Please do not use !?_ or your License Plate Number: Analyzing Password Policies in German Companies

Eva Gerlitz, *Fraunhofer FKIE;* Maximilian Häring, *University of Bonn;*
Matthew Smith, *University of Bonn, Fraunhofer FKIE*

https://www.usenix.org/conference/soups2021/presentation/gerlitz

This paper is included in the Proceedings of the
Seventeenth Symposium on Usable Privacy and Security.

August 9–10, 2021

978-1-939133-25-0

Open access to the Proceedings of the
Seventeenth Symposium on Usable Privacy
and Security is sponsored by

Please do not use !?_ or your License Plate Number: Analyzing Password Policies in German Companies

Eva Gerlitz*
Fraunhofer FKIE

Maximilian Häring*
University of Bonn

Matthew Smith
University of Bonn, Fraunhofer FKIE

Abstract

Password composition policies (PCPs) set rules that are intended to increase the security of user-chosen passwords. We conducted an online survey and investigated the employee-facing authentication methods of 83 German companies and the extracted 64 PCPs. We compared the password policies to recommendations proposed by institutions and related work. We found that many companies still require several character classes to be used as well as mandating regular password changes. Short and complex passwords are more often enforced than alternative mechanisms, such as minimum-strength requirements, that related work found more usable. Many of the policies were in line with recommendations given through the German Federal Office for Information Security (BSI). At the same time, there is high heterogeneity in the reported elements. Based on a selection of the main elements (password age, complexity, minimal length), at most seven out of the 64 PCPs are identical. The company size does not seem to play a significant role in the configuration of the PCPs.

1 Introduction

Passwords as a security measure are the daily reality of users working with computers, and even with technologies like FIDO2, they will likely stay for a while. It is well known that users sometimes choose weak passwords regarding their security effect. Websites and companies thus try to prevent this by using password composition policies (PCPs). These policies constrain the passwords users can choose, e.g., by preventing commonly chosen passwords. However, poorly chosen PCPs can be detrimental to usability and security [34]. A large body of work looks at PCPs in end user-facing websites, e.g., [15, 30, 31, 38], and how users cope with PCPs, e.g., [23, 26, 29, 33]. In this paper, we look at this topic from the view of those who manage PCPs. We conducted a survey with IT staff from 83 German companies. We focused on employee-facing PCPs since their passwords often protect accounts of great value for hackers (e.g., espionage or access to large amounts of user data).

To help companies with the creation of PCPs, organizations like the American NIST (National Institute of Standards and Technology) [22], OWASP (Open Web Application Security Project) [2] or the German BSI (Bundesamt für Sicherheit in der Informationstechnik, the German Federal Office for Information Security) [6] provide guidelines. To analyze if and how these guidelines affect the creation of PCPs, we surveyed what PCPs our participants used for company-wide user accounts or company email accounts and what information sources they used during the creation of the PCPs. We also surveyed what their experiences and perceptions of the PCPs is. We found a very heterogeneous set of PCPs with a surprising number of creative and unique PCP elements.

In this paper, we

- give an overview of the PCP landscape of 83 German companies.
- look at possible influences on and of PCPs.
- compare the identified PCP elements with recommendations.

The rest of the paper is structured as follows: First, we give an overview of relevant related work regarding PCPs and their effects on the resulting passwords, then we describe the methodology of the study, followed by the results, discussion, and directions for future work.

* These authors contributed equally to this work.

USENIX Symposium on Usable Privacy and Security (SOUPS) 2021.
August 8–10, 2021, Virtual Conference.

2 Related Work

There is a large body of literature on various aspects of password authentication. In the following section, we discuss previous work that is most relevant to our study. First, we give a short overview of the analysis of existing PCPs on websites, followed by the effects of different PCPs on end user behavior. Finally, we give a summary of guidelines given by organizations like NIST or the BSI.

2.1 Analysis of Existing Password Composition Policies

In 2010, Florêncio et al. [21] examined the password policies of 75 websites, including top, high and medium traffic sites as well as banks, universities and government sites. They calculated the minimum strength of each password policy using the cardinality of the minimum character set required and the minimum length given in the policy. Afterward, they analyzed if different characteristics correlate with stronger password policies but found no correlation between the website's size, the number of users, or the frequency of attacks. Instead, they noted a strong inverse correlation between password policy strength and sites that accept advertising and sponsored links. The authors hypothesized the necessity of those websites to have high usability to keep users on their site.

Mayer et al. [28] replicated and extended this study in 2016 by analyzing the password policies of the same websites as visited by Florêncio et al. and additionally investigating a corresponding sample of German websites. They noted that the average strength of the password policies had grown significantly in the US, except for websites that display third-party advertisements. In all samples, inverse correlation was found for users visiting a website with a clear competitor regarding their service. While comparing the password policies of German websites and those from the US, the authors noted a much smaller median of policy strength on German websites, with especially weak policies on banking websites.[1]

2.2 Effect of Password Policies on Usability and Password Strength

Komanduri et al. [27] studied password strength and user sentiment across four password composition policies in 2011. For this, they invited 5000 people to participate in an online study where participants had to create a password that they had to recall two days later. The policies requested a certain length (8 vs. 16 characters) either alone, with an additional complexity requirement, or the non-existence of dictionary words in the chosen password. They found that participants across all conditions used at least 2.2 digits, while symbols mainly were used if a policy requested to do so. Also, requiring a high complexity led to passwords with a higher entropy than other policies. At the same time, high complexity and the ban of dictionary words made password creation more complicated, with only 17.7% of participants being able to create a password in one try compared to the 52 to 84 % with other policies. The authors noted a correlation between storing passwords and the use of higher-entropy passwords. Of the four tested password composition policies, the one asking for at least 16 characters but not requiring anything else seemed to be the best trade-off between usability and security of the resulting passwords.

The same approach was followed by Kelley et al. [26] in 2012, Shay et al. in 2014 [33] and 2016 [34] and Tan et al. [35] in 2019. Kelley et al. [26] tested the effect of policies of different lengths and complexities as well as the presence of password blocklists, which varied in their size and complexity. They found that larger and more complex blocklists lead to stronger passwords. Shay et al. [34] examined 15 password policies by inviting 20,000 participants. Their password composition policies included policies that required only a minimal length or a length in combination with complexity or a certain number of words. The authors found that requiring a longer password with less complexity made it easier for participants to create and recall them while being less likely to be guessed. While experimenting with password blocklists, they noted that substring blocklists made passwords more secure without making recalling them more difficult. Policies that only requested minimal lengths were found to be usable; however, many of the resulting passwords were very weak. Additionally, the authors found the frequently used PCP consisting of a minimum of 8 characters and one character of each character class to be less usable and secure than some other tested policies. Based on their findings, the authors also gave recommendations for service providers regarding password composition policies.

Tan et al. [35] tested 21 policies, including composition requirements, blocklist requirements (using different lists and four different matching algorithms) and minimum-strength requirements (i.e., the number of guesses needed). During creation, a password meter showed compliance with the requirements and gave additional hints on how to increase the security once compliance was met. The study was completed by 6477 participants. The authors found that character-class requirements are annoying while simultaneously resulting in passwords that can be easily cracked using state-of-the-art password-cracking tools. They additionally saw usability differences when comparing different blocklists and found no benefit of requiring four character classes in addition to a large blocklist. They recommend using a minimum-strength requirement in combination with a length requirement and rate the benefit of minimum-strength higher than blocklists in protecting against offline attacks.

Ur et al. [37] asked 49 participants to create an account for

[1] It should be noted, though, that the login for costumers is protected by rate-limiting. Further actions need to be approved by a second factor [8].

fictitious banking, email, and news websites while thinking aloud to understand common password patterns and users' misconceptions about password strength. The authors found that while some weak passwords were created consciously, most were a result of misconceptions, e.g., that a "!" at the end makes a password more secure or that hard to spell words are securer than easy ones. Additionally, many participants demonstrated misconceptions regarding possible attacks, believing that personal data as passwords is secure as long as it is not known publicly. When confronted with policies that required the participants to add numbers or symbols to their password, which they had not included before, many simply appended one to their password.

Inglesant et al. [25] let 32 staff members of two different companies keep a password diary for one week and interviewed them regarding the details of each password. They found that the policies existing in 2010 were too complex, which, in the worst case, harmed the (organizational) productivity.

In a study of passwords collected from over 25.000 members of their university, Mazurek et al. [29] found the passwords of people who were annoyed by the complex password composition policy to be weaker.

In 2010, Zhang et al. [39] examined whether password expiration meets its intended purpose. For this, they analyzed a data set consisting of 7700 accounts. They found that 41% of the new passwords can be broken with knowledge about previous passwords for the same account within seconds, and 17% of the accounts can be broken into with five online password guesses.

