"As soon as it's a risk, I want to require MFA": How Administrators Configure Risk-based Authentication (Extended Version)

Philipp Markert, Theodor Schnitzler, Maximilian Golla[⋆], and Markus Dürmuth[‡]
Ruhr University Bochum, ⋆ Max Planck Institute for Security and Privacy, ‡ Leibniz University Hannover

Abstract

Risk-based authentication (RBA) complements standard password-based logins by using knowledge about previously observed user behavior to prevent malicious login attempts. Correctly configured, RBA holds the opportunity to increase the overall security without burdening the user by limiting unnecessary security prompts to a minimum. Thus, it is crucial to understand how administrators interact with off-the-shelf RBA systems that assign a risk score to a login and require administrators to configure adequate responses.

In this paper, we let $n = 28$ system administrators configure RBA using a mock-up system modeled after Amazon Cognito. In subsequent semi-structured interviews, we asked them about the intentions behind their configurations and experiences with the RBA system. We find that administrators want to have a thorough understanding of the system they configure, show the importance of default settings as they are either directly adopted or depict an important orientation, and identify several confusing wordings. Based on our findings, we give recommendations for service providers who offer risk-based authentication to ensure both usable and secure logins for everyone.

1 Introduction

Password-based authentication is still the dominant form of user authentication, despite severe weaknesses such as phishing attacks [40, 48], password reuse attacks [14, 23], and their guessability [46, 56]. Password alternatives such as biometric authentication [34, 66], graphical passwords [6, 55], or security keys [12, 19] all have their own set of drawbacks that so-far have prevented their widespread adoption [9, 27].

To improve user security, services deployed additional protection mechanisms to reinforce passwords, for example, by using *multi-factor authentication* (MFA) [13, 22, 30], proactive *password-reuse checks* [36, 44, 52], and *risk-based authentication* (RBA) [16, 20, 63]. Several authorities, such as the NCSC [41], NIST [24], and others [7], all mention risk-based authentication as one of the key concepts to minimize account compromises.

RBA is a method for strengthening user authentication on the server's side without involving the user (except for rare cases). Thus, it offers the potential to increase the security of accounts without burdening the legitimate user. However, RBA comes at the cost of being a privacy-invasive technique that requires login behavior monitoring and client-side fingerprinting [8, 65]. At the moment of password entry, RBA monitors a variety of signals, such as the source IP, user-agent, login time, and further information about the user's machine, e.g., obtainable via client-side fingerprinting. This information is then compared with the user's profile from past logins, as well as profiles from typical attacks. Based on this information, a risk level is computed [20, 26].

The configuration of an RBA system requires administrators to decide how the system should treat logins with different risk levels. We consider this a non-trivial configuration task as it interferes with usability and security requirements that directly impact the user. In this work, we study how administrators interact with configuration interfaces for RBA. We focus on professionals not specialized in the administration of RBA, which we assume is rather common in small and medium-sized enterprises. To the best of our knowledge, our work is the first to study the rationale behind configuring RBA systems on the administrators' side. Thus, we keep our research exploratory and follow three broad research questions. RQ1: *How do administrators configure RBA?* (e.g., risk-level behavior, when and how to notify), RQ2: *Which obstacles and misunderstandings do they encounter?*, and RQ3: *What*

USENIX Symposium on Usable Privacy and Security (SOUPS) 2022.
August 7–9, 2022, Boston, MA, United States.

is the impact of previous exposure to other RBA systems and how do different requirements influence administrators?

In our two-part study, we assigned $n = 28$ administrators a configuration task for adjusting risk level behavior and RBA notification settings in an enterprise scenario. The configuration tool they worked with resembled the look-and-feel of Amazon Cognito, the system that *Amazon Web Services* (AWS) offers to its customers. Subsequently, we interviewed participants about their intentions behind the configurations they made, their interaction with the configuration system, and potential obstacles they encountered while completing their task. To facilitate the recruitment of system administrators from different continents, the study was conducted online using a video conferencing tool and an online web interface accessible to the participants.

Our results suggest that system administrators want to deeply understand how risk-based authentication systems work in order to be able to make informed decisions. For example, the tool we used hid some complexity behind generic phrases such as *low*, *medium*, and *high risk*, which was criticized by several participants. Additionally, our study identifies issues with room for improvement and other topics to be explored by future research—both in more detail and in a larger variety of RBA configuration systems. In summary, our paper makes the following key contributions:

- Through an in-depth qualitative evaluation of interviews, we complement existing knowledge about risk-based authentication by providing insights into administrators' decision processes.
- Our study shows that system administrators desire detailed information about risk levels and the ability to make fine-grained configurations in order to ensure appropriate risk level behavior.
- Our findings unveil several issues to be explored by future research, while at the same time indicating first recommendations for service providers to ensure usability and security of RBA systems.

2 Related Work

In this paper, we study how administrators configure risk-based authentication. Since there are no other studies to the best of our knowledge, we align this section along prior work about RBA and studies focusing on system administrators.

Risk-based Authentication In 2010, Google added a new feature to protect their users from suspicious account activity [16], and while, e.g., Facebook also stated to employ risk-based authentication [45], not much was publicly known about its inner working. In 2016, Microsoft started to offer risk-based conditional access to its Azure AD customers and supported risk events like unfamiliar locations, impossible travel, IP addresses with suspicious activity, and users with leaked credentials [49]. At the same time, Hurkała [29], Bonneau et al. [10], and Freeman et al. [20] discussed the potentials of RBA. The latter also presented a prototype and found that an algorithm based on the user's IP address and user-agent history has a recall rate of up to 89% and a false-positive rate of 10%. Later, other features like the round-trip time of IP packets were found to be useful [47, 62].

Wiefling et al. [63] showed that verification codes sent via email are the de-facto standard for login challenges enforced by RBA. In a subsequent study, they demonstrated that providing this code in the subject can reduce the login time [64]. A study by Doerfler et al. [17] evaluated the efficacy of login challenges at preventing account takeovers. They found that up to 94% of phishing-rooted hijacking attempts and even 100% of automated hijacking attempts can be prevented. As shown by Wiefling et al. [61], RBA is perceived as more secure than passwords but also more usable than multi-factor authentication. While the latter poses an even higher security standard, increasing its adoption is a research field on its own. Rates of the Google user-base from 2018 show that less than 10% have MFA enabled [39]. In response, Google decided to auto-enable MFA for 150 million users in October 2021 [31].

Studies with System Administrators Studies with system administrators as their focus group have investigated different aspects. For example, Xu et al. [67] studied how administrators resolve common "access denied" issues and found that missing feedback can cause trial-and-error approaches. Xu and Zhou [68] surveyed characteristics of common configuration errors in an attempt to support administrators in making fewer errors. Similarly, Dietrich et al. [15], who investigated security misconfigurations, found missing documentation to be one of the root causes. Studies focusing on the update process [35, 38, 53] also find that administrators struggle to find useful information about updates although they perceive them as eminent for solving their tasks. This aligns with our findings of administrators criticizing the lack of information.

Studies analyzing tools used by administrators [33, 37, 54] highlighted the importance of usability as it can have a direct impact on security. This is especially important as administrators may have a technical background, but their mental models can be incorrect [28, 32]. Verdi et al. [59] further confirmed the importance of usability: the networking monitoring tool they analyzed received an average SUS score of 49, and the surveyed administrators complained about missing help and sometimes even failed to complete the provided task. In our study, all participants finished the task. Still, the usability of the tested RBA interface was also not assessed to be perfect. One part of this is whether administrators prefer graphical or command-line interfaces to complete their tasks. Towards this end, Voronkow et al. [60] found that 60% actually prefer a graphical interface.

Table 1: Options to configure risk-based authentication offered by cloud providers and access managers.

	Service	Automated Risk Levels	Behavior Defaults	Behavior Modifiable	Notifications Modifiable	Custom Policies
Cloud Providers	Alibaba Cloud	*only internally*	–	–	–	○
	Amazon Web Services	low, medium, high	●	●	●	○
	Google Cloud Platform	*only internally*	–	–	–	●
	IBM Cloud	low, medium, high, very high	●	●	●	●
	Microsoft Azure	no risk, low, medium, high	●	●	–	●
	Oracle Cloud	low, medium, high	–	●	●	●
	Tencent Cloud	–	–	–	–	○
Access Managers	CyberArk	non detected, low, medium, high, undetermined	–	●	●	●
	ForgeRock	–	–	–	–	●
	Ilantus	–	–	–	–	●
	Micro Focus	–	–	–	–	●
	Okta	low, medium, high	–	●	●	●
	Auth0 (Okta)	low, medium, high, neutral	–	●	●	●
	OneLogin	0–100	–	●	●	●
	Ping Identity	low, medium, high	–	●	●	●
	Thales	–	–	–	–	●

●: Offers the option, ○: Partially offers the option, –: Does not offer the option.

