A laboratory study on user perceptions of authentication schemes. *International Journal of Human-Computer Studies* 133 (2020), 26–44. doi:10.1016/j.ijhcs.2019.08.006.

-
- [348] Zimmermann, V., Gerber, N., Kleboth, M., von Preuschen, A., Schmidt, K., and Mayer, P. The Quest to Replace Passwords Revisited – Rating Authentication Schemes. In *Proceedings of the 12th International Symposium on Human Aspects of Information Security & Assurance (HAISA 2018)*, Dundee, UK (Plymouth, UK, 2018), Plymouth University, pp. 38–48.
- [349] Zimmermann, V., Gerber, N., Mayer, P., Kleboth, M., von Preuschen, A., and Schmidt, K. Keep on rating – On the systematic rating and comparison of authentication schemes. *Information & Computer Security* 27, 5 (2019), 621–635. doi:10.1108/ICS-01-2019-0020.
- [350] Zimmermann, V., Gerber, P., Marky, K., Böck, L., and Kirchbuchner, F. Assessing Users’ Privacy and Security Concerns of Smart Home Technologies. *i-com* 18, 3 (2019), 197–216. doi:10.1515/icom-2019-0015.
- [351] Zimmermann, V., Gerber, P., and Stöver, A. That Depends – How Context Affects User Perceptions of Authentication Schemes.
- [352] Zimmermann, V., Marky, K., and Renaud, K. Hybrid password meters for more secure passwords - A comprehensive study of password meters and nudges. *Behaviour & Information Technology* (submitted).
- [353] Zimmermann, V., and Renaud, K. Moving from a ‘human-as-problem’ to a ‘human-as-solution’ cybersecurity mindset. *International Journal of Human-Computer Studies* 131 (2019), 169–187. doi:10.1016/j.ijhcs.2019.05.005.
- [354] Zimmermann, V., and Renaud, K. The Nudge Puzzle: Matching Nudge Interventions to Cybersecurity Decisions. *ACM Transactions on Computer-Human Interaction* 28 (2021), 7:1–7:45. doi:10.1145/3429888.
- [355] Zviran, M., and Haga, W. J. Password Security: An Empirical Study. *Journal of Management Information Systems* 15, 4 (1999), 161–185. doi:10.1080/07421222.1999.11518226.

List of Abbreviations

Abbreviation	Explanation
<i>APA</i>	American Psychological Association
<i>ATI</i>	Affinity for Technology Interaction scale
<i>ACCESS</i>	Authentication Choice Support System
<i>BPS</i>	British Psychological Society
<i>CTSS</i>	Compatible Time-Sharing System
<i>df</i>	Degrees of Freedom
<i>EFPA</i>	European Federation of Psychologists' Association
<i>H AIS-Q</i>	Human Aspects of Information Security Questionnaire
<i>HTML</i>	Hypertext Markup Language
<i>M / \bar{x}</i>	Mean
<i>Md / \tilde{x}</i>	Median
<i>MIT</i>	Massachusetts Institute of Technology
<i>NIST</i>	National Institute of Standards and Technology
<i>PIN</i>	Personal Identification Number
<i>PCCP</i>	Persuasive Cued Click Points
<i>PSD2</i>	European Payment Services Directive
<i>RFID</i>	radio-frequency identification
<i>SD/ σ</i>	Standard Deviation
<i>SeBIS</i>	Security Behavior Intentions Scale
<i>SUS</i>	System Usability Scale
<i>TAM</i>	Technology Acceptance Model
<i>WPR</i>	"What's the problem represented to be" approach