Similarly, Habib et al. [24] found that 67% of the participants from an online survey self-reported creating their new password by modifying their previous one; most prominent was capitalizing a letter, which was done by 30%. Still, according to self-reports, regular password changes do not seem to lead to weaker passwords. 82% of their participants agreed that frequent password expiration secures accounts against unauthorized persons.

Using a more theoretical methodology, Shay and Bertino [32] presented an algorithm to simulate the effect of policies on security. As input, it takes details of the policy (e.g., length, per-character entropy, expiration), details of users (e.g., probability that a user remembers a seven-digit password after seeing it for the first time) and details of the service. With this, they offer administrators the possibility of testing a PCP concerning various properties of their organization.

Blocki et al. [14] presented an algorithm that takes a sample of users' preferred passwords as input. Based on this, it creates an ideal policy that maximizes the minimum entropy of the resulting distribution of passwords.

2.3 Official Recommendations

Table 2 (Appendix) shows the recommendations for password policies given by the American NIST [22] and the German BSI [5]. A few months after we conducted this study, the BSI changed their recommendations substantially in the area of PCPs. We will thus refer to their recommendations that were present during our study as "old", and the revised BSI recommendations as "new".

NIST published very specific recommendations, for example, regarding the minimum number of characters a password should have (minimal length), the minimum number of characters a password should be allowed to have (minimal maximal length), the number of character classes covered in a password (complexity), the time after which a password needs to be reset (maximal age), which characters should be allowed as part of the password (allowed characters) and which elements should be prohibited from usage (blocklist). On the other hand, both BSI recommendations consciously keep the recommendations vague to leave room for interpretation, for example, by stating that passwords of suitable quality should be chosen, without defining "quality". The BSI guidelines use as a basis the ISO 270012 and can be used as a help to implement it [12]. Additionally to the guideline, the BSI published implementation notes [11] with examples on how the policies could be built. The examples are based on a combination of minimal length and character classes, e.g., length of 20 to 25 and two character classes, or length of 8 to 15 and four character classes.

3 Methodology

To investigate the current state of password composition policies in German companies, we conducted a survey in late 2019 using Qualtrics [10].

The survey aimed at people who are responsible for PCPs in German companies. The questionnaire was offered in German and English since employees at this level can be from an international context.

3.1 Survey Design

Since our target audience is usually very busy and extremely hard to recruit for research purposes, we paid special attention to keeping the survey as short as possible. Thus, our survey was designed to take around ten minutes. Since the PCP is a sensitive piece of information and we were concerned that companies would not share them with us, we did not collect any information that could identify the company and any personal data from the person taking part in the study. While this would have been interesting data, we did not want to jeopardize either the company or our participants if our systems were breached.

Our survey, which can be found in Appendix A, was structured into three sections.

The first part asked whether the company uses a company-wide account per user that is centrally managed, e.g., for logging onto workstations, email, communications platforms, and the like. If this was the case, we asked for the authentication method(s) with which this account is secured, e.g., passwords, biometrics, or tokens. For those companies that did not have such central accounts, we opted to study the authentication methods for the company email accounts as we were fairly certain that most companies would have such a service. This way, we were able to include these companies and observe possible differences in these application areas.

The second section included questions regarding the authentication method(s) details, e.g., if a password composition policy is used, who created the PCP, and asked for the PCP itself. We encouraged participants to copy and paste their policy if they were allowed. Further, we asked for our participants' opinions on the authentication methods' security and usability.

The last section contained general information and demographic questions. As noted above, we collected only very minimal demographic information.

3.2 Survey Testing

Since we set ourselves a strict time limit for the survey, it proved challenging to formulate short enough questions to not slow down the survey but also unambiguous enough to gather useful data. The survey underwent five internal iterations, and we conducted a pilot study with the VP of Security of a large multi-national company and an administrator responsible for a small organization. We integrated the feedback from the pilot study into the final version of the survey.

3.3 Recruitment

We recruited our participants through several channels. The most effective channel by far was a newsletter sent by the BSI ($n = 69$ valid data sets of 83 valid data sets in total). We also recruited via contacts of two german digital associations (Bitkom [1] and Cyber Security Cluster Bonn [4]) and personal contacts.

Our survey was targeted at the person within the company responsible for the authentication system and the PCP. Since we had no way of contacting them directly, we clearly stated that only these people could fill out the survey and requested that the survey link be passed on to this person within the company. As we offered opt-out options, we believe that an accidental non-decision maker would have to have had malicious intent to affect the results negatively. In total, 110 participants took part. The participants were not compensated for their time.

3.4 Data Quality

Since we expected a heterogeneous set of PCPs and asked questions that are either sensitive or broad and thus may not apply to every participant, we included "Other", "I do not wish to make a statement", and "I do not know" options to questions (see also Section 3.6 and Section 6). This way, we wanted to prevent that the participants leave after facing a question that they could or did not want to answer.

Before analyzing the data, we checked for duplicate companies by using the company demographics and policies. We saw nothing to suggest that one company participated more than once. We also manually went through all the complete answers and excluded one participant who gave answers in the open texts, which led to the conclusion that they had not understood the previous questions. We also excluded 25 participants who had a completion rate lower than 50%. Most (21) of them closed the survey after answering whether there is a centrally managed account and, if so, what authentication methods can be used to log in.

In the end, we were left with 83 complete, valid data sets, and 77 policies.

3.5 Data Analysis

In our analysis, we separated the 77 password policies by their usage for a centrally managed account ($n = 64$, in the following called PWA for "Password Account") and those applying for email accounts ($n = 13$).

To analyze the PCPs, we used open coding as described by Corbin et al. [19]. Even though the policies mainly were enumerations of several elements and did not allow much room for interpretation (with few exceptions like "no *easy* passwords"), two researchers independently coded the policies to reduce errors. As suggested by Campbell et al. [16], we developed a code book by separately coding a small set of policies (n = 10) and comparing the codes. This then served as a base for future codes. After coding all the remaining policies independently, the codes were compared, and the inter-coder agreement was calculated using Cohen's kappa coefficient (κ) [18]. Our agreement was 81.48. A value above 0.75 is considered a good level of coding agreement [20]. We were able to resolve all conflicts.

We found a large set of possible properties concerning a PCP which we used to categorize each PCP based on its attributes (e.g., minimal length = 8 characters, minimal age = 1 day, maximal age = 90 days). We opted against calculating the strength of the PCPs as done by Florêncio et al. [21], and Mayer et al. [28], which only describes the theoretical size of the possibilities. It is also acknowledged by Florêncio et al. and Mayer et al. themselves that this is not a good metric to calculate the resulting password strength.

However, we discuss compliance with recommendations regarding PCPs from related work.

This study contributes data with an exploratory approach guiding to further research themes. Trends and interesting data were only very rarely tested on statistical significance to reduce the problems of multiple comparisons analysis.

When looking for statistical significance, we corrected the results with the Bonferroni–Holm method, also taking tests into account that we did but do not report. In the following sections, the stated p is that after correction. Percentages are reported rounded.

3.6 Ethics

The companies were asked details of their authentication methods and policies, which could give indications of vulnerabilities. To keep risks of exposure low, we did not collect any information that could identify a company or individual.

If participants included their company's name in one of the free text fields, we anonymized the answer before analyzing it. Additionally, the respondents were given an explicit option to answer questions with "I do not want to answer". Before the survey began, there was an introduction to the study, and participants had to consent.

The Research Ethics Board of our university reviewed and approved our study.

4 Results

In the following section, we will present the results of the survey. First, we present the demographics and the authentication methods used by the participating companies. This is followed by an analysis of the present password composition policies and their different components that respect to Table 2. We conclude this section with an overview of the potential impacts different authentication methods have.

4.1 Demographics

Table 3 (Appendix) shows the size (number of employees) of the participating companies ($n = 83$) and the number of desktop clients the participants had to handle.

We asked the participants what situations regarding their emails apply to their companies. 60 (72%) stated that employees can access their emails outside the company network. In 51 (61%) companies, emails can be accessed through a web login. In 28 (34%) cases, the employees do not need to know their password to access their emails, e.g., because of pre-configured mail clients.

On average, it took the participants 11 minutes to complete the survey.

4.2 General Authentication Setting

Of our 83 participants, 68 (81.93%) reported the use of company wide accounts of which all were secured at least with passwords. Ten (12%) companies additionally made use of biometric authentication (two face recognition, seven fingerprint and one palm vein recognition). One mentioned that face recognition is allowed on mobile devices. 29 (35%) participants stated that they use hardware tokens in addition to passwords. Eight (10%) participants reported they offer authentication with passwords, biometrics and tokens. Apart from this, two (2%) participants mentioned (device) certificates.