3 Real-World RBA Systems

In this section, we describe how the risk-based authentication systems of different real-world service providers are implemented and which configuration options they offer. For our analysis, which is summarized in Table 1, we considered five factors. First, we determined which *automated risk levels* are provided by the services, i.e., what are the potential output variables of the function that calculates a risk for a new login. *Behavior defaults* describes, if actions are suggested by the services that should be taken in response to the calculated risk levels, i.e., high risk login attempts are blocked by default. The third factor, *behavior modifiable*, describes if it is possible to modify the actions taken in response to the calculated risk levels. The fourth factor, *notifications modifiable*, considers whether the provider allows administrators to adjust how to inform the user about the actions taken in response, for example, by customizing notifications. Finally, we checked if the service providers allow for *custom policies*, which can be used to implement custom logic, e.g., block certain IP ranges, devices, or users. The results depicted in Table 1 are shown for two groups, *cloud providers* and *access managers*.

Cloud providers offer a range of services to enable customers to move IT infrastructure into their data centers and easily scale services. In contrast, access managers have an intentionally narrow focus on access-related services like identity management and MFA. As such, they close a gap by offering their service to enterprises that are already in the cloud but need features their cloud providers do not offer. To get an overview of a representative group of providers, we consulted the Gartner "Magic Quadrant for Cloud Infrastructure and Platform Services" [5] and the "Magic Quadrant for Access Management" [51].

Alibaba Cloud and Google Cloud Platform do offer RBA only internally, without an option for the customers to configure it. Microsoft Azure provides their customers with four risk levels, allows them to modify the behavior for each of them, and provides a default behavior which blocks all login attempts which are deemed as low, medium, or high risk. Notifications sent to users cannot be modified while custom policies based on various login information like the IP address, device, and the calculated risk level are supported. IBM Cloud offers all checked options, Oracle only does not provide a default behavior. Tencent is the only cloud provider supporting only custom policies based on the IP address, but no automated risk levels or any form of RBA in general. In contrast, most access managers like CyberArk, Okta, Auth0 (acquired by Okta [43]), OneLogin, and Ping Identity support RBA with all the described functionalities. Since they rely on custom configurations, none of them provides a default behavior. ForgeRock, Ilantus, Micro Focus, and Thales do not support RBA, yet.

In this study, we decided to focus on Amazon, the market leader in cloud computing according to Gartner [5] and others [11, 50]. We tested the "adaptive authentication" feature from Amazon, which is part of its paid AWS service Cognito [4]. Cognito's adaptive authentication provides three automated risk levels and a default behavior which is similar to IBM and Microsoft. It also allows to modify this behavior and the sent notifications. Custom policies are supported but only in the form of allow- and blocklists for certain IP ranges. Hence, based on the options it offers, AWS depicts an average representative in the group of cloud providers. To study Cognito's adaptive authentication interface, we built a self-hosted copy of it. In Section 4.2, we provide a detailed description of the tested interface and all of its components.

4 Method

This section describes our user study design, the tested scenarios, and the recruitment process and discusses our ethical considerations and the limitations of our findings.

4.1 Study Structure

The study was designed as an online study due to the ongoing COVID-19 pandemic and to facilitate the participation of an international audience. Prior to the main study, we conducted a pilot study with four participants to ensure that the procedure works as intended. The study, which was offered in both English and German, was split into two parts. First, we sent participants a link that led them to a website hosted on our servers where they configured a risk-based authentication system using an interface based on AWS. Afterward, they answered 26 multiple choice questions. For this first part, we observed a mean completion time of 11 minutes. In the second part, we conducted an interview, which took 33 minutes on average. *Zoom* was used throughout the study with no interaction except for a short introduction during the first part. We decided not to record the hands-on task to prevent participants from feeling monitored and avoid influencing them. Below, we outline the general structure of both parts. For a detailed description, please refer to Appendix A and B.

Part 1: Hands-on Task

1. *Agenda:* After welcoming the participants via Zoom, we briefly summarized the structure of the study and provided them with the link to the first part. We also told participants that they could seek our help at any time during the study. Still, we asked them to only do this if they do not know how to continue and that they should rather approach and solve the task like any other task they would get at work.
2. *Consent Form:* The first page on the website contained the consent form, which contained all the basic information about the study and informed participants that they could withdraw from the study at any time.
3. *Scenario:* After consenting, participants saw information about the fictitious company *MediaShop Corporation*, which they should imagine working for, and an email from their supervisor telling them about their task. This information changed depending on the scenario (see Section 4.3).
4. *Configuration:* Using a configuration interface, participants configured the risk-based authentication (depicted in Figure 2 in Appendix D). The upper settings specified a behavior for each of the three risk levels and whether or not a notification should be sent to the user. Below, the participants could adjust the wording of the notifications. We describe this interface in more detail in Section 4.2.
5. *Usability:* After the configuration, participants filled out the 10 items of the System Usability Scale (**SUS1**–**SUS10**). To ensure the quality of the data, we also included an attention check (**AC**) which all 28 participants passed.
6. *Security Knowledge:* To assess the participant's security knowledge, we asked a variant of the *Web-use Skill Measure* [25], which we expanded using common security terms from the NCSC glossary [42].
7. *Demography:* The first part concluded with the demographics (**D1**–**D6**). In addition to basic personal information, we also collected information about participant's employment, including their current job title, work experience, and the size of the company they work for.

Part 2: Interview

1. *Introduction:* We started the second part by describing the general outline of the interview. We highlighted that there are no wrong or right answers, and we are solely interested in perceptions and opinions. We also asked if we were allowed to record the interview. All participants agreed.
2. *Warm-up:* The interview started with two questions (**Q1** & **Q2**) about the participants' job to allow them to familiarize themselves with the situation. We also used these questions to double-check participants' eligibility.
3. *Risk Level Configuration:* Questions **Q3** to **Q8** covered the part of the configuration which defines the behavior for the risk levels. We asked about the reasoning for the chosen settings and if there were any difficulties. Participants who clicked on the link to the info page were asked about their reasons and whether or not the page was helpful.
4. *Notification Wording:* We now focused on the wording of the notifications. Questions **Q9** to **Q13** were similar to the previous ones and covered the reasoning, potential issues, and any consulted help.
5. *Risk-based Authentication:* After asking participants about their settings, we intended to learn about general aspects in regard to risk-based authentication. First, we asked participants how they incorporated the scenario to understand how it affected their settings (**Q14**). Afterward, **Q15** focused on prior experience with such notifications and if it may have played a role during the configuration. This question was added after the pilot study, where three of four participants mentioned this aspect without being specifically asked about it. We concluded this block with question **Q16** about any prior experiences with risk-based authentication.
6. *Improvements:* For the last set of questions (**Q17**–**Q21**), we shifted the focus back to the system participants have used to make their settings. We asked participants to assess the offered granularity of the options, potential obstacles, as well as the most positive and most negative aspects of the system. Finally, we let participants describe how the system would look like if they could change it in any way.
7. *Debriefing:* We finished the interview by answering any final questions the participants had and explained the background of the study. As part of this, we also showed participants the original system, which is part of AWS.

4.2 Configuration Interface

The central aspect of the first part of the study was the configuration of the RBA system. The user interface for this can be seen in Figure 2 in Appendix D. It consists of two components, a decision matrix defining the behavior according to the risk levels and text boxes to customize notifications. The layout of this interface is modeled after the risk-based authentication system of AWS Cognito (cf. Section 3). All aspects of the risk level and notification configuration match the Cognito interface, including texts, links, tooltips, help pages, and the overall design. We only removed the configuration of the `From` and `Reply-To` email addresses, as well as the allow- and blocklists for certain IP ranges, because we wanted to focus on adjustments which are made based on personal experience and judgement.

Risk Level Configuration. The decision matrix maps the three risk levels (*low*, *medium*, *high*) to one of four actions (*allow*, *optional MFA*, *require MFA*, *block*) and a binary decision depicting whether or not the user should be notified. If a risk is set to *allow*, any correct login the system assigns to this risk level will be granted. If set to *optional MFA*, users who have set up a second factor will be challenged to provide it. For users who have not registered a second factor, the system will continue without a challenge, i.e., the login flow is identical to *allow*. If the behavior for a risk level is set to *require MFA*, users have to provide a second factor; users who have not registered a second factor are blocked. Similarly, *block* prevents all logins. The default setting, which we adopted from AWS, allows low risk logins, whereas medium and high risk are set to *optional MFA*. Notifications are sent in all three cases. In addition to the general description of the matrix, a link to a page with further information about risk-based authentication is provided. This page is again a copy of the documentation AWS provides and contains information for each of the four behaviors and the feature that the user can be notified.

Notification Configuration. By default, AWS sends a notification email after every login attempt to the user. A login is registered after entering the correct username and password and pressing the login button, independent of the successful login and risk-level configuration.