List of Figures

1	Outline of the dissertation.	3
2	Research step 1.	4
3	Categorization of knowledge-based authentication schemes. <i>Note:</i> The list is not exhaustive. For references to the examples listed in the figure, see Appendix 10.5.	7
4	Categorization of biometric authentication schemes. <i>Note:</i> The list is not exhaustive. For references to the examples listed in the figure see Appendix 10.5. *DNA can be clustered differently.	10
5	Categorization of token-based authentication schemes. <i>Note:</i> The list is not exhaustive. For references to the examples listed in the figure see Appendix 10.5.	14
6	Research step 2.	16
7	Procedure of the rating process conducted to compare and select authentication schemes for this research as described in [348].	17
8	Research step 3.	18
9	Study procedure of the laboratory study on user perceptions of authentication schemes, adapted from [347].	21
10	Research step 4.	23
11	Study procedure of the online study on the influence of the context and type of scheme on user perceptions of authentication schemes [351].	26
12	Interim conclusion.	28
13	Research step 5.	30
14	Research step 6.	35
15	Study procedure of the three field studies exploring the influence of password nudges on password creation as described in [252]. Images included with permission of © 2018 Cambridge University Press.	38
16	Research step 7.	41
17	Research step 8.	45
18	Graphical depiction of the process to differentiate the nudge from related interventions before analysing ethical implications for the selected intervention types as described in [249].	45
19	Research step 9.	49
20	Screenshot of the password creation hybrid nudge condition with dynamic information based on Ur <i>et al.</i> [313].	51
21	Study procedure of the online study analysing the single and joint effects of nudging and information provision in hybrid nudges. <i>Note:</i> Figure also included [354].	52
22	Research step 10.	57
23	Study procedure of the online study analysing hybrid password meters based on different biases and heuristics as described in Manuscript 7 that has been under review by the journal <i>Behaviour & Information Technology</i> at the time of the publication of this dissertation [352]. The exemplary screenshots show varying password scores to illustrate differences in the feedback.	61
24	Conclusion	63
25	Research Step 11.	73
26	Graphical depiction of the problematization approach and the "Cybersecurity, Differently" principles derived thereof, adapted from [353].	74

List of Tables

1	Differences between the criteria for a nudge and related interventions.	43
2	Overview of the conditions and interventions tested in the study to analyse the single and joint effect of simple nudges and information provision. <i>Note:</i> This is an adapted version of a table also included in [354].	52
3	Descriptive values and Mann-Whitney-U test results of the password strength values, M = Mean, SD = Standard deviation, Md = median, Z = standardized test statistic, df = degrees of freedom, p = level of significance, r = Effect size. <i>Note:</i> This is an adapted version of a table also included in [354].	54
4	Descriptive values and Mann-Whitney-U test results of the password entropy values, M = Mean, SD = Standard deviation, Md = Median, Z = standardized test statistic, df = degrees of freedom, p = level of significance, r = Effect size. <i>Note:</i> This is an adapted version of a table also included in [354].	54
5	Descriptive values and Mann-Whitney-U test results of the password length values, M = Mean, SD = Standard deviation, Md = Median, Z = standardized test statistic, df = degrees of freedom, p = level of significance, r = Effect size. <i>Note:</i> This is an adapted version of a table also included in [354].	55
6	Vita of Verena Zimmermann.	110
7	Reference list of exemplary knowledge-based authentication schemes.	115
8	Reference list of exemplary biometric authentication schemes.	116
9	Reference list of exemplary token-based authentication schemes.	117

Appendix

Appendix A: Author Vita

Verena Zimmermann, born 23 June 1990 in Dieburg, Germany, is a researcher in the field of Human Factors in Safety and Security. After completing her Master's degree in psychology she worked as a doctoral researcher at in the research group "Work and Engineering Psychology" at Technische Universität Darmstadt and within the German National Research Center for Applied Cybersecurity (ATHENE).