15 (18%) of the surveyed participants do not use company-wide accounts. These participants answered questions regarding their companies' email passwords. We will take a closer look at these policies in Section 4.3.2.

4.3 Password Composition Policies

In the following, all presented results only refer to PCPs used for the companies' user accounts (for regular employees), unless stated otherwise ($n = 68$).

63 (93%) of the participants stated that users are allowed to set their own account password. Two (3%) mentioned that the password is given to the user and cannot be changed by themselves. We could not find any standing out property of these two participants. In two (3%) cases, it is explicitly mentioned that an initial password is generated by the system and is changeable later; one company directly demands a change.

From the 68 participants who use company-wide accounts, we were able to extract 64 password composition policies. The remaining participants did not define a policy but gave a general description of how a policy could look. 59 (92%) of the policies get enforced technically. In two of the four companies, where this is not the case, participants mentioned that they use awareness trainings for their employees to counter the problem of common passwords.

Twenty-nine (45%) participants were part of the password policy creation process and 15 (23%) stated that the PCP was created by their predecessor.

As was expected based on the recruiting procedure, many (55%) participants who were part of the creation process of the password composition policy relied on the BSI as an influence in the PCP creation process. Figure 4 (Appendix) shows which other inspirations were used by our participants who were part of the creation process. Some other sources mentioned were ANSSI (French National Agency for the Security of Information Systems), ISO 27001, or PCI DSS (Payment Card Industry Data Security Standard). The option "Expert Panels" did not concern any specific panel but was given as a non-explicit, "consulting with experts".

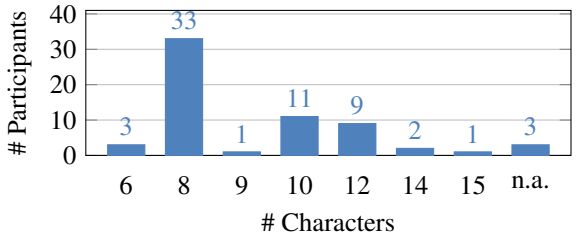

Figure 1: Minimal character length of passwords (PWA). "n.a." means no answer was given

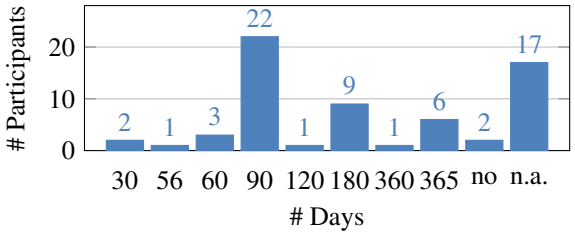

Figure 2: Password rotation cycle in days (PWA). "no" indicates participants who explicitly mentioned not using expiring passwords. "n.a." means no answer was given

4.3.1 PCP Components

In the following section, we present which elements were present in the PCPs that refer to companies' user accounts. For this, we follow the policy elements mentioned by official recommendations, as summarized in Table 2. An overview of the elements "minimal length", "password age" and "complexity" can be found in Table 4 in the Appendix.

Length Sixty-one (95%) companies use length requirements to ensure a secure password. While a minimum length is widespread (54 companies, 84%), some participants also mentioned fixed lengths (11%). However, the data for maximal and fixed length was not always clear, for example, in case of participants who stated: "Password length 8. [...]". In these cases, we count them as minimal and maximal length. In 33 (52%) companies, the participants mentioned eight characters to be their minimal password length. Eleven (17%) companies require 10 characters and 9 (14%) participants stated their minimal length to be 12 characters. Figure 1 gives an overview of how many participants mentioned which minimal length.

Password Complexity The complexity of a password depends on the number of character classes being used to create the password. For this, five different character classes can be used: uppercase letters (A-Z), lowercase letters (a-z), numbers (0-9), special characters including the space character, as well as the remaining Unicode characters that are alphabetic but not uppercase or lowercase (e.g., Chinese symbols). The latter one was only mentioned by one participant who indicated the complexity to be "Windows Password complexity", which includes all Unicode characters; so in the following, we will concentrate on the first four character classes.

Overall, 57 (89%) companies give constraints regarding the complexity. Seventeen (27%) companies require their users to build passwords using characters from all four classes. In one case, two characters of each class were demanded, one company requires a mixture between one or two characters per class, and in the other companies, one character of each class was sufficient.

Thirty-two (50%) participants mentioned that in order to fulfill their policy, characters from 3 of the four classes need to be present in a password. Fifteen (23%) specified which classes need to be covered, while 17 (27%) accepted a password as long as any three classes were present.

In five (8%) cases, participants stated that a complexity requirement is in place but did not specify, how this requirement looks.

Seven (7%) of our participants did not mention any requirements regarding complexity. However, none of them explicitly mentioned not using one.

Password Age and Password History As suggested by the BSI during the time of our study, 45 (70%) of our participants stated to force their users to change their passwords regularly. The top three rotation cycles were 90 days (34%), 180 days (14%) and 365 days (11%). Two participants explicitly mentioned not using a password expiration. While the percentage for 90 days (34%) is similar to what Habib et al. [24] found (28%), our peak at 180 and 365 days cannot be found in their sample. All password rotation cycles can be seen Figure 2.

Thirty (47%) participants reported a password history to prevent users from reusing previously used passwords. Twenty-seven (42%) of them check whether the passwords are identical, whereas three (3%) companies require significant changes, where it is, for example, not sufficient to increase a number within the old password to be accepted. Most mentioned was a history of 10 passwords (14 %) and 24 or 5 passwords (6% each). Figure 3 shows how many participants mentioned which number of previous passwords are stored.

When presented with the need to change their password and not be allowed to reuse a certain amount of their last passwords, users might counter this by changing their password several times in a row until they are allowed to use their original password again [7]. Because of this, companies use a minimal password age, as mentioned by 15 (23%) participants, so that users cannot change their password within this time period [7]. The minimal ages range from 24 hours (eleven companies) to 14 days (one company).

Allowed Characters NIST [22] recommends allowing all printing ASCII characters, the space character, and Unicode

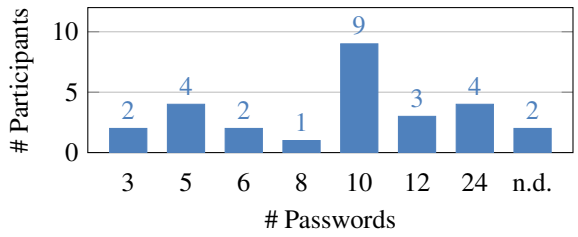

Figure 3: How many previously used passwords are not allowed to be reused (exact match). (PWA) "n.d." means no detail was given but a password history was mentioned.

characters for user-chosen passwords. As most participants only mentioned which characters are not allowed or which classes should be covered, it is hard to draw reasonable conclusions about the allowed character sets. However, one participant mentioned that a password needs to cover the "Windows Password complexity", which includes Unicode characters, e.g., "from Asian languages" [9].

Blocklists Twenty-six (41%) participants affirmed the question whether passwords were checked against common or leaked passwords. However, we did not ask for details, so we do not know how the comparison is made technically or which lists are used for this purpose.

Rarely Encountered We also found several constraints, which were only mentioned in at most three policies and were constraints to particular cases. We believe some of these are used since certain characters might break backend processing or serve as substitute for blocklists (e.g., to prevent passwords consisting of personal information such as nicknames). The atypical constraints included: (1) No colloquial language of any language, (2) No words of any language written backwards, (3) Certain special characters like € or umlauts, (4) Not more than 2 characters or sequences in series, (5) The last 20 passwords need to differ significantly from the new password, (6) Not more than 2 characters which appear in the same series in your name, (7) Not your license plate number.[2]

4.3.2 Additional Policies

In addition to PCPs, that define the user's chosen passwords on company accounts, some participants also mentioned additional PCPs, such as those for administrator accounts.

All participants who do not use company-wide accounts answered the questions regarding their email accounts.

We will present both extra sets of PCPs in the following paragraphs. Be reminded that both sets are excluded from the analysis above.

[2]While this was only mentioned by two participants, it is in fact mentioned in the implementation notes offered by the BSI [11].

General Three participants mentioned two different password composition policies. Two of them applied stricter rules if an account belonged to an administrator. In both cases, the minimal length was increased to 16 characters (while the regular accounts were required to use 12 respectively 10 characters).

One company requires its employees to use at least 20 character long passphrases for SSH and PGP/GPG keys.

Email Passwords Companies that do not use company-wide accounts for their employees were asked about their email passwords. We received 15 (18% of all responses) answers and were able to extract 13 password composition policies. Though we did not ask whether the PCPs are given by an email provider, six (46%) indicated that they were part of the creation process. The primary influence was the BSI, own knowledge, and expert panels (each one mentioned three times).