On AWS, as well as in our user study, text boxes allow to modify the subject and the body for these notifications for each of the three risk level outcomes: (1) login is allowed, (2) MFA is required, and (3) login is blocked. Note, *optional MFA* is covered by either the notification for allowed logins or those that require MFA. For the default notifications, the email subjects for allowed and MFA logins are both set to "New login attempt" while "Block login attempt" is used for blocked logins. The body of the default notifications is shown in Listing 1 and only differs in the first sentence, which describes the risk level outcome. For example, for allowed logins the sentence is: "*We observed an unrecognized sign-in to your account with this information.*" The rest of the text includes the login time, device name, and location. The notification also instructs the user to change their password and click a link if they do not recognize the login. The email also includes another link that a user can (optionally) visit to tell the system that the login was legitimate. An administrator can add or remove template placeholders variables like `{city}` from a predefined list that can be found in the official AWS documentation [3]. To mimic this behavior, we also included a link to a self-hosted version of this message template page and observed if the participants visited it.

Listing 1: Default RBA notification message.

```
<risk level outcome>
Time: {login-time}
Device: {device-name}
Location: {city}, {country}
If this login was not by you, you should change
your password and notify us by clicking
on {one-click-link-invalid}.
If this login was by you, you can follow
{one-click-link-valid} to let us know.
```

4.3 Scenarios

We used four real-world scenarios with varying focuses to cover different circumstances system administrators may face, how they affect the configuration of the RBA, and if the tested system allows administrators to configure RBA in situations with varying requirements. Without knowing that there were four different ones, each participant randomly saw one scenario before the configuration phase. Please refer to Appendix A for the exact wording used in each scenario.

Neutral (N): In this scenario, participants were told that they are the system administrator of the *MediaShop Corporation*, where they are responsible for the online shop hosted at *dresscode.com*. An email from their supervisor Jo further informs them that it is their task to complete the configuration of the risk-based authentication.

Security (S): The background information given in this scenario is identical to the neutral scenario with one exception: the supervisor mentions a recent hack in an email that emerged from a password reuse attack. To prevent similar incidents in the future, risk-based authentication should be set up.

Usability (U): This scenario is again based on the neutral one. The only difference is given in an email where the supervisor highlights that customers should not be annoyed by the introduction of the RBA.

Neutral In-House (NI): Unlike the first three cases, participants in this scenario were not told that they administrate the online shop but "the login system 'VPN-Guard' that the employees use to work from home." Apart from this, the scenario is similar to the neutral one in that it does not introduce any focus on security or usability.

Table 2: Demographic information of participants ($n = 28$).

Age		**Gender**	
Minimum	30	Female	2
Maximum	55	Male	26
Median	40		
Degree		**Experience**	
High School	5	2–3 years	3
Training	9	4–5 years	3
Bachelor's	8	6–10 years	3
Master's	6	11–15 years	10
		>15 years	9
Residency		**Company**	
Germany	17	10–49 employees	4
USA	6	50–250 employees	5
Other	5	>250 employees	19

4.4 Recruitment and Demographics

The recruitment for our study targeted a special audience in the form of system administrators. On top of that, we conducted a qualitative study with an expected duration close to an hour which we assumed would further reduce the willingness to participate. Hence, we utilized multiple channels to get in contact with potential candidates and shared the information to the study on *LinkedIn*, the German pendant *XING*, the subreddits *r/sysadminjobs*, and *r/SampleSize*, as well as personal contacts in industry. We decided not to require prior experiences with RBA to include participants who have not worked with such a system but potentially could in the future. To also include those where sysadmin tasks only make up a certain part of their daily job, which often applies to small companies, we only required participants to work at least partially in the field of system administration. In cases where the background of the participants was not obvious to us, we asked for additional information, e.g., their *LinkedIn* profile.

We recruited a total of $n = 28$ participants for the study through the described channels. While saturation was reached after 21 participants, we decided to conduct the already scheduled seven additional interviews. The study took place in December 2021 and lasted 48 minutes on average. Each participant received a $45,- Amazon voucher as compensation. The demographics of the participants are shown in Table 2. Participants were between 30 and 55, with 40 years being the average. In terms of the gender distribution, we anticipated a shift towards male-identifying participants and tried to mitigate this by proactively contacting persons with other identities. Still, we ended up with a majority (26; 93%) who identified as male; we note this in our limitations section. Most participants resided either in Germany (17; 61%) or the United States (6; 21%). The distribution of degrees was more equal, ranging from 18% for high school to 32% for training, with the latter being the typical degree for system admins in Germany. Two-thirds of the participants (19; 68%) have worked as a system admin for at least 11 years and work in a company with more than 250 employees.

To assess the participants' security knowledge, we asked them to rate their familiarity with 9 security related items. The basis for this scale is the Web-use skill Measure [25], which we expanded with terms from the NCSC glossary [42]. The results of this assessment are shown in Table 4 in Appendix C. Overall, we observe high ratings ranging from 4.5 to 4.8; a Cronbach's α of 0.80 indicates a good level of internal consistency. The term *challenge response* is the only outlier (3.9), suggesting a slightly lower understanding of this term. Still, a composite score of 4.6 demonstrates a high familiarity with security-related terms and confirms our expectations since all participants have a strong background in IT.

4.5 Ethical Considerations

Our institution does not have an Institutional Review Board (IRB) governing this kind of study. Still, we ensured that our study would meet all requirements for such an approval, e.g., participants were told upfront about the study procedure, had to actively consent to participate, and were able to withdraw at any time. To further ensure the ethics of our research, we designed it to conform to the principles described in the Menlo Report [58] and stored all data in accordance with the General Data Protection Regulation (GDPR) [18].

4.6 Limitations

We planned our study to provide a high level of ecological validity, still, there are some limitations which we note in this section. First, our demography is shifted towards male-identifying participants despite our efforts to proactively recruit a diverse sample. Still, the distribution of system administrators is disparate in general: according to the German Federal Employment Agency only 11% of currently employed system administrators identify as female [21], the U.S. Bureau of Labor Statistics puts this proportion at 17% [57].

Secondly, participants mostly resided in Germany and the USA which can be attributed to our recruiting channels. We were not able to observe any differences in the responses across the described demographics, yet, our findings may not be representative for all system administrators.

In terms of the framing and the context of the study, we are limited by the fact that participants configured the risk-based authentication for a fictional company. Hence, participants did not have to fear any negative implications, e.g., due to potentially insecure settings and may have not taken the task as serious as if they would have configured a real-world system. Still, we believe that the insights we got are valid as they align across the group of participants. Moreover, during the interview some participants even described that they spent minutes to think about additional changes they could make on the configuration page but finally continued without any.

Finally, we studied the interface of AWS Cognito, which is only one of several available RBA systems. Thus, all findings apply primarily to AWS, and future research is needed to generally confirm them. However, as shown in Section 3, the solutions have many commonalities, so certain findings are applicable across them. For example, four services, including AWS, use the three risk levels *low*, *medium*, *high*. In response to **Q16**, seven participants also confirmed that they have worked with a similar solution before.

5 Results

We now present the results of our study, concentrating on how administrators configure risk-based authentication. Table 3 provides an overview of risk level behavior and notification configurations administrators chose in the first part of our study. We start with presenting configurations for each of the two blocks, followed by analyses of participants' reasoning behind the configuration based on the interviews during the second part of the study. Participants' responses in these interviews were separately labeled by two coders who then met to resolve differences and to create the final codebook shown in Appendix E, Table 6–9.

5.1 Risk Level Configuration

In the default configuration, low-risk logins are always allowed. For medium- and high-risk, the user is prompted to confirm the login with MFA, if it is activated for their account (*optional MFA*). By default, there is no enforcement of multi-factor authentication, nor are any login attempts blocked.

5.1.1 Configured Risk Level Behavior

Participants' risk level behavior configurations are summarized in the first block of Table 3. Overall, only one participant (N-P6) went with the defaults here. All others configured stronger measures for at least one of the three levels. Low-risk login behavior was changed by 19 participants, most of whom selected *optional MFA*; six even increased the measures to *require MFA*. For medium-risk logins, 23 overruled the default risk level behavior (*optional MFA*) and required multi-factor authentication instead. All participants who made changes chose a stronger option for high-risk logins: 17 participants required MFA for such login attempts, 10 chose to block them. In total, 11 participants selected a configuration with incrementally stronger measures on each risk level.

Figure 1 provides an overview of the risk level behavior configuration for all three risk levels separately for our four studied scenarios. Although our study was designed for an in-depth qualitative analysis, and the group sizes do not allow conclusions about (significant) differences between the groups, we can still observe a couple of interesting tendencies. For low-risk login attempts, the usability scenario is the only one in which none of the participants required multi-factor authentication. For login attempts classified as high-risk, more than half of the participants of the security scenario configured blocking, which is more than in any other scenario.