2018 - today	Associate in the RTG "Privacy and Trust for Mobile Users" at Technische Universität Darmstadt
2016 - today	Doctoral Researcher, Research Group Work and Engineering Psychology (Prof. Joachim Vogt), Technische Universität Darmstadt, Germany
2019	IANUS Research Award for scientific-technological peace and security research (2nd place)
2018	Research Award of the Department of Human Sciences at Technische Universität Darmstadt for an innovative doctoral research proposal
2018	Research stay, Division of Cybersecurity (Prof. Karen Renaud), Abertay University, Dundee, Scotland
2016	August-Euler Award for outstanding scientific contributions in the area of aviation and air traffic
2015	Research stay, Safety Science Innovation Lab (Prof. Sidney Dekker), Griffith University, Brisbane, Australia
2012 - 2015	Master of Science in Psychology, Technische Universität Darmstadt, Germany Master's thesis: "Public Value eines Flughafens in der Wahrnehmung verschiedener Interessensgruppen"
2014	Internship, Human Factors Training, Lufthansa Aviation Training (formerly Lufthansa Flight Training), Frankfurt, Germany
2009 - 2012	Bachelor of Science in Psychology, Technische Universität Darmstadt, Germany Bachelor's thesis: "Spatial Processing of Traffic Information in Sport Aviation – A Randomized Controlled Study of Mental Rotation Processes in Non-Pilot Participants"
2012	Semester abroad at the University of Hertfordshire, Hatfield, UK
2009	Abitur (A-Levels) at Lichtenberggymnasium Darmstadt

Table 6: Vita of Verena Zimmermann.

Appendix B: List of Publications

2021

- Zimmermann, V.**, Gerber, P., and Stöver, A. That Depends – How context Affects User Perceptions of Authentication Schemes. (in preparation).
- Zimmermann, V.**, and Renaud, K. The Nudge Puzzle: Matching Nudge Interventions to Cybersecurity Decisions. *ACM Transactions on Computer-Human Interaction* 28 (2021), 7:1-7:45. doi:10.1145/3429888.
-

2020

- Marky, K., Schmitz, M., **Zimmermann, V.**, Herbers, M., Kunze, K., and Mühlhäuser, M. 3d-auth: Two-factor authentication with personalized 3d-printed items. In *Proceedings of the 38th ACM Conference on Human Factors in Computing Systems (CHI '20)*, Honolulu, HI, USA (New York, NY, USA, 2020), ACM, pp. 1-12. doi:10.1145/3313831.3376189.
- Marky, K., **Zimmermann, V.**, Funk, M., Daubert, J., Bleck, K., and Mühlhäuser, M. Improving the Usability and UX of the Swiss Internet Voting Interface. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*, Honolulu, HI, USA, (New York, NY, USA), ACM, pp. 1-13. doi:10.1145/3313831.3376769.
- Marky, K., **Zimmermann, V.**, Stöver, A., Hoffmann, P., Kunze, K., and Mühlhäuser, M. All in One! User Perceptions on Centralized IoT Privacy Settings. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*, Honolulu, HI, USA, (New York, NY, USA, 2020), ACM, pp. 1-8. doi:10.1145/3334480.3383016.
- Renaud, K., and **Zimmermann, V.** How to nudge in cybersecurity. *Network Security* 2020, 11 (2020), 20-20. doi:10.1016/S1353-4858(20)30132-X.
- Zimmermann, V.**, and Gerber, N. The password is dead, long live the password – A laboratory study on user perceptions of authentication schemes. *International Journal of Human-Computer Studies* 133 (2020), 26–44. doi:10.1016/j.ijhcs.2019.08.006.
- Zimmermann, V.**, Marky, K., and Renaud, K. Hybrid password meters for more secure passwords- A comprehensive study of password meters and nudges. *Behaviour & Information Technology* (submitted).
-