The median minimal length mentioned by participants in this group was 10 characters (range from 8 to 30 characters). Three participants mentioned very large minimal lengths: One (8%) needed 20 characters, and one (8%) only accepts passwords if they are 30 or more characters long. Additionally, the modes of the complexity were 3 and 4 character classes, and the median of the rotation cycle was 180 days (range from 180 to 365 days).

Interestingly, three of the email participants mentioned that employees generate their passwords with a password generator, whereas only one participant from the account group said so.

We also found one company that differentiated between regular and administrator accounts. They increased the minimal length from 12 to 16 characters and decreased the maximum age from 90 days for regular employees to 45 days for administrator accounts.

4.4 Effects of Authentication Methods

We asked participants how they rated password, biometric and token authentication concerning their security and usability impact and how often they encounter problems with the systems (e.g., forgotten passwords or lost hardware token). Additionally, we asked for their overall satisfaction with all used authentication methods combined. Following related work (Section 2), we assumed that unusable policies would lead to more problems and eventually to a lower satisfaction of the responsible person.

4.4.1 Influence on Security and Usability

Figure 5 (Appendix) shows what influences the use of the PCPs (n=64), biometrics (n=10), and hardware token (n=29) has on the sensed security and usability of the authentication system as well as how often problems arise for each method.

It can be seen, according to the participants, passwords lead to problems more often and form the lower bound of usability and security.

However, when comparing the scores, one has to keep in mind that biometrics and token were never used alone but always in combination with passwords. As we first asked for details about the password policies and only later for details of biometrics and token, the observed scores of biometrics and token may be compared to the security and usability of passwords. When looking at the number of problems arising from the authentication methods, tokens seem superior to biometrics and passwords. This is similar to the results of Abbott and Patil [13], who found token to have the second-highest UX rating when comparing different 2FA mechanisms.

We could identify a negative correlation between the reported impact on the security of the policy and the problems with passwords ($Spearman - \rho = -0.41294, p = 0.00938$), so the better the perceived security, the fewer perceived problems. The connection of the perceived user-friendliness and the problems is not statistically significant after correction ($Spearman - \rho = -0.3339, p = 0.05125$).

4.4.2 Satisfaction of Authentication Methods

We asked participants how satisfied they are with the present overall authentication system. Figure 6 (Appendix) shows the observed scores depending on whether a company offers passwords alone or in combination with a token. Due to small numbers, we excluded those participants using passwords, biometrics, and token (n = 8) and participants making use of passwords and biometrics (n = 2). It seems that participants using passwords in combination with tokens are more satisfied than participants using only passwords.

When concentrating on passwords, participants who stated that they created the password composition policy on their own had a median satisfaction of 4.00 (average: 3.41, sd: 0.95), while participants who were not part of the password policy creation had a median satisfaction of 3.00 (average: 2.94, sd: 0.92). This was not statistically significant (Mann–Whitney U test: $U = 624.5, p = 0.21824$).

However, a negative correlation can be reported between the number of problems with passwords and the satisfaction with the overall authentication system ($Spearman - \rho = -0.38639, p = 0.01409$); the fewer the problems, the more satisfied were the participants.

4.5 2-Factor Authentication (2FA)

If the participants stated to use more than one mechanism to authenticate accounts, we asked whether the methods are used in combination, for example, for 2-factor authentication. Of the 35 participants whose companies allow other methods than passwords, 18 stated that the methods are used as multiple factors and five companies leave the decision which method to use for authentication up to the employees.

5 Discussion

In the following section, we discuss the results. For this, we first compare the PWA PCPs to recommendations given by organizations like the BSI or NIST. This is followed by comparing the PCPs to recommendations and lessons learned from related work. After this, we analyze factors that might have influenced the PCPs when being created.

5.1 Compliance with Recommendations and Usability

5.1.1 NIST and BSI

In the following section, we discuss and compare the password composition policies and their elements with recommendations by NIST and the BSI, as seen in Table 2. We can only compare elements that are concretely covered by the corresponding guideline. The old BSI [5] recommendations do not make concrete recommendations regarding the length, complexity, or minimal password age and only require them to be “sufficient”, or “appropriate”. The same applies to the element “Quality”, which was introduced with the new BSI recommendations [6] and is expected to be “appropriate”. As the implementation notes [11] give concrete examples we compare those to the policies.

We again want to point out that the old BSI recommendation included a password expiration period and a complexity requirement during the time of the study. At the same time, NIST did not have such requirements.

Anecdotally, one participant mentioned following guidelines given by the PCI DSS (Payment Card Industry Data Security Standard) but does not consider it reasonable.

Length NIST recommends at least 8 characters for user chosen passwords. Thirty-three (52%) policies exactly fulfill this part and 57 (89%) require 8 or more characters. Three (5%) companies go against this recommendation and require only six characters.

We cannot make any statement on the minimal maximal length of 64 characters mentioned by NIST. No participant mentioned fulfilling this, but again, this does not mean that the companies only allow shorter passwords.

Complexity NIST recommends not setting complexity requirements in the form of character classes, while the old BSI recommendation suggested using a “sufficient” complexity. Fifty-seven (89%) participants mentioned complexity requirements, most often with the necessity to include three character classes in the password, so the vast majority of the companies were more in line with the old BSI recommendations than

with NIST. In the newer BSI [6] recommendation, the explicit mentioned need to establish mandatory rules of the requirement of some complexity was removed,[3] and only included that a password needs to have some level of "quality". We thus believe that our dataset can serve as a baseline for future studies observing the development of policies in companies.

Independent of what a policy requires, Tan et al. [35] found that users tend to include more character classes in their passwords.

Length and Complexity As mentioned before, the implementation notes of the BSI [11] give examples of how to combine length and complexity requirements. There are two we can use: First, a length requirement of 8 to 12 and 3 character classes. Twenty-seven (42%) participants match this. Second a length requirement of 20 to 25 and 2 character classes. None of the reported policies is equal to that example.

Password Age At the time the study was conducted, the BSI recommended regular password changes. Forty-five (70%) companies follow this recommendation and force users to change their passwords regularly. In contrast, NIST updated their recommendations in 2017, in line with results from research [39], and since advised against regular password change. Instead, a change should be forced whenever there is evidence of a compromise [22]. The new BSI recommendations are in line with this, but still recommend regular password changes, in case password cannot be checked against compromises [11].

Two (3%) participants explicitly mentioned foregoing password rotation.

Regular password changes cause users to develop mechanisms to make these changes less painful. These mechanisms, in turn, can lead to new rules added to the policies. One example of this are password histories, meaning users cannot reuse a certain number of their passwords. While most companies mentioned that the passwords as a whole are not allowed to be reused, three (4.69%) participants mentioned that significant changes are necessary. We did not ask for details of the technical implementation. The naive solution for this would be to store all passwords in plain text. This is obviously not a good idea, and systems that can be used to mitigate this issue have been proposed [17].

Another example of added rules due to a regular password change are minimal ages for passwords where users are not allowed to change their passwords within a specific period, to prevent them from cycling through and going back to their old password right away. However, this rule has the negative side-effect that users cannot change their passwords even if they assume it was compromised. Most likely, administrators could still make changes to the users' passwords instead of the users themselves, but it adds an additional step to the process. Minimal ages were mentioned by 15 (23%) participants.

Since regular password changes can cause a number of follow-up problems, it seems good that NIST and BSI no longer recommend this, and it will be interesting to see how and when companies adopt this change.

[3] Although the newer BSI [6] recommendations do not mention the complexity requirement as a mandatory rule, rules making use of complexity and length are used in the examples of the implementation notes [11].

5.1.2 Recommendations from Related Work

As summarized in Section 2, Shay et al. [34] tested 15 PCPs with over 20,000 participants. Tan et al. [35] tested 21 policies with over 6000 participants. Both research projects examined the strength of the policies by measuring the password guessability and the usability by looking at the user sentiment, the dropout rates, creation, and recall. Based on their findings, they gave recommendations for service providers at which we will have a detailed look in the following. It has to be noted that their recommendations are based on only two studies, and thus, further research is needed to explore the discovered aspects further. Additionally, not all elements that we discovered in the policies were studied or mentioned in the recommendations.

Avoid Using Length-Only Requirements Some of the tested policies by Shay et al. [34] only included length requirements. These policies seemed to be usable, and while some of the resulting passwords were quite strong, many others were very weak. Therefore, the authors suggest introducing further requirements, even if the minimal length is high. When looking at the three elements "minimal length", "complexity" and "maximum age", we only found two participants who only made statements about the length. One requested a minimal length of 6 characters but demanded the password not to include parts of the username or character/number sequences. The other company uses a minimal length of 12 characters and also forbids easy-to-guess passwords. It has to be researched what effect these additional elements have on the resulting passwords.