Table 3: Summary of RBA configurations. See Table 5 in Appendix C for the configuration made by each participant.

		Risk Level Behavior			
		Allow	**Optional MFA**	**Require MFA**	**Block**
Risk Level	**Low**	9*	13	6	0
	Medium	0	5*	23	0
	High	0	1*	17	10

		Notification Configuration	
		Do Not Notify	**Notify**
Risk Level	**Low**	7	21*
	Medium	2	26*
	High	1	27*

* Default

On the opposite, four participants who were all in one of the two *neutral* scenarios (see Table 5 in Appendix C) configured the same behavior for all three risk levels (*require MFA*).

5.1.2 Rationale Behind Configuration

When participants were asked to explain the rationale behind the configurations they made (**Q4**), the reasons of 14 participants revolved around multi-factor authentication and when to activate it. Six participants stated to always require MFA regardless of the risk levels. For two of them, N-P5 and NI-P5, security was a key factor for their MFA configuration. Both of them referred to the ease of use of multi-factor authentication and did not see it as a burden for their users.

> "*I chose to require MFA because from my experience, users don't find it that hard to use, and it really increases the security. So that's why I chose that for everyone, not just for low and medium risk.*" (N-P5)

Participants' personal attitudes also played a role among those requiring MFA, e.g., N-P7 expressed to be generally cautious in the light of any type of risk.

> "*As soon as it's a risk, I want to require MFA.*" (N-P7)

Two participants said they would always *offer* MFA to the users of their system (*optional MFA*) because they preferred MFA in general but refrained from requiring it due to the context being an online shop. They mainly pointed out that an online shop application was less sensitive than other systems.

> "*[...] it is dresscode.com, had it been my bank, maybe blocked would be more prudent.*" (N-P1)

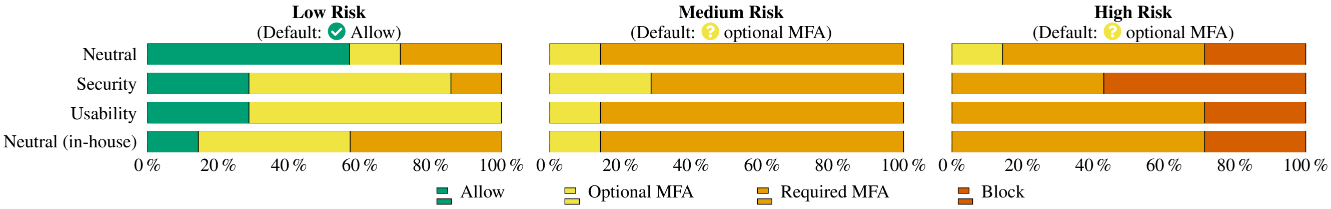


Figure 1: Overview of the risk level behavior configuration. For all risk levels, participants tend to increase the default provided by AWS. In the neutral scenario, participants chose a less strict configuration, especially in contrast to security and in-house.

The remaining participants whose justification involved MFA, basically mentioned medium or high risk to be appropriate for requiring multi-factor authentication.

In total, 11 participants have configured RBA with stronger measures for each risk level (see Section 5.1.1). For 10 of them, this incremental increase was the justification for their configuration, i.e., they wanted a stronger requirement the higher the risk was classified.

User experience when using the system was named by four participants being a reason for their configuration. This aspect is most likely connected with the considered application being an online shop, since user experience was mostly viewed in the light of customer satisfaction. That is, these participants were rather careful in bothering users with MFA or even blocking access since they feared disadvantages for their business when users preferred their less intrusive competitors.

> "*Blocking is of course extremely invasive. I mean, I would bounce our customers and we don't want that. Maybe they go to a competitor.*" (N-P3)

Six participants mentioned examples, e.g., situations which they had experienced before, that represent triggers for RBA events. These situations include login attempts from new geographical locations ($n = 5$), e.g., in the case of travel, and logins from previously unknown devices ($n = 3$). Participants used such examples to make risk level assessments for login attempts more tangible and reasoned what action they would require. Therefore, their configurations likely incorporate realistic scenarios that are relevant in the context but may also involve a risk of being too narrowed to specific anecdotes, losing sight of the broader threat landscape.

Four participants referred to having taken reactions into account they had when experiencing real-world RBA systems. While three of them mentioned their own experience from a user's perspective for services such as Netflix or PayPal, participant N-P4 stated to have followed the practice of their own company from the administrator side.

> "*When choosing the settings, I more or less followed the way we do it at ours [company]. For example, we aim to protect external access with MFA.*" (N-P4)

5.1.3 Obstacles in the Configuration

Q6 to Q8 were designed to capture obstacles participants faced during configuration and if and how they solved them. Some difficulties already became apparent when participants explained their choice in Q4. Six participants misunderstood the *optional MFA* setting when configuring the risk level behavior. For example, S-P7 interpreted optional as a decision that can be made by users in their account settings.

> "*I have interpreted this so that the user can decide whether they want to use it or not, so that they specify this somewhere in the settings beforehand, whether they want it or not. As a result, users can also control how secure they want to be.*" (S-P7)

Four participants misunderstood the concept of risk levels which became apparent when, e.g., participant U-P2 referred to different users being categorized as different risk levels. While we must keep in mind such issues when interpreting our participants' responses as a whole, we judged that none of the misconceptions qualified for invalidating entire responses.

Eight participants mentioned that being unsure about the risk level computation affected their choice (Q4). This is consistent with responses to Q6, in which the same participants named the unclear functionality of the levels a difficulty.

Further issues include missing specific descriptions of individual items ($n = 4$), and missing options for the risk level behavior configuration ($n = 3$). As an example, N-P4 asked for the ability to configure an MFA method (e.g., enforcing the use of a security key) to be required for confirming the login with MFA. S-P2 mentioned the lack of a test environment to simulate their configuration from a user's perspective.

Missing information about specific items is also reflected in the use of the provided help pages (Q7). Out of 15 participants who clicked on the help link, six participants responded they were looking for information about the risk level behavior, five participants searched for information about how the risk levels work. The remaining four participants accessed the help page out of curiosity for no specific reason.

Finally, responses to Q8 indicate that the level of information provided in our study was largely appropriate and complete. Only two participants mentioned they used external help (Google and Wikipedia) for rather small issues, and the remaining 26 participants did not use any sources of information from outside our study.

5.2 Notification Configuration

By default, AWS sends a notification after every login attempt, independent of a successful login and risk-level configuration. The second block of Table 3 provides an overview of the changes to the default notification configuration.

5.2.1 When to Notify

Overall, 20 participants have not changed the defaults suggested by AWS when to notify the user. Seven participants turned off the default notifications for low-risk sign-in attempts, two of which also turned off notifications for medium-risk attempts. Most notably, one participant U-P4 turned everything upside down and opted not to notify the user for high-risk login attempts but for the two lower risk levels.

Their preference not to annoy users in the case of a negligible risk along with the danger of notification fatigue motivated seven participants to disable the notification email for low-risk login attempts:

> "*If you get bombarded with login notifications, you get annoyed. [...] why would you look at the high risk notification unless you make it screaming? So I chose to only notify when there's a reason.*" (N-P1)

Only two of these seven participants allowed low-risk logins to proceed without MFA, while four configured optional MFA. NI-P7 even required MFA for such low-risk logins.

Out of the two participants who disabled the email even for medium-risk logins, both configured MFA to be required. The participant who turned off notifications for high-risk login attempts assumed high-risk logins to originate from "hijacked" accounts. Thus, an attacker might be able to fool the system by clicking a link in the email to report that the login was legitimate (cf. Section 5.2.2). However, they did not go into detail how such accounts could be recovered:

> "*I don't know if I'm giving away information there. If I have a hijacked account, and I send a notification, which the attacker can get and click—'Yes, it's really me.'—How it goes on then?*" (U-P4)

Interestingly, a similar scenario is mentioned by Google in a talk by Grzergor Milka [39], where immediately deleting the "*Security alert: A new login on ...* " notification, might cause an increase of the security risk score.

Across scenarios (i.e., focus on security or usability) one can observe a tendency towards sending more notifications in the security scenario, and less in the usability-focused scenario. However, due to the quantitative focus of the study, no statistical significant difference can be observed.

5.2.2 Content and Wording

The default notification text, which slightly differs by the risk level outcomes, can be found in Listing 1. All emails are also depicted in full length in Figure 2 in Appendix D.

We observed 12 participants who decided not to change the default notification or its subject. Reasons for not changing the text are either the notification being similar to those sent by popular service providers or the default is seen as sufficient in the amount of detail it contains:

> "*I found the mail to be basically fine. Of course you can still customize it individually, but in the end, the users get the information they need.*" (N-P6)

N-P4 also gave an additional justification for not touching the notification text, namely, the fear that a change will likely cause a lot of issues in future updates:

> "*I know from experience that if you put software somewhere and tinker with it, it will break by the third update at the latest. [...] Especially when working with placeholders, things go wrong so easily.*" (N-P4).