2019

- Gerber, N., **Zimmermann, V.**, and Volkamer, M. Why Johnny Fails to Protect his Privacy. In *Proceedings of the 2019 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, (New York, NY, USA, 2019), IEEE, pp. 109-118. doi:10.1109/EuroSPW.2019.00019.
- Menig, A., **Zimmermann, V.**, and Vogt, J. (2019). Digital Transformation of the Workplace — Risk or Opportunity? In *Digitalisation and Communication: Societal Trends and the Change in Organisations, Science Policy Paper No. 6*, C. Reuter, T. Schultz and C. Stegbauer, , Eds., Mercator Science-Policy Fellowship-Programme, Frankfurt am Main, Germany, 2019, pp. 29-34.
- Renaud, K., and **Zimmermann, V.** Encouraging password manager use. *Network Security* 2019, 6 (2019), 20-20, doi:10.1016/S1353-4858(19)30075-3.
- Renaud, K., and **Zimmermann, V.** Nudging folks towards stronger password choices: providing certainty is the key. *Behavioural Public Policy* 3, 2 (2019), 228–258. doi:10.1017/bpp.2018.3.
- Zimmermann, V. (née Schochlow)**, and Dekker, S. The 1980s and Onward: Normal Accidents and High Reliability Organizations. In *Foundations of Safety Science* S. Dekker, Ed., CRC Press, Boca Raton, FL, USA, 2019, pp. 267-304. doi:10.4324/9781351059794.
-

-
- Zimmermann, V.**, Dickhaut, E., Gerber, P., and Vogt, J. (2019). Vision: Shining Light on Smart Homes – Supporting Informed Decision-Making of End Users. In *Proceedings of the 2019 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, (New York, NY, USA, 2019), IEEE, pp. 149-153. doi:10.1109/EuroSPW.2019.00023.
- Zimmermann, V.**, Gerber, N., Mayer, P., Kleboth, M., von Preuschen, A., and Schmidt, K. Keep on rating—on the systematic rating and comparison of authentication schemes. *Information & Computer Security* 27, 5 (2019), 621–635. doi:10.1108/ICS-01-2019-002
- Zimmermann, V.**, Gerber, P., Marky, K., Böck, L., and Kirchbuchner, F. Assessing users’ privacy and security concerns of smart home technologies. *i-com* 18, 3 (2019), 197–216. doi:10.1515/icom-2019-0015
- Zimmermann, V.**, and Renaud, K. Moving from a ‘human-as-problem’ to a ‘human-as-solution’ cybersecurity mindset. *International Journal of Human-Computer Studies* 131 (2019), 169–187. doi:10.1016/j.ijhcs.2019.05.005.
-

2018

- Gerber, N., **Zimmermann, V.**, Henhapl, B., Emeröz, S. and Volkamer, M. Finally Johnny Can Encrypt: But Does This Make Him Feel More Secure? In *Proceedings of the 13th International Conference on Availability, Reliability and Security (ARES 2018)*, Hamburg, Germany (New York, NY, USA, 2018), ACM, pp. 1-10. doi:10.1145/3230833.3230859.
- Gerber, N., **Zimmermann, V.**, Henhapl, B., Emeröz, S., Volkamer, M., and Hilt, T. Nutzerwahrnehmung der Ende-zu-Ende-Verschlüsselung in WhatsApp. *Datenschutz und Datensicherheit-DuD* 42, 11 (2018) 680-685. doi:10.1007/s11623-018-1024-z.
- Marky, K., Mayer, P., Gerber, N., and **Zimmermann, V.** (2018). Assistance in Daily Password Generation Tasks. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers (UbiComp’18)*, Singapore, Singapore) (New York, NY, USA, 2018), ACM, pp. 786-793. doi:10.1145/3267305.3274127.
- Renaud, K., and **Zimmermann, V.** Ethical guidelines for nudging in information security & privacy. *International Journal of Human-Computer Studies* 120 (2018), 22–35. doi:10.1016/j.ijhcs.2018.05.011
- Renaud, K., and **Zimmermann, V.** Guidelines for ethical nudging in password authentication. *SAIEE Africa Research Journal* 109, 2 (2018), 102–118. doi:10.23919/SAIEE.2018.8531951
- Zimmermann, V.**, Bennighof, M., Edel, M., Hofmann, O., Jung, J., and von Wick, M. (2018) ‘Home, Smart Home’—Exploring End Users’ Mental Models of Smart Homes. In *Proceedings of the Mensch und Computer 2018-Workshopband, Dresden, Germany* (Bonn, Germany, 2018) R. Dachsel and G. Weber, Eds., Gesellschaft für Informatik, pp. 407-417. doi:10.18420/muc2018-ws08-0539.
- Zimmermann, V.**, Felscher-Suhr, U., and Vogt, J. Public Perceptions of Frankfurt Airport’s Value – A Survey Approach. *Journal of Air Transport Management* 67 (2018), 46-54. doi:10.1016/j.jairtraman.2017.11.005
- Zimmermann, V.**, Gerber, N., Kleboth, M., von Preuschen, A., Schmidt, K., and Mayer, P. The Quest to Replace Passwords Revisited – Rating Authentication Schemes. In *Proceedings of the 12th International Symposium on Human Aspects of Information Security & Assurance (HAISA 2018)*, Dundee, UK (Plymouth, UK, 2018), Plymouth University, pp. 38-48.
- Zimmermann, V.**, Shevkova, O., Keil, U., Schneider, M., and Baum, C. Psychische Belastung/Beanspruchung, Irritation und Beinahe-Unfälle bei Straßenbahn- und Stadtbahnfahrern/innen. In *Psychologie der Arbeitssicherheit und Gesundheit. Voneinander lernen und miteinander die Zukunft gestalten. 20. Workshop 2018*, R. Trimpop, J. Kampe, M. Bald, I. Seliger and G. Effenberger, Eds., Asanger, Kröning, Germany, 2018, pp. 257-260.
-