Do not Concentrate on Character Classes Tan et al. [35] included three policies that only contained length and complexity requirements. They found that the resulting passwords could be cracked with equal success rates, independent of the number of character classes required in the password. The authors conclude that users, at least when seeing a password meter, tend to choose longer and more complex passwords than required. When using a large blocklist, additional character class requirements do not seem to have any positive effect. The authors also found that with the use of a minimum-strength requirement, it is more usable to increase the length requirement or minimum-strength threshold compared to requiring more character classes to defend against offline attacks. In our sample, 61 of the 64 companies that also use a centrally managed user account and reported a pol-

icy mentioned using a character class requirement. Twenty-eight (46%) of them additionally mentioned to use blocklists.

If You are Using Comp8, Replace It. The PCP "comp8" included at least 8 characters, a complexity of 4, and no dictionary words. When testing this very common policy against others, Shay et al. [34] found three other PCPs to be more usable and more secure at the same time: "2class12" (minimal length of 12 characters, complexity 2), "3class12" (minimal length of 12 characters, complexity 3), and "2word16" (minimal length of 16 characters, at least 2 words). Looking at our sample, we find 10 (16%) participants who mentioned policies that match the "comp8" policy. Contrary, only four (6%) participants make use of "3class12" and no policy matches "2class12" or "2word16". It is interesting to see that while research offers good alternatives, many companies do not seem to adopt them.

Blocklists Some of the policies tested by Shay et al. [34] prohibited passwords from containing substrings from a prespecified list. They noted that this seemed to make creating a new password more difficult. However, this did not apply to the recall of passwords. Thus, they argue that including substring blocklists in the PCP is suitable if passwords do not expire too often. Tan et al. [35] found differences between different wordlists and matching algorithms (e.g. case-insensitive). They recommend not combining blocklist and character class requirements, especially when any password is rejected that exactly matches one included in a public leak.

Of the participating companies, 32 (50%) either confirmed whether they check the user-chosen passwords against commonly used passwords or mentioned to prohibit users from using certain sequences in their passwords as, for example, the company name, character/numerical sequences or dictionary words. Twenty-five (78%) of them additionally require regular password changes, that Shay et al. consider as an unfavorable combination [34].

Twenty-eight (46%) companies used a blocklist in combination with complexity requirement, which is not recommended by Tan et al. [35]. We do not have further insight into (a) what lists were used nor (b) how the comparison takes place (are numbers and digits included in a full string comparison / is the comparison case-insensitive?). Since different implementations seem to affect the security of the resulting passwords, but also on the time needed to create passwords, as found by Tan et al. [35], we believe future work should look at the actual implementation of blocklists within companies.

Minimum-Strength Requirements According to Tan et al. [35], minimum-strength requirements (i.e., number of guesses needed) are beneficial for password creation and, at the same time, result in strong and easy to remember passwords overall. When needing protection against offline attacks, the authors recommend their so-called "1c12+NN10" policy. In comparison to blocklist requirements, minimum-strength policies seem to combine better protection with improved usability.

None of our participants explicitly mentioned checking the passwords against a guessing attack. 5 participants required not to use "easy to guess" passwords. However, they did not specify if and how this is checked. It could thus be that the user is not supported fulfilling this rule. It remains to be seen if and how this relatively new knowledge about minimum-strength requirements will be applied in PCPs within the following years.

Ur et al. [37] recommended supporting users in developing good approaches for creating passwords and teaching them to correctly judge their decisions regarding password strength instead of creating PCPs that focus on character-class structures. Two of the participants explicitly mentioned awareness trainings for their employees regarding passwords. We did not ask about this explicitly, so there are probably more participants who, in fact, use awareness trainings.

Inglesant et al. [25] studied the effect of the PCP of two different companies, from which one had a very complex policy. It required their employees to use passwords consisting of 7 to 8 characters, a complexity of 3, no dictionary words, not exchanging o with 0 or i with 1, an expiration of 120 days, and significant changes to the 12 previous passwords. Many participants from this company expressed frustration and negative feelings towards password creation and recall. The authors note that the unusability arises from the combination of complexity, regular password changes, and the necessity to make significant changes to a previous password. Even though they conducted the study 10 years ago, we still saw many policies that have the potential to create user frustration as they consist of many elements that all have to be kept in mind when creating the password (cf. section 4.3.1). We hope that the revised BSI recommendations help in creating more user-friendly password composition policies.

5.2 Factors

Before conducting the survey, we had two themes that we assumed would influence the PCPs in companies:(1) **company size**: Tiefenau et al. [36] showed a significant difference between small and large organizations regarding formal update processes. We thus hypothesized that this difference could also be seen in other areas. (2) **the consulted recommendations (e.g., BSI or NIST)**, as they differ in details (cf. Section 2.3).

5.2.1 Size of Company

Table 3 shows that the amount of clients managed by the participant grows with the company size based on the number of employees. The amount of participants reporting a company-wide account does not increase with the number of employees.

Although the percentages at the two poles are clearly different, the small number of companies per bin does not allow us to draw conclusions. But we find that the use of a company-wide account does not seem like something extraordinary, even for small companies.

When separating the small and medium-sized companies (G_small, n=23) from the big companies (G_big, n=40) according to the definition from the EU [3], we do not see a big difference in the PCPs. They do not vary much in terms of the mean length (G_small = 9.8 vs. G_big = 8.95) or the mean maximum age (G_small = 156.8 vs G_big = 141). We also found no difference in the modes of the complexity when simplified to only the number of required character classes (1,2,3,4): G_small = 3 vs. G_big = 3.

5.2.2 Consulted Recommendations

As already touched in Section 4.3 and can be seen in detail in Figure 4, the participants reported using different sources as a basis for their policies. In 19 out of 29 cases, the participants listed more than one source of information besides 'Own Knowledge', statements from the 'Other' text field included. 2 participants reported only one institution besides 'Own Knowledge'.

Slightly more than half of the participants who took part in the creation of the PCP (16, 55%) used the BSI as a basis for the policy. Fourteen of those were recruited via the BSI newsletter, so that we might see a recruitment bias here.

When looking at the minimal length, we could not identify a big difference between those who use the BSI as an inspiration and were part of the creation process (G_BSI, n = 16) and those who do not (G_nBSI, n = 13): We found a mean minimal length of 10.5 characters for G_BSI and of 9.7 characters for G_nBSI. Nevertheless, we saw deviating means for the maximum age of passwords (225.00 days for G_BSI, 135 days for G_nBSI). The BSI guidelines suggested a regular password change. The high number could be based on the will to mitigate this measure but does not explain why the mean number of days for G_nBSI is so low.

Potentially, this points to the need to further separate the participants into groups classified by aspects like industrial sector or based on a risk evaluation.

The mode of the complexity is 3 for both groups.

5.2.3 Self-Made Policies

During analysis of the data, we found that a higher amount of policies that require a minimal length of 8 characters were mentioned by participants who were not part of the creation process. (31% in selfmade vs. 69% in not-selfmade). Overall, the mean minimal length of the not-self-made PCPs is also slightly lower (10.1 characters vs. 8.5 characters). A possible explanation might be the time a policy was created and official recommendations during those times.

Most (24 of 29) of the participants who helped create the policy reported having based the policy on 'Own Knowledge'. Future research should investigate this further as it is open to discuss whether this is a situation of "writing your own crypt library" or not. It should be noted that the used Software often provides options for policies, and so this probably heavily influences the policies (e.g., through default values).

PCPs of participants who were part of the creation process contained more elements when compared to the PCPs of participants who stated that they were not involved. This mainly included forbidden password elements as not using the username or license plate numbers. This may be an artifact of the methodology (e.g., recall bias).

While it is tempting to see a causal relationship here, be aware that the hypotheses around the self-made aspect are build from the data, so this phenomenon should be investigated in a separate study.

5.3 Heterogeneity

One of the main observation we made is that there is large heterogeneity in the landscape of PCPs, that we reported in Section 4.3. All PCPs used for company-wide accounts arrange in the area of 6-15 characters minimum (with a clear peak at a minimal character length of 8, Figure 1), an expiration range from 30 to 365 days (with two peaks at 90 and 180 days, Figure 2) and a password history of 3 to 24 passwords (with a peak at a history of 10, Figure 3). Yet we only found two PCPs that were identical concerning all mentioned elements as password history, forbidden words, etc. As this could be an artifact of our methodology, as discussed in section 6, we focused on the most common combinations of the three elements "complexity requirement", "minimal length" and "password rotation" as well as the number of PCPs that mentioned this combination (cf. Table 1). As mentioned in Section 4.3.1, some participants specified which character classes need to be covered in case they had the complexity requirement "3 out of 4". We merged all of them for Table 1 and only looked for the number of required character classes. Using the three elements, we found 41 out of 540 (9 different minimal lengths * 10 different maximum ages * 6 different numbers of required character classes (1 to 4, not stated, unspecific) possible combinations in the PWA policies

6 Limitations

Our study, like most surveys, has several important limitations.