In contrast, 16 participants decided to change the text. The considerations when changing or tweaking the default template include: (1) adding details (e.g., username or IP), (2) improving the wording, (3) adding context (e.g., shop name), (4) preventing phishing, and (5) a distrust in the location.

One participant acknowledged that designing such notification requires a lot of time and effort and might also involve other departments and some testing.

> "*I'm trying to make it understandable, which can be a challenge, so in real life, I probably would have spent more time and also work with the communications people and tested it.*" (N-P1)

Add Details. Noteworthy, eight participants considered adding more details to be important. Most often, participants wanted to add the following: the username to increase trust by addressing the receiver individually, the IP address or event ID, in both cases, to enable easier debugging, and some form of contact information to support the user.

> "*It is important to have an event ID so you can assign it afterwards.*" (NI-P1)

Of course, the details participants added are influenced by the template placeholders that AWS lists in the official documentation [3]. It was accessed by eight participants of which five added details. An additional three added details but did not check the documentation and even one participant who decided against changing the notification suggested the importance of providing a lot of details.

Improve Wording. Overall, four participants noted the importance of changing the wording of the message. Here, the motivation was either to make sure the notification is understandable or to highlight certain aspects as NI-P4 describes:

> "*[...] I just made it a little more urgent, saying 'hey, you have to do something' [...]*" (NI-P4).

Add Context. In total, three participants remarked (depending on their scenario) the importance of context in the email subject and/or the body. For example, N-P1 who changed the subject to "New login attempt to dresscode.com" said:

> "*I added some context, that it was from dresscode.com in the subject, so it stands out a little bit more.*" (N-P1)

On the other hand, NI-P2 who changed the intro of the notification to "We observed an unrecognized VPN sign-in attempt" explained the motivation as follows:

> "*I can imagine the MediaShop has many different types of accounts and systems. [..] But here we're specifically talking about the VPN. So that's why I narrowed in on that.*" (NI-P2)

Prevent Phishing. Interestingly, three participants were concerned about phishing, suggesting to remove the two hyperlinks and increase trust by adding the username.

> "*Because normal phishing emails just go out without your username.*" (NI-P2).

Location Distrust. Finally, two participant wanted to add the word "Approximate" in front of the word "Location". They explained that IP-based geolocation cannot be trusted.

> "*The location is never 100% accurate. That database changes far too often, and it can be changed arbitrarily. Sometimes, when I have a new IP, it goes back to somewhere in Kansas or whatever the center point of America is. So the word 'approximate' is important.*" (S-P5).

5.3 Other Influential Factors

There are several additional factors that may have influenced participants' RBA configurations. In this context, we are particularly interested in effects of the scenario itself (**Q14**), and participants' prior experience with RBA, both from a user's (**Q15**) and administrator's perspective (**Q16**).

Incorporating the Scenario. When we asked participants whether they had incorporated the scenario, 16 participants stated they had done so, whereas 12 had not considered it. Among the participants who considered it, eight described that they had considered the context of a company with an online shop more generally. Four participants stated that they made a trade-off weighing the security of the online shop and its usability when configuring the RBA settings.

> "*When you have an online shop, you have lots of customers so it's a balance [...] you always want to have this nice and easy experience, but at the same time you want to protect the customers.*" (S-P2)

Another four participants considered the scenario when making the configurations but at the same time admitted they would have requested additional information in a real-world setting. However, S-P7 further added that even then the decision to deploy RBA would probably not have been overruled.

> "*I might have asked if it was certain that it really was a hack. But let's put it this way, if the boss says turn it on, then you turn it on.*" (S-P7)

From those participants who did not incorporate the scenario, the vast majority stated to have followed a rather general approach that was not influenced by specific properties of the described scenario ($n = 10$). Two participants explained that they used experience from their current job as a background to configure the RBA appropriately.

Previous Experience with RBA. In the pilot study, three of the four participants mentioned that they followed a login notification they received, without being specifically asked about it. Hence, we decided to ask participants if their approach was similarly influenced by such real-world notifications; 22 confirmed while 6 negated. Of the former, 16 participants described that the information in the notification text should reflect the information present in real-world notifications.

> "*I actually think that Facebook does a pretty good job of these. If I remember correctly, their emails look a lot like this and include most of these things, you know, time, device, location.*" (NI-P2)

Five participants emphasized that their configuration, i.e., the behavior in response to the risk level, was chosen such that it matches services they use.

A different aspect not directly related to the configuration, but still highlighted by five participants, is the abuse of such notifications for phishing. This risk is further enabled by the fact that even legitimate notifications, like the default text used by AWS, contain links. As we could already observe in Section 5.2.2, some participants tried to mitigate this, e.g., by removing the links. A second challenge, described by two participants, is the risk of notification fatigue caused by login notifications being sent too often.

Regarding the administrator's perspective, 16 participants did not have experience with RBA systems before. From the remaining 12 participants who already had such experiences, seven stated that the system they worked with was similar to the one used in our study. While none of them worked with AWS, we had participants who worked with Microsoft Azure that offers a similar level of detail. In contrast, five participants reported differences, most of which were subject to variations in the levels of detail, such as the way how different risk levels are presented.

5.4 Using the System

We used the System Usability Scale (SUS) to assess the usability of the RBA system in our study. The mean score across all participants was 75 ($SD = 13$), i.e., "above average" usability (>68). Still, this shows that there is room for improvement. Hence, we will now provide insights into participants' feedback on using the system and investigate which aspects are already satisfying and which can be improved.

Generally speaking, 13 participants rated the settings options as overall sufficient. While most responses to **Q17** remained rather unspecific, five participants appreciated the simplicity of the settings, and two emphasized that the configuration granularity was a good fit for the scenario showcasing a small business environment. Simplicity aspects were again referred to when we asked participants what they remembered most positively about the system (Q20). Here, simplicity was named 14 times in different flavours, often in conjunction with clarity of how settings were presented ($n = 7$). Other positive aspects concerned certain features ($n = 7$), e.g., the tooltips for optional and required MFA, and that settings can be adjusted to the context of the scenario ($n = 4$).

On the downside, 18 participants missed certain items in the settings (Q17). Note that the total number of mentions is larger than the number of participants as they could rate the options as collectively sufficient and at the same time state that they were missing something. Of the 12 participants who preferred to have more actions in response to risk levels (Q17), seven declared that this circumstance hindered them from configuring the RBA settings the way they wanted (Q18). In response to Q19, the same participants mentioned this lack as the most negative aspect of the system. Seven participants referred to missing descriptions when asked about obstacles. For four of them, this was the most negative aspect.

When we asked participants what they would change and how a perfect system would look like (Q21), 10 wished for adjustable risk levels. Five participants wanted to be able to configure multi-factor authentication in more detail. These responses are largely in line with comments to previous questions, e.g., with participants demanding the ability to further specify the MFA requirements (cf. Section 5.1.3). Some participants also asked for certain features, including a monitoring solution on the administrator side ($n = 4$) and a preview function of the final notification ($n = 3$).

6 Discussion

Overall, we identified several issues with the RBA system of AWS concerning key aspects like the meaning of risk levels and the configuration interface. Moreover, we saw a tendency to increase the defaults and observed a basic intuition for usability requirements. In the following, we like to discuss the implications of our findings in more detail and how they apply to Amazon Cognito and RBA systems in general.

6.1 Risk Levels

Most prominently, we highlight the need for a clear description of the risk levels in an RBA system and how many different levels there are. AWS's interface allows defining actions for three risk levels (*low*, *medium*, *high*). However, a fourth outcome is that the system assesses the login as "not risky at all" and does not enforce any additional security mechanisms. IBM, Microsoft, and CyberArk prevent this confusion by making this lowest risk level part of the configuration.

Second, administrators demand insights into the calculation of the risk levels, arguing that it is crucial for an informed decision. In our study, we saw participants overcoming this problem by guessing how the risk levels work, which may lead to inaccurate and potentially insecure configurations. Others argued that they must treat all levels equally if they cannot distinguish them. This may not lead to an insecure decision, yet it contradicts the initial goal of RBA in limiting security prompts for users. Others emphasized that a thorough description would be a "must-have" when deciding on a solution. Hence, service providers should also be interested in providing a complete and comprehensive documentation.

Third, we observed administrators who wanted to adjust the calculation of the risk levels and configure a more fine-granular behavior. We emphasize that fewer participants brought up this aspect, which appeared to have a more in-depth understanding of RBA. The majority was able to configure RBA according to their needs and emphasized the simplicity of the evaluated system. Hence, service providers who want to offer this feature may want to provide an additional "expert mode". This mode would allow professionals with special requirements to make more fine adjustments, while others could still use a simpler user interface.