- Gerber, N., and **Zimmermann, V.** Security vs. privacy? user preferences regarding text passwords and biometric authentication. In *Proceedings of the Mensch und Computer 2017-Workshopband, Regensburg, Germany* (Bonn, Germany, 2017), M. Burghardt, R. Wimmer, C. Wolff, and C. Womser-Hacker, Eds., Gesellschaft für Informatik e.V., pp. 279-287, doi: 10.18420/muc2017-ws05-0405.
- Renaud, K., **Zimmerman, V.**, Maguire, J., and Draper, S. Lessons learned from evaluating eight password nudges in the wild. In *Proceedings of the LASER Workshop: Learning from Authoritative Security Experiment Results (LASER 2017), Arlington, VA, USA* (Berkeley, CA, USA, 2017), USENIX Association, pp. 25–37, isbn: 978-1-931971-41-6.
- Renaud, K., and **Zimmerman, V.** Enriched nudges lead to stronger password replacements... but implement mindfully. In *Proceedings of the 2017 Information Security for South Africa (ISSA), Johannesburg, South Africa* (New York, NY, USA, 2017), IEEE, pp. 1–9. doi:10.1109/ISSA.2017.8251779.
- Zimmermann, V.**, and Gerber, N. “if it wasn’t secure, they would not use it in the movies” -security perceptions and user acceptance of authentication technologies. In *Proceedings of the 5th International Conference on Human Aspects of Information Security, Privacy, and Trust (HAS2017), Vancouver, Canada* (Cham, Switzerland, 2017), T. Tryfanos, Ed., Springer, pp. 265–283. doi:10.1007/978-3-319-58460-7_18.
- Zimmermann, V.**, Henhapl, B., Gerber, N., and Enzmann, M. (2017). Promoting Secure Email Communication and Authentication. In *Proceedings of the Mensch und Computer 2018-Workshopband, Regensburg, Germany* (Bonn, Germany, 2018) M. Burghardt, R. Wimmer, C. Wolff and C. Womser-Hacker, Eds., Gesellschaft für Informatik, pp. 269-277. doi:10.18420/muc2017-ws05-0404.
- Zimmermann, V.**, Henhapl, B., Volkamer, M., and Vogt, J. Ende-zu-Ende sichere E-Mail-Kommunikation. *Datenschutz und Datensicherheit-DuD* 41, 5 (2017) 308-313. doi: 10.1007/s11623-017-0781-4.