The participants were invited over the newsletter sent by the BSI. Therefore, most of our participants are likely to be already interested in security-relevant topics. As the participation was voluntary and not remunerated, this effect is even heightened. As already stated in Section 3.3, we do not know for sure whether the participants were responsible for the

Complexity At least one char	Min. Length	Max. Age	Policies
3 classes	8	90 days	7
4 classes	8	90 days	5
3 classes	10	90 days	4
4 classes	8	n.a.	3
3 classes	8	180	3
4 classes	12	n.a.	2
4 classes	10	90	2
3 classes	8	60	2
3 classes	8	365	2
3 classes	8	n.a.	2
3 classes	12	365	2
		Sum:	34 of 64

Table 1: Most common PCP element combinations of complexity (at least one character of each class), minimal length and maximal age.
Policies gives the number of policies that showed this element combination. All other combinations only occurred once in the data set.

PCPs. It is also possible that multiple participants come from the same company. We searched the data for evidence but could not identify such.

As with any survey, participants may have selected the first answer that seemed appropriate without deeply thinking about their true beliefs and behavior. We tried to mitigate this the answers were randomized wherever meaningful. We also designed the survey to an expected time of 10 minutes.

The self-report and the recall bias has most likely affected our results. It is possible that participants answered in an effort to be more socially desirable. To reduce this bias, we kept the surveys anonymous and asked for as little demographic data as possible. It is also reasonable that they did not answer the free text answers in full detail, not because of bad faith but because they forgot it, misremembered parts, or because a detailed answer would have taken too much time to answer. This might have influenced the heterogeneity within the PCPs. To compensate for this in our analysis, we thus separated not mentioning something and explicitly excluding something. We believe this leads to a more optimistic estimation of the policies.

Due to differences in company structures, there are likely questions that are not in the direct area of responsibility of the participants among the diverse set of questions asked. We tried to mitigate this with a clear statement at the start of the survey to ensure that at least password policies are present in this area. If the participant did not know the answer, there was the answer option “I do not know”.

As a consequence of the heterogeneity of the PCPs, it is difficult to compare them. Many of our comparisons happen on large groups, only taking a few possible elements into account, disregarding their combination.

7 Conclusion

We conducted a survey to evaluate the current status quo of password composition policies in German companies. Our main finding is that there is high heterogeneity in the PCPs. While we cannot draw causal conclusions, it seems likely that the clashing nature of the national (BSI) and international guidelines (NIST) and the vague nature of the national BSI guidelines contribute to this heterogeneity. There is also a high prevalence of PCP elements that the research community, as well as NIST consider harmful, such as password expiry and character class requirements. When comparing the PCPs to recommendations made by related work, we found many that are very likely to be user-unfriendly. While the new BSI guidelines might fix the latter issue, they still require an analysis of the companies situation and leaf room for interpretation. Thus we do not believe that they will reduce the heterogeneity. While heterogeneity itself is not necessarily harmful and could even have security advantages, we doubt that the policy differences are based on conscious decisions and believe that they are more likely the product of a best-effort process. We recommend further research in this area and test whether more concrete recommendations lessen the burden on decision-makers within companies, who currently have to make many decisions that require a high level of domain knowledge.

8 Future Work

As mentioned in Section 2.3, the BSI changed their recommendation after the first survey was conducted. In future work, we will monitor how this guideline change affects the PCPs of companies.

Most of our findings indicate trends that need further research with other methodologies to validate them. We encountered several PCP elements that can not easily be enforced technically. We plan to examine whether custom enforcement mechanisms were implemented or whether the hope is that users follow the instructions even though they are not enforced. It is also open how users perceive the difference. While most policies were found to be similar in the big picture, they differed in their details. While this may be for good reasons, it also shows how diverse the landscape is, and further research is needed to see whether this is needed or if the benefit of one usable policy is higher.

The surveyed participants were all employees of German companies. Further research has to be conducted to validate the findings across cultures and study the influence of different local recommendations.

We were able to show a difference in the PCPs reported by their creators compared to PCPs created by somebody else than the participant. Still to be researched is the process of creating this policy, how big the personal factor is (and should be), and what role official recommendations play.

Acknowledgments

This work was partially funded by the Werner Siemens Foundation.
We thank Christian Tiefenau, Dirk Backofen, and Alexander Häring for their help, domain knowledge, and input. We thank our reviewers and shepherd, who helped improve the paper a lot, and all participants who took the time to answer our questions.

References

[1] About | Bitkom e.V. https://www.bitkom.org/EN/About-us/About-us.html. Accessed: May 31, 2021.

[2] Authentication cheat sheet. https://cheatsheetseries.owasp.org/cheatsheets/Authentication_Cheat_Sheet.html. Accessed: May 31, 2021.

[3] COMMISSION RECOMMENDATION of 6 May 2003 concerning the definition of micro, small and medium-sized enterprises. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32003H0361. Accessed: May 31, 2021.

[4] Cyber Security Cluster Bonn. https://cyber-security-cluster.eu/. Accessed: May 31, 2021.

[5] IT-Grundschutz Compendium - Final Draft, 1 February 2019. https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Grundschutz/International/bsi-it-gs-comp-2019.pdf?__blob=publicationFile&v=1. Accessed: May 31, 2021.

[6] IT-Grundschutz-Kompendium Februar 2020. https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Grundschutz/Kompendium/IT_Grundschutz_Kompendium_Edition2020.pdf?__blob=publicationFile&v=1. Accessed: May 31, 2021.

[7] Minimum password age. https://docs.microsoft.com/en-us/windows/security/threat-protection/security-policy-settings/minimum-password-age. Accessed: February 27, 2020.

[8] Nur fünf Zeichen fürs Banking-Passwort? https://www.heise.de/-4935773. Accessed: June 02, 2021.

[9] Password must meet complexity requirements. https://docs.microsoft.com/en-us/windows/security/threat-protection/security-policy-settings/password-must-meet-complexity-requirements. Accessed: May 31, 2021.

[10] Qualtrics. https://www.qualtrics.com. Accessed: May 31, 2021.

[11] Umsetzungshinweise zum Baustein: ORP.4. Identitäts- und Berechtigungsmanagement. https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Grundschutz/Umsetzungshinweise/Umsetzungshinweise_2021/Umsetzungshinweis_zum_Baustein_ORP_4_Identitaets_und_Berechtigungsmanagement.pdf?__blob=publicationFile&v=1. Accessed: May 31, 2021.

[12] Zuordnungstabelle ISO zum modernisierten IT-Grundschutz. https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Grundschutz/Kompendium/Zuordnung_ISO_und_modernisierter_IT_Grundschutz.html?__blob=publicationFile&v=1. Accessed: May 31, 2021.

[13] Jacob Abbott and Sameer Patil. How mandatory second factor affects the authentication user experience. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pages 1–13, New York, NY, USA, 2020. Association for Computing Machinery.

[14] Jeremiah Blocki, Saranga Komanduri, Ariel Procaccia, and Or Sheffet. Optimizing password composition policies. In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*, EC '13, page 105–122, New York, NY, USA, 2013. Association for Computing Machinery.

[15] Joseph Bonneau and Sören Preibusch. The Password Thicket: Technical and Market Failures in Human Authentication on the Web. In *WEIS*, 2010.

[16] John L. Campbell, Charles Quincy, Jordan Osserman, and Ove K. Pedersen. Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociological Methods & Research*, 42(3):294–320, 2013.

[17] Rahul Chatterjee, Joanne Woodage, Yuval Pnueli, Anusha Chowdhury, and Thomas Ristenpart. The typtop system: Personalized typo-tolerant password checking. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, CCS '17, pages 329–346, New York, NY, USA, 2017. Association for Computing Machinery.

[18] Jacob Cohen. A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1):37–46, 1960.

[19] Juliet M. Corbin and Anselm Strauss. Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology*, 13(1):3–21, 1990.

[20] Joseph L. Fleiss, Bruce Levin, and Myunghee C. Paik. *Statistical methods for rates and proportions*. john wiley & sons, 2013.

[21] Dinei Florêncio and Cormac Herley. Where do security policies come from? In *Proceedings of the Sixth Symposium on Usable Privacy and Security (SOUPS)*, New York, NY, USA, 2010. Association for Computing Machinery.