6.2 Interface

The Amazon Cognito interface uses two terms that are crucial but, at the same time, not self-explanatory: *optional MFA* and *block*. The former defines a behavior where users who have MFA enabled are prompted, while users who have not, are still allowed to login. However, nine participants misinterpreted it such that the user is asked during the login whether or not they like to use MFA. Hence, they argued that it cannot prevent an attack because the MFA prompt can simply be skipped, and legitimate users would likely skip it for convenience reasons. We emphasize that hovering over the term "optional MFA" on the configuration interface will display a tooltip with a short explanation, just like on the original AWS implementation. Moreover, the term is also explained in more detail on the provided help page. Regarding the tooltip, none of the nine participants who misunderstood the term noticed the tooltip, as there is no visual indicator present. Seven of those nine participants noticed the information on the help page; the other two did not visit the page. To minimize the risk for misinterpretation AWS should describe the term "optional MFA" more prominently, e.g., as part of the main interface, since it is crucial for a thorough understanding of the configuration.

The term "block" also caused confusion among the administrators. In contrast to optional MFA, the general idea of denying the login was clear to all. However, details of the actual consequence were not. For example, SP-6 extensively reasoned about how long the block will last and whether it is combined with some sort of rate-limiting. The participant concluded that blocking attempts is not an option unless its consequences are fully understood, again highlighting the need for a profound documentation, similar to the risk levels. In contrast to the term "optional MFA", which is unique to AWS, blocking logins is an option all RBA services provide. Especially since it is the most invasive outcome, service providers should describe in detail of how it is implemented.

Regarding the template placeholder variables, we had participants who wanted an easier-to-use interface. While some found the approach easy and understandable, others struggled with using the variables surrounded by curly brackets and suggested preferring a drag-and-drop-based solution. Moreover, we observed that one participant misunderstood the `{one-click-link-invalid}` variable and asked why the email should contain a "non-working link." This also aligns with a statement by N-P4, who describes the granularity of the configuration interface as inconsistent: the risk level behavior is configured via radio buttons while the notification templates can be changed arbitrarily. When providing a single configuration page for both the risk levels and the notifications, as AWS does, one solution could again be an additional expert mode that would enable the use of placeholders. A second solution is to keep the configuration of the risk level and the modification on separate pages; this is what IBM, Oracle, and all access managers do.

6.3 Spicing Up Defaults

Interestingly, only one of the 28 participants did not increase the risk level behavior. It seems that the defaults AWS provides (low risk: *allow*, medium risk: *optional MFA*, high risk: *optional MFA*) are perceived as too lax. Especially 10 participants stand out who blocked access for high-risk logins which could also be caused by false positives, e.g., a login from another country during vacation. A user would have no other option than to contact the helpdesk (or order at another online shop). Moreover, it is distinct that many participants prefer to prompt the user for MFA even for low-risk logins: 19 went with either "optional MFA" or "require MFA".

Our findings highlight the need for a correctly balanced RBA configuration to be able to increase security while at the same time limiting notifications to a minimum. This is also supported by AWS's documentation, which recommends keeping "*the advanced security features in audit mode for two weeks before enabling actions*" to observe and train the login behavior before deciding on what to enforce and block [2]. In November 2021, AWS changed its defaults to "block" for all risk levels [1]. This way, enabling and using the defaults is no longer a valid option, potentially leading to more administrators who audit the logins before deciding on any actions.

6.4 Cooperation and Usability

It is pleasant to see that some administrators are aware of usability requirements, e.g., some participants took a moment to consider the impact of their work on the end-user. We noted a preference for easy-to-understand notifications, and a few participants even decided not to send notifications that could be considered unnecessary or unhelpful. While participants' primary concern was on common tasks in their responsibility like debugging (i.e., adding an event ID), we also observed an awareness to cooperate with other departments, e.g., "*the communications people*". Ultimately, this might lead to a more secure system. However, such an approach cannot be taken for granted as it is hard to follow for most smaller IT departments. For example, S-P5 summarized that it is most important to minimize the time spent with the configuration: "*you know, my time is forever compromised*." Hence, it should be the goal to reduce the workload by providing useful default notifications and guidelines.

7 Summary & Future Work

In this study, we investigate how administrators configure risk-based authentication, which issues they face, and how different requirements influence their decisions. Generally, we observed an urge of administrators to increase the default security parameters of RBA systems. We learned that some of these often unnecessary changes are owed to undefined risk levels and confusing wordings like "optional MFA." As small- to medium-sized enterprises cannot rely on trained specialists, our research reveals the need for easier-to-use configuration interfaces that support administrators in making more informed decisions, e.g., by highlighting the impact of the various configuration options. We observed that administrators are aware of potential usability issues, as some of our participants considered the impact of their work on the end-user. Still, guidance should be provided when possible.

Based on our findings, we identified multiple research directions for the design of RBA systems:

- Defaults are crucial as administrators sometimes struggle to decide which risk level behavior is reasonable and which notifications are necessary. One approach could be to have trained professionals predefine defaults based on the requirements of common scenarios, e.g., online shopping. Similarly, a guided and an expert mode could be developed to allow administrators to customize the settings according to their prior experience and knowledge.
- It needs to be investigated how terms that are open to interpretation, such as "low risk," "optional MFA," and "link-invalid" can be explained in a meaningful way.
- Administrators want to understand the implications of their configurations. It could be tested if a simulation that depicts the user's perspective provides these insights.
- Regarding the notification design, we identified a lack of consensus across participants, suggesting that future work needs to explore how to design RBA notifications, i.e., which information to include.

Acknowledgments

We thank Julian Vogt for his help with the implementation of the study website. We also thank our shepherd and the reviewers for their insightful comments and feedback. This research was supported by the research training group "Human Centered Systems Security" sponsored by the state of North Rhine-Westphalia and funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2092 CASA – 390781972.

References

[1] Amazon Web Services, Inc. Amazon Cognito: Amazon Cognito Launches New Console Experience, November 2021. https://aws.amazon.com/about-aws/whats-new/2021/11/amazon-cognito-console-user-pools/, as of June 9, 2022.

[2] Amazon Web Services, Inc. Amazon Cognito: Developer Guide – Audit Mode for Two Weeks, June 2021. https://docs.aws.amazon.com/cognito/latest/developerguide/cognito-user-pool-settings-advanced-security.html, as of June 9, 2022.

[3] Amazon Web Services, Inc. Amazon Cognito: Developer Guide – Message Templates, June 2021. https://docs.aws.amazon.com/cognito/latest/developerguide/cognito-user-pool-settings-message-templates.html, as of June 9, 2022.

[4] Amazon Web Services, Inc. Amazon Cognito: Developer Guide – Using Adaptive Authentication, June 2021. https://docs.aws.amazon.com/cognito/latest/developerguide/cognito-user-pool-settings-adaptive-authentication.html, as of June 9, 2022.

[5] Raj Bala, Bob Gill, Dennis Smith, Kevin Ji, and David Wright. Gartner Magic Quadrant for Cloud Infrastructure and Platform Services. Report G00736363, Gartner, Inc., July 2021.

[6] Robert Biddle, Sonia Chiasson, and Paul C. Van Oorschot. Graphical Passwords: Learning from the First Twelve Years. *ACM Computing Surveys*, 44(4):19:1–19:41, August 2012.

[7] Joseph R. Biden Jr. Executive Order on Improving the Nation's Cybersecurity, May 2021.

[8] Joseph Bonneau, Edward W. Felten, Prateek Mittal, and Arvind Narayanan. Privacy Concerns of Implicit Secondary Factors for Web Authentication. In *Who Are You?! Adventures in Authentication Workshop*, WAY '14, Menlo Park, California, USA, July 2014. USENIX.

[9] Joseph Bonneau, Cormac Herley, Paul C. Van Oorschot, and Frank Stajano. The Quest to Replace Passwords: A Framework for Comparative Evaluation of Web Authentication Schemes. In *IEEE Symposium on Security and Privacy*, SP '12, pages 553–567, San Jose, California, USA, May 2012. IEEE.

[10] Joseph Bonneau, Cormac Herley, Paul C. Van Oorschot, and Frank Stajano. Passwords and the Evolution of Imperfect Authentication. *Communications of the ACM*, 58(7):78–87, June 2015.

[11] Canalys. Global Cloud Services Spend Exceeds US$50 Billion in Q4 2021, February 2022. https://www.canalys.com/newsroom/global-cloud-services-q4-2021, as of June 9, 2022.

[12] Stéphane Ciolino, Simon Parkin, and Paul Dunphy. Of Two Minds about Two-Factor: Understanding Everyday FIDO U2F Usability through Device Comparison and Experience Sampling. In *Symposium on Usable Privacy and Security*, SOUPS '19, pages 339–356, Santa Clara, California, USA, August 2019. USENIX.