2016 and earlier

- Santel, C. G., Gerber, P., Mehringskoetter, S., **Zimmermann, V. (née Schochlow)**, Vogt, J. and Klingauf, U. How Glider Pilots misread the FLARM Collision Alerting Display: a Laboratory Study. *Aviation Psychology and Applied Human Factors* 4, 2 (2014), 86 – 97. doi:10.1027/2192-0923/a000060.
- Zimmermann, V. (née Schochlow)**, Neumann, S., Braun, K., and Volkamer, M. Bewertung der GMX/Mailvelope-Ende-zu-Ende-Verschlüsselung. *Datenschutz und Datensicherheit-DuD* 40, 5 (2016) 295-299. doi:10.1007/s11623-016-0599-5
- Zimmermann, V. (née Schochlow)**, Santel, C. G., Weber, C., Vogt, J., and Klingauf, U. Kollisionsvermeidung im Luftsport: Eine experimentelle Studie der Mensch-Maschine-Schnittstelle eines populären Kollisionswarnsystems in der allgemeinen Luftfahrt. In M. Grandt and S. Schmerwitz (Eds). In *Proceedings of the 54th Fachausschusssitzung Anthropotechnik: Fortschrittliche Anzeigesysteme für die Fahrzeug- und Prozessführung, Koblenz, Germany* (Bonn, Germany, 2012), Deutsche Gesellschaft für Luft- und Raumfahrt, pp. 157-173.

Appendix C: Reference List of Authentication Schemes

The following tables provides references to the authentication schemes listed as examples in Figure 3, Figure 4 and Figure 5. The references given are either the developers of a scheme where available, or articles that describe and/ or evaluate the scheme. For general concepts and schemes for which no specific author can be identified, such as the password scheme, no reference is given.

Authentication Scheme	References
3D Gesture	[81]
Android Pattern Unlock	[114, 312]
Associative Questions	[246]
Blonder Scheme	[30]
Challenge Questions	[153, 154]
Cued Click Points	[54]
Déja Vu	[75]
Draw-a-Secret	[150]
GriDSure	[104]
Mnemonic Password	[180, 340]
Musical Password	[106, 179, 236]
Passfaces	[40, 229]
PassGo	[301]
Passlogix	[221]
Passobject Schemes	[287]
Passphrase	[162, 280]
PassPoints	[333]
PassShapes	[330]
Persuasive Cued Click Points	[51, 52]
Picture Password	[148]
PIN	general concept
Recall-a-Story	[188]
Tactile Password	[174]
Text password	general concept
Use your illusion	[130]
Visual Identification Protocol (VIP)	[67]
Weinshall	[328]

Table 7: Reference list of exemplary knowledge-based authentication schemes.

Authentication Scheme	References
DNA	[149]
Earshape Recognition	[341]
EEG	[1, 307]
Eye Movement	[169]
Face Recognition	[309]
Finger Nail Bed and Nail Plate Recognition	[176]
Fingerprint	[144, 145]
Gesture Recognition	[242, 265]
Hand Geometrics	[178]
Hand Vein Pattern	[177]
Heart Rate (Variability)	[6, 283]
Iris Recognition	[334]
Keystroke Dynamics	[159, 207]
Knuckle Shape	[177]
Odour	[28]
Retina Recognition	[35]
Signature Dynamics	[103, 195]
Speech Dynamics	[279]
Voice Recognition	[13]
Walk/Gait Recognition	[121]

Table 8: Reference list of exemplary biometric authentication schemes.

Authentication Scheme	References
Chip Card	[82]
ID Card	general concept
Ironkey	[168]
Key	general concept
Magnetic Stripe Card	[172]
MPAuth	[191]
One-Time Codes / TAN	[120]
Paper Card	general concept
Phoolproof	[227]
PhotoTAN	[220]
Physical Object	general concept
Pico	[135, 289]
QR/ Bar Code	[291, 335]
SMS TAN	[209]
Transponder	[181]
YubiKey	[343]

Table 9: Reference list of exemplary token-based authentication schemes.