[22] Paul A. Grassi, James L. Fenton, Elaine M. Newton, Ray A. Perlner, Andrew R. Regenscheid, William E. Burr, Justin P. Richer, Naomi B. Lefkovitz, Jamie M. Danker, Yee-Yin Choong, Kristen K. Greene, and Mary F. Theofanos. NIST Special Publication 800-63b: Digital Identity Guidelines. https://doi.org/10.6028/NIST.SP.800-63b. Accessed: May 25, 2021.

[23] Hana Habib, Jessica Colnago, William Melicher, Blase Ur, Sean Segreti, Lujo Bauer, Nicolas Christin, and Lorrie Cranor. Password creation in the presence of blacklists. *Proc. USEC*, page 50, 2017.

[24] Hana Habib, Pardis E. Naeini, Summer Devlin, Maggie Oates, Chelse Swoopes, Lujo Bauer, Nicolas Christin, and Lorrie F. Cranor. User behaviors and attitudes under password expiration policies. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS) 2018*, pages 13–30, Baltimore, MD, August 2018. USENIX Association.

[25] Philip G. Inglesant and Martina A. Sasse. The true cost of unusable password policies: password use in the wild. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, page 383–392, New York, NY, USA, 2010. Association for Computing Machinery.

[26] Patrick G. Kelley, Saranga Komanduri, Michelle L. Mazurek, Richard Shay, Timothy Vidas, Lujo Bauer, Nicolas Christin, Lorrie F. Cranor, and Julio Lopez. Guess again (and again and again): Measuring password strength by simulating password-cracking algorithms. In *2012 IEEE Symposium on Security and Privacy*, pages 523–537. IEEE, 2012.

[27] Saranga Komanduri, Richard Shay, Patrick G. Kelley, Michelle L. Mazurek, Lujo Bauer, Nicolas Christin, Lorrie F. Cranor, and Serge Egelman. Of passwords and people: Measuring the effect of password-composition policies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 2595–2604, New York, NY, USA, 2011. Association for Computing Machinery.

[28] Peter Mayer, Jan Kirchner, and Melanie Volkamer. A second look at password composition policies in the wild: Comparing samples from 2010 and 2016. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS) 2017*, pages 13–28, Santa Clara, CA, July 2017. USENIX Association.

[29] Michelle L. Mazurek, Saranga Komanduri, Timothy Vidas, Lujo Bauer, Nicolas Christin, Lorrie F. Cranor, Patrick G. Kelley, Richard Shay, and Blase Ur. Measuring Password Guessability for an Entire University. In *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security*, CCS '13, pages 173–186, New York, NY, USA, 2013. Association for Computing Machinery.

[30] Sören Preibusch and Joseph Bonneau. The password game: Negative externalities from weak password practices. In *International Conference on Decision and Game Theory for Security*, pages 192–207. Springer, 11 2010.

[31] Tobias Seitz, Manuel Hartmann, Jakob Pfab, and Samuel Souque. Do Differences in Password Policies Prevent Password Reuse? In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI '17, pages 2056–2063, 05 2017.

[32] Richard Shay and Elisa Bertino. A comprehensive simulation tool for the analysis of password policies. *International Journal of Information Security*, 8(4):275–289, 08 2009.

[33] Richard Shay, Saranga Komanduri, Adam L. Durity, Phillip Huh, Michelle L. Mazurek, Sean M. Segreti, Blase Ur, Lujo Bauer, Nicolas Christin, and Lorrie F. Cranor. Can long passwords be secure and usable? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 2927–2936, New York, NY, USA, 2014. Association for Computing Machinery.

[34] Richard Shay, Saranga Komanduri, Adam L. Durity, Phillip S. Huh, Michelle L. Mazurek, Sean M. Segreti, Blase Ur, Lujo Bauer, Nicolas Christin, and Lorrie Faith Cranor. Designing password policies for strength and usability. *ACM Transactions on Information and System Security (TISSEC)*, 18(4):13, May 2016.

[35] Joshua Tan, Lujo Bauer, Nicolas Christin, and Lorrie F. Cranor. Practical recommendations for stronger,

more usable passwords combining minimum-strength, minimum-length, and blocklist requirements. In *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*, CCS '20, page 1407–1426, New York, NY, USA, 2020. Association for Computing Machinery.

[36] Christian Tiefenau, Maximilian Häring, Katharina Krombholz, and Emanuel von Zezschwitz. Security, availability, and multiple information sources: Exploring update behavior of system administrators. In *Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020)*, pages 239–258. USENIX Association, August 2020.

[37] Blase Ur, Fumiko Noma, Jonathan Bees, Sean M. Segreti, Richard Shay, Lujo Bauer, Nicolas Christin, and Lorrie F. Cranor. "I Added'!'at the End to Make It Secure": Observing Password Creation in the Lab. In *Eleventh Symposium On Usable Privacy and Security (SOUPS) 2015*, pages 123–140, Ottawa, July 2015. USENIX Association.

[38] Ding Wang and Ping Wang. The emperor's new password creation policies. In *European Symposium on Research in Computer Security*, pages 456–477. Springer, 11 2015.

[39] Yinqian Zhang, Fabian Monrose, and Michael K. Reiter. 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 '10, pages 176–186, New York, NY, USA, 2010. Association for Computing Machinery.

Appendix

A Survey

Consent

Welcome!

Thank you for taking time to participate in our study. The study is conducted by the team of Prof. Dr. Matthew Smith at the University of Bonn.

In this Study we want to find out more about the current state of authentication methods, in particular password policies in various companies. We will not ask for any personal information or data that could identify your company. Further, we will only report anonymous aggregated information. The goal of our research is to identify the needs of industry and develop supporting measures to increase IT-security. With your participation you will make an valuable contribution to this goal.

The survey is addressed to persons who are responsible for authentication and password policies in companies.

The survey will take 5-10 minutes, is voluntary and can be canceled at any time. If you have any questions, please contact -mail-.

By continuing with the study, you confirm that you are at least 18 years old and consent to your data being used anonymously. As the data is collected anonymously, it is not possible to delete any data after taking the survey.

Accounts

- Is there a company-wide account per user, that is managed centrally? (E.g., for logging into the workstation, communication platform, email or the like.)
  Yes / No
  - If yes:
    * What can this account be used for? (Multiple answers possible.)
      Email / Workstations / Communication platform (SharePoint, Slack etc.) / VPN into corporate network / Access to shared corporate data (e.g., Active Directory) / Other: [Free text]
    * Which methods can be used to log in? Please check the applicable.
      Password or PIN / Biometrics (e.g., Fingerprint, Face recognition) / Hardware Token (e.g. Smartcard, Token, Smartphone)
    * In there any other method in use that it not listed?
      Yes, the following: [Free text]/ No
    * You stated, that there are several methods in use which enable your employees to log in. Are the methods used in combination (e.g., 2FA)?
      Yes, the methods are used in combination (2FA) / No, the employees can choose one of the methods / Other: [Free text]/ I do not know / I do not wish to make a statement

Passwords

You stated that there is no company-wide account with wich the employees can log into several services.
The following questions regard the email accounts of your employees and their passwords (Webmail, imap, pop, etc.).
Or
You stated, that your employees use passwords/PINs to log in. The following questions regard these passwords/PINs.

- How are passwords handled?
  Users can choose them themselves / Passwords are created by a system, and users cannot change them / I don't want to make a statement / Other: [Free text]

- What specification (also called password policy) do passwords need to fulfill (e.g., at least x characters, new password needs to be selected after x days, etc.)

  This question is the main focus of our research. Please be as detailed as possible. If possible and allowed, please copy your specification into the following text box. At this point we want to remind you, that the data is managed anonymously. It will not be possible to identify your company.
  [Free text]

- Are these specifications enforced by the system?
  Yes / No / There are no specifications / I do not know / I do not wish to make a statement / Partially:[Free text]

- Optional: What reasons spoke against the introduction of a password policy?
  [Free text]

- Are users prevented from picking passwords that belong to the most common passwords?
  Yes / No / Other:[Free text] / I do not know / I do not wish to make a statement

- Who created the specifications (password policies) for the passwords?
  Myself / My predecessor / Somebody else: [Free text]/ I do not know / I do not wish to make a statement / There are no specifications

- What are the specifications based on? (Multiple answers possible.)
  Own knowledge / Expert panels / Exchange with other companies / NIST (National Institute of Standards and Technology) / BSI (Bundesamt für Sicherheit in der Informationstechnik) / OWASP (Open Web Application Security Project) / Other: [Free text]/ I do not know / I do not wish to make a statement

- How do the password policies impact the user-friendliness of the authentication system?
  1: Very negative – 5: Very positive

- How do the password policies impact the security of the authentication system?
  1: Very negative – 5: Very positive

- How often do passwords cause problems in your company (e.g., forgotten passwords, etc.)?
  1: Very rarely – 5: Very often

- Is there a policy which specifies how the passwords are stored in the system (hash function, length of the salt, etc)?
  Yes / No / I do not know / I do not wish to make a statement

- Is there a process which initiates an update of the policy on how to store passwords?
  Yes / No / I do not know / I do not wish to make a statement

- Optional: How are stored passwords protected? We are particularly interested in the hash and salt functions which are used.