[13] Jessica Colnago, Summer Devlin, Maggie Oates, Chelse Swoopes, Lujo Bauer, Lorrie Faith Cranor, and Nicolas Christin. "It's Not Actually That Horrible": Exploring Adoption of Two-Factor Authentication at a University. In *ACM Conference on Human Factors in Computing Systems*, CHI '18, pages 456:1–456:11, Montreal, Quebec, Canada, April 2018. ACM.

[14] Anupam Das, Joseph Bonneau, Matthew Caesar, Nikita Borisov, and XiaoFeng Wang. The Tangled Web of Password Reuse. In *Symposium on Network and Distributed System Security*, NDSS '14, San Diego, California, USA, February 2014. ISOC.

[15] Constanze Dietrich, Katharina Krombholz, Kevin Borgolte, and Tobias Fiebig. Investigating System Operators' Perspective on Security Misconfigurations. In *ACM Conference on Computer and Communications Security*, CCS '18, pages 1272–1289, Toronto, Ontario, Canada, October 2018. ACM.

[16] Pavni Diwanji. Google: Detecting Suspicious Account Activity, March 2010. https://security.googleblog.com/2010/03/detecting-suspicious-account-activity.html, as of June 9, 2022.

[17] Periwinkle Doerfler, Kurt Thomas, Maija Marincenko, Juri Ranieri, Yu Jiang, Angelika Moscicki, and Damon McCoy. Evaluating Login Challenges as a Defense

Against Account Takeover. In *The World Wide Web Conference*, WWW '19, pages 372–382, San Francisco, California, USA, May 2019. ACM.

[18] The European Parliament and the Council of the European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). Official Journal of the European Union, L 119/1, April 2016.

[19] Florian M. Farke, Lennart Lorenz, Theodor Schnitzler, Philipp Markert, and Markus Dürmuth. "You still use the password after all" – Exploring FIDO2 Security Keys in a Small Company. In *Symposium on Usable Privacy and Security*, SOUPS '20, pages 19–35, Virtual Conference, August 2020. USENIX.

[20] David Mandell Freeman, Sakshi Jain, Markus Dürmuth, Battista Biggio, and Giorgio Giacinto. Who Are You? A Statistical Approach to Measuring User Authenticity. In *Symposium on Network and Distributed System Security*, NDSS '16, San Diego, California, USA, February 2016. ISOC.

[21] German Federal Employment Agency. Employees by Occupation (KldB 2010) – Germany (Quarterly Figures), June 2021. https://statistik.arbeitsagentur.de/DE/Navigation/Statistiken/Fachstatistiken/Beschaeftigung/Beschaeftigung-Nav.html, as of June 9, 2022.

[22] Maximilian Golla, Grant Ho, Marika Lohmus, Monica Pulluri, and Elissa M. Redmiles. Driving 2FA Adoption at Scale: Optimizing Two-Factor Authentication Notification Design Patterns. In *USENIX Security Symposium*, SSYM '21, pages 109–126, Virtual Conference, August 2021. USENIX.

[23] Maximilian Golla, Miranda Wei, Juliette Hainline, Lydia Filipe, Markus Dürmuth, Elissa Redmiles, and Blase Ur. "What was that site doing with my Facebook password?" Designing Password-Reuse Notifications. In *ACM Conference on Computer and Communications Security*, CCS '18, pages 1549–1566, Toronto, Ontario, Canada, October 2018. ACM.

[24] Paul A. Grassi, James L. Fenton, and William E. Burr. Digital Identity Guidelines – Authentication and Lifecycle Management: NIST Special Publication 800-63B, June 2017.

[25] Eszter Hargittai and Yuli Patrick Hsieh. Succinct Survey Measures of Web-Use Skills. *Social Science Computer Review*, 30(1):95–107, February 2012.

[26] Cormac Herley and Stuart Schechter. Distinguishing Attacks from Legitimate Traffic at an Authentication Server. Technical Report MSR-TR-2018-19, Microsoft, June 2018.

[27] Cormac Herley and Paul C. Van Oorschot. A Research Agenda Acknowledging the Persistence of Passwords. *IEEE Security & Privacy*, 10(1):28–36, January 2012.

[28] Dennis G. Hrebec and Michael Stiber. A Survey of System Administrator Mental Models and Situation Awareness. In *SIGCPR Conference on Computer Personnel Research*, SIGCPR '01, pages 166–172, San Diego, California, USA, April 2001. USENIX.

[29] Adam Hurkała and Jarosław Hurkała. Architecture of Context-Risk-Aware Authentication System for Web Environments. In *International Conference on Informatics Engineering and Information Science*, ICIEIS '14, pages 219–228, Lodz, Poland, September 2014. ACM.

[30] Roger Piqueras Jover. Security Analysis of SMS as a Second Factor of Authentication. *ACM Queue*, 18(4):37–60, August 2020.

[31] Guemmy Kim. Google: Making You Safer With 2SV, March 2022. https://blog.google/technology/safety-security/reducing-account-hijacking/, as of June 9, 2022.

[32] Katharina Krombholz, Karoline Busse, Katharina Pfeffer, Matthew Smith, and Emanuel von Zezschwitz. "If HTTPS Were Secure, I Wouldn't Need 2FA" – End User and Administrator Mental Models of HTTPS. In *IEEE Symposium on Security and Privacy*, SP '19, pages 246–263, San Francisco, California, USA, May 2019. IEEE.

[33] Katharina Krombholz, Wilfried Mayer, Martin Schmiedecker, and Edgar Weippl. "I Have No Idea What I'm Doing" – On the Usability of Deploying HTTPS. In *USENIX Security Symposium*, SSYM '17, pages 1339–1356, Vancouver, British Columbia, Canada, August 2017. USENIX.

[34] Leona Lassak, Annika Hildebrandt, Maximilian Golla, and Blase Ur. "It's Stored, Hopefully, on an Encrypted Server": Mitigating Users' Misconceptions About FIDO2 Biometric WebAuthn. In *USENIX Security Symposium*, SSYM '21, pages 91–108, Virtual Conference, August 2021. USENIX.

[35] Frank Li, Lisa Rogers, Arunesh Mathur, Nathan Malkin, and Marshini Chetty. Keepers of the Machines: Examining How System Administrators Manage Software Updates For Multiple Machines. In *Symposium on Usable Privacy and Security*, SOUPS '19, pages 273–288, Santa Clara, California, USA, August 2019. USENIX.

[36] Lucy Li, Bijeeta Pal, Junade Ali, Nick Sullivan, Rahul Chatterjee, and Thomas Ristenpart. Protocols for Checking Compromised Credentials. In *ACM Conference on Computer and Communications Security*, CCS '19, pages 1387–1403, London, United Kingdom, November 2019. ACM.

[37] Salvatore Manfredi, Mariano Ceccato, Giada Sciarretta, and Silvio Ranise. Do Security Reports Meet Usability? Lessons Learned from Using Actionable Mitigations for Patching TLS Misconfigurations. In *Workshop on Education, Training and Awareness in Cybersecurity*, ETACS '21, pages 1–13, Virtual Conference, August 2021. IEEE.

[38] Florin Martius and Christian Tiefenau. What Does This Update Do to My Systems? – An Analysis of The Importance of Update-Related Information to System Administrators. In *Workshop on Security Information Workers*, WSIW '20, pages 1–12, Virtual Conference, February 2020. USENIX.

[39] Grzergor Milka. Anatomy of Account Takeover. In *USENIX Enigma Conference*, Enigma '18, Santa Clara, California, USA, January 2018. USENIX.

[40] Katharine Murphy. Google Detecting 18 Million Malware and Phishing Messages per Day Related to COVID-19, July 2020. https://www.theguardian.com/australia-news/2020/jul/14/google-detecting-18m-malware-and-phishing-messages-per-day-related-to-covid-19, as of June 9, 2022.

[41] National Cyber Security Centre. Cloud Security Guidance: Identity and Authentication, November 2018. https://www.ncsc.gov.uk/collection/cloud-security/implementing-the-cloud-security-principles/identity-and-authentication, as of June 9, 2022.

[42] National Cyber Security Centre. NCSC Glossary: Definitions for Common Cyber Security Terms, December 2021. https://www.ncsc.gov.uk/information/ncsc-glossary, as of June 9, 2022.

[43] Okta, Inc. Okta Completes Acquisition of Auth0, May 2021. https://www.okta.com/press-room/press-releases/okta-completes-acquisition-of-auth0, as of June 9, 2022.

[44] Bijeeta Pal, Mazharul Islam, Marina Sanusi, Nick Sullivan, Luke Valenta, Tara Whalen, Christopher A. Wood, Thomas Ristenpart, and Rahul Chatterjee. Might I Get Pwned: A Second Generation Compromised Credential Checking Service. In *USENIX Security Symposium*, SSYM '22, Boston, Massachusetts, USA, August 2022. USENIX.