  We want to remind you that the data is gathered anonymously and we are not able to link it to your company.
  [Free text]

Biometrics

You stated, that your employees use biometrics to log in. The following questions regard this method.

- What kind of biometrics are in use?
  Fingerprint / Iris / Face recognition / Other: [Free text]/ I do not wish to make a statement

- How does the biometric authentication impact the user-friendliness of the authentication system?
  1: Very negative – 5: Very positive

- How does the biometric authentication impact the security of the authentication system?
  1: Very negative – 5: Very positive

- How often does the use of biometric authentication cause problems?
  1: Very rarely – 5: Very often

- Optional: Do you wish to provide us with additional information about this topic?
  [Free text]

Hardware Token

You stated, that your employees use a hardware token to authenticate. The following questions regard this token.

- Does the token support FIDO2?
  Yes / No / I am not sure / I do not wish to make a statement

- How does the token impact the user-friendliness of the authentication system?
  1: Very negative – 5: Very positive

- How does the token impact the security of the authentication system?
  1: Very negative – 5: Very positive
- How often does the usage of the token cause problems?
  1: Very rarely – 5: Very often
- Optional: Do you wish to provide us with additional information about this topic?
  [Free text]

Demographics

- Please check the conditions which apply to your company. (Multiple answers possible.)
  There are employees who can access their emails outside the company network /
  There are employees who can access their emails using a weblogin /
  There are employees who do not need to know the password for accessing their emails, e.g., as the email-client is pre-configured
- Is there any additional security for emails? (e.g., encryption in combination with a smartcard)
  Yes, obligatory / Yes, voluntary / No / I do not wish to make a statement
- How many employees work in your company?
  1-9 / 10-49 / 50-249 / 250-499 / 500-999 / 1000 or more / Not sure / I do not wish to make a statement
- How many desktop clients do you manage?
  1-9 / 10-49 / 50-249 / 250-499 / 500-999 / 1000 or more / Not sure / I do not wish to make a statement
- How many employees in your company work full time on IT-security topics?
  0 / 1 / 2-5 / 6-10 / 11-20 / 21 or more / Not sure / I do not wish to make a statement
- How satisfied are you with your authentication system?
  1: ☹ – 5: ☺
- Has this questionnaire motivated you to update parts of your authentication system in the near future? If yes, which parts?
  Password Policies / Security measures for stored passwords / Adding biometrics / Adding hardware token / No / Other: [Free text]

B Additional Figures and Tables

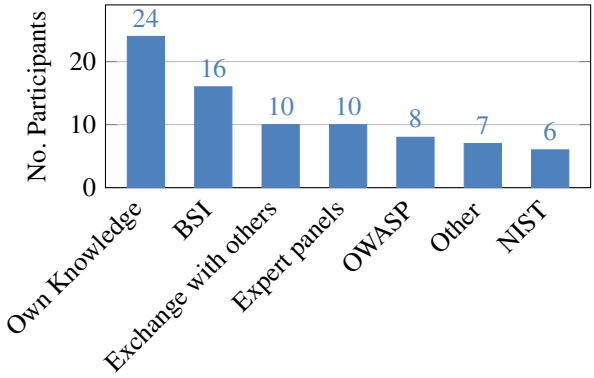

Figure 4: Sources of inspirations for the password composition policies reported from participants who took part in the creation (n=29, PWA). Multiple answers were possible.

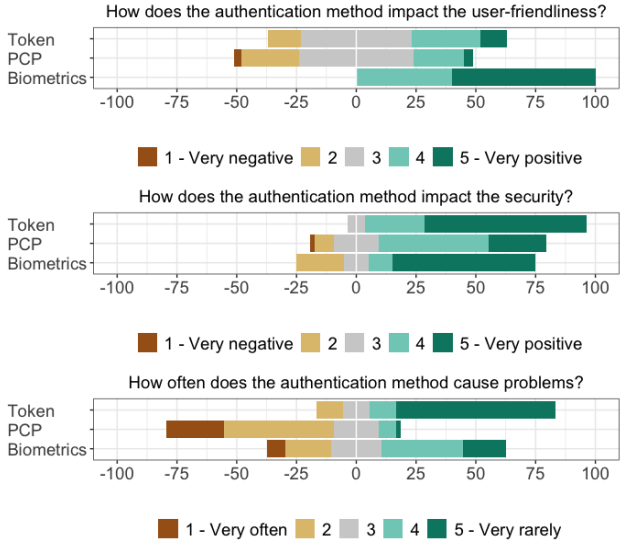

Figure 5: Impact of the different authentication methods: token ($n = 29$), the PCP ($n = 64$) and biometrics ($n = 10$) on the perceived user-friendliness, security and frequency of problems. Only PWA policies are used for the figures.

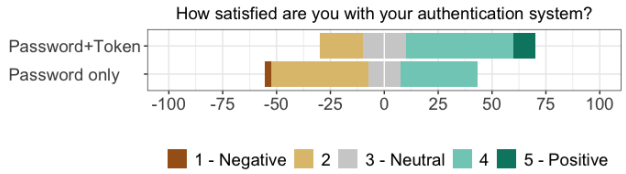

Figure 6: Satisfaction of participants with their overall authentication system, depending on methods in use: Passwords only (n = 33) and passwords in combination with token (n = 20). Numbers on the x-axis are percentages. Only PWA policies are used for this figure.

Policy Elements	**NIST** (2020) [22]	**BSI Old** (2019) [5]	**BSI New** (2020) [6]
Quality	-	-	Appropriate
Minimal length	8	Sufficient	-
Minimal maximal length	64	-	-
Complexity	Advised against	Sufficient	-
Maximal age	Advised against	Appropriate	-
Allowed characters	All ASCII & Unicode characters	-	-
Blocklist	At least: - Leaked passwords - Dictionary words - Repetitive or sequential characters - Context-specific words	-	- Easy to guess - Common passwords - Reused passwords

Table 2: Recommendations for password policies by different organizations, split by their elements. The BSI revised their recommendation in 2020.

		No. of Participants		**No. of Participants per No. of Managed Clients**						
		Company-wide	**Email**	**0-9**	**10-49**	**50-249**	**250-499**	**500-999**	**>= 1000**	**n.a.**
Company size	1- 9	5	3	6	1					
	10 - 49	7	5	2	9					1
	50 - 249	12	1			11	1			1
	250 - 499	7	2			1	7			1
	500 - 999	6	1					6	1	
	>= 1000	30	3	2				5	23	3
	No answer	1	0							

Table 3: Number of participants depending on the Company Size and Number of Managed Clients. Empty fields indicate 0. n.a. = No answer

	Element	Account Policies n= 64	Mail Policies n= 13	Additional Policies n = 4
Minimal Length	6	3	2	-
	8	33	4	-
	9	1	-	-
	10	11	2	-
	12	9	1	-
	14	2	1	-
	15	1	-	-
	16	-	-	3 (Admin)
	20	-	1	1 (Passphrases)
	30	-	1	-
	Unspecific	1	-	-
	N.A.	3	1	-
	Any minimal length	61	12	4
Maximal Age (Days)	30	2	-	-
	42	-	1	-
	45	-	-	1 (Admin)
	56	1	-	-
	60	3	-	-
	90	22	4	-
	120	1	-	-
	180	9	2	-
	360	1	-	-
	365	6	1	-
	Explicitly not	2	-	-
	N.A.	17	5	3
	Any maximal age	45	8	1
Character Classes (>= 1 each)	Special character	1	-	-
	2 (Letter, Digit)	-	1	-
	2 (Digit, Special)	2	-	-
	3 (Capital, Digit, Special)	3	-	-
	3 (Capital, Lowercase, Special)	3	-	-
	3 (Capital, Lowercase, Digit)	7	1	1 (Admin)
	3 (Letter, Digit, Special)	2	-	-
	Any 3 out of 4	16	2	-
	Any 3 out of 5 (incl. Unicode)	1	-	-
	4 (Capital, Lowercase, Digit, Special)	15	3	1 (Admin)
	4 (at least 2 each)	1	1	-
	4 (Capital and Lowercase: at least 1; Digit+Special: at least 2)	1	-	-
	Unspecific	5	1	1 (Admin)
	N.A.	7	4	1
	Any character class requirement	57	9	3

Table 4: Number of policies with the different elements of “minimal length”, “maximal age” and “character classes”. The four standard character classes are: Lowercase, Capital, Digit, and Special character.