[45] Christopher Palow. After Watching This Talk, You'll Never Look at Passwords the Same Again, November 2013. http://www.meetup.com/HNLondon/events/150289672/, as of June 9, 2022.

[46] Sarah Pearman, Jeremy Thomas, Pardis Emami Naeini, Hana Habib, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, Serge Egelman, and Alain Forget. Let's Go in for a Closer Look: Observing Passwords in Their Natural Habitat. In *ACM Conference on Computer and Communications Security*, CCS '17, pages 295–310, Dallas, Texas, USA, October 2017. ACM.

[47] Esteban Rivera, Lizzy Tengana, Jesús Solano, Alejandra Castelblanco, Christian López, and Martín Ochoa. Risk-Based Authentication Based on Network Latency Profiling. In *ACM Workshop on Artifical Intelligence and Security*, AISec '20, pages 105–115, Virtual Conference, November 2020. ACM.

[48] Kevin Shalvey. A Hacker Stole More than $55 Million in Crypto after a bZx Developer Fell for a Phishing Attack, November 2021. https://www.businessinsider.com/hacker-steals-55-million-in-crypto-after-bzx-phishing-attack-2021-11, as of June 9, 2022.

[49] Alex Simons. Azure AD Identity Protection: Risk-Based Conditional Access Policies, March 2016. https://techcommunity.microsoft.com/t5/azure-active-directory-identity/azure-ad-identity-protection-is-in-public-preview-whoop-whoop/ba-p/244242, as of June 9, 2022.

[50] Synergy Research Group. As Quarterly Cloud Spending Jumps to Over $50B, Microsoft Looms Larger in Amazon's Rear Mirror, February 2022. https://www.srgresearch.com/articles/as-quarterly-cloud-spending-jumps-to-over-50b-microsoft-looms-larger-in-amazons-rear-mirror, as of June 9, 2022.

[51] Henrique Teixeira, Abhyuday Data, and Michael Kelley. Gartner Magic Quadrant for Access Management. Report G00740722, Gartner, Inc., November 2021.

[52] Kurt Thomas, Jennifer Pullman, Kevin Yeo, Ananth Raghunathan, Patrick Gage Kelley, Luca Invernizzi, Borbala Benko, Tadek Pietraszek, Sarvar Patel, Dan Boneh, and Elie Bursztein. Protecting Accounts From Credential Stuffing With Password Breach Alerting. In *USENIX Security Symposium*, SSYM '19, pages 1556–1571, Santa Clara, California, USA, August 2019. USENIX.

[53] Christian Tiefenau, Maximilian Häring, Katharina Krombholz, and Emanuel von Zezschwitz. Security, Availability, and Multiple Information Sources:

Exploring Update Behavior of System Administrators. In *Symposium on Usable Privacy and Security*, SOUPS '20, pages 239–258, Virtual Conference, August 2020. USENIX.

[54] Christian Tiefenau, Emanuel von Zezschwitz, Maximilian Häring, Katharina Krombholz, and Matthew Smith. A Usability Evaluation of Let's Encrypt and Certbot: Usable Security Done Right. In *ACM Conference on Computer and Communications Security*, CCS '19, pages 1971–1988, London, United Kingdom, November 2019. ACM.

[55] Sebastian Uellenbeck, Markus Dürmuth, Christopher Wolf, and Thorsten Holz. Quantifying the Security of Graphical Passwords: The Case of Android Unlock Patterns. In *ACM Conference on Computer and Communications Security*, CCS '13, pages 161–172, Berlin, Germany, November 2013. ACM.

[56] Blase Ur, Fumiko Noma, Jonathan Bees, Sean M. Segreti, Richard Shay, Lujo Bauer, Nicolas Christin, and Lorrie Faith Cranor. "I Added '!' at the End to Make It Secure": Observing Password Creation in the Lab. In *Symposium on Usable Privacy and Security*, SOUPS '15, pages 123–140, Ottawa, Ontario, Canada, July 2015. USENIX.

[57] U.S. Bureau of Labor Statistics. 11. Employed Persons by Detailed Occupation, Sex, Race, and Hispanic or Latino Ethnicity, January 2022. https://www.bls.gov/cps/cpsaat11.htm, as of June 9, 2022.

[58] U.S. Department of Homeland Security. The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research, August 2012. https://www.caida.org/publications/papers/2012/menlo_report_actual_formatted/, as of June 9, 2022.

[59] Fábio Luciano Verdi, Hélio Tibagí de Oliveira, Leobino N. Sampaio, and Luciana A. M. Zaina. Usability Matters: A Human-Computer Interaction Study on Network Management Tools. *Transactions on Network and Service Management*, 17(3):1865–1874, September 2020.

[60] Artem Voronkov, Leonardo A. Martucci, and Stefan Lindskog. System Administrators Prefer Command Line Interfaces, Don't They? An Exploratory Study of Firewall Interfaces. In *Symposium on Usable Privacy and Security*, SOUPS '19, pages 259–271, Santa Clara, California, USA, August 2019. USENIX.

[61] Stephan Wiefling, Markus Dürmuth, and Luigi Lo Iacono. Verify It's You: How Users Perceive Risk-based Authentication. *IEEE Security & Privacy*, 19(6):47–57, November 2021.

[62] Stephan Wiefling, Markus Dürmuth, and Luigi Lo Iacono. What's in Score for Website Users: A Data-Driven Long-Term Study on Risk-Based Authentication Characteristics. In *Financial Cryptography and Data Security*, FC '21, pages 361–381, Virtual Conference, March 2021. Springer.

[63] Stephan Wiefling, Luigi Lo Iacono, and Markus Dürmuth. Is This Really You? An Empirical Study on Risk-Based Authentication Applied in the Wild. In *International Conference on ICT Systems Security and Privacy Protection*, IFIP SEC '19, pages 134–148, Lisbon, Portugal, June 2019. IFIP.

[64] Stephan Wiefling, Tanvi Patil, Markus Dürmuth, and Luigi Lo Iacono. Evaluation of Risk-based Re-Authentication Methods. In *International Conference on ICT Systems Security and Privacy Protection*, IFIP SEC '20, pages 280–294, Virtual Conference, September 2020. IFIP.

[65] Stephan Wiefling, Jan Tolsdorf, and Luigi Lo Iacono. Privacy Considerations for Risk-Based Authentication Systems. In *International Workshop on Privacy Engineering*, IWPE '21, pages 320–327, Virtual Conference, September 2021. IEEE.

[66] Flynn Wolf, Ravi Kuber, and Adam J. Aviv. "Pretty Close to a Must-Have": Balancing Usability Desire and Security Concern in Biometric Adoption. In *ACM Conference on Human Factors in Computing Systems*, CHI '19, pages 151:1–151:12, Glasgow, Scotland, United Kingdom, April 2019. ACM.

[67] Tianyin Xu, Han Min Naing, Le Lu, and Yuanyuan Zhou. How Do System Administrators Resolve Access-Denied Issues in the Real World? In *ACM Conference on Human Factors in Computing Systems*, CHI '17, pages 348–361, Denver, Colorado, USA, May 2017. ACM.

[68] Tianyin Xu and Yuanyuan Zhou. Systems Approaches to Tackling Configuration Errors: A Survey. *ACM Computing Surveys*, 47(4):70:1–70:41, July 2015.

Appendix

A Study Part 1: Hands-on Task

Scenario

For participants in the neutral, security, and usability treatment

In this scenario, you are a system administrator of the MediaShop Corporation, a company with 300 employees. There you administrate the online shop www.dresscode.com, which sells both cheap and expensive clothing. You have just received an email from your supervisor Jo:

For participants in the neutral treatment

> Hey Alex,
>
> did you know that our login management system supports risk-based authentication? I just activated it, but not sure which settings are the best for us. Could you please complete the setup? I'm sure you will do fine.
>
> Regards,
> Jo

For participants in the security treatment

> Hey Alex,
>
> not sure if you heard it, but a hacker was able to log in to one of our customers accounts. As far as we know, the customer reused their password and the hacker got it from a hacked database. Afterwards, the hacker ordered lots of expensive jewelry using the account. My boss wants me to make sure that this should never happen again! I just activated the risk-based authentication in our login management system, could you please complete the setup for me?
>
> Regards,
> Jo

For participants in the usability treatment

> Hey Alex,
>
> did you know that our login management system supports risk-based authentication? We should give it a try. Could you please complete the setup? But make sure our customer support doesn't receive a ton of emails because of frustrated customers.
>
> Regards,
> Jo

For participants in the neutral (in-house) treatment

In this scenario, you are a system administrator of the MediaShop Corporation, a company with 300 employees. There you administrate the login system 'VPN-Guard' that the employees use to work from home. You have just received an email from your supervisor Jo:

> Hey Alex,
>
> did you know that VPN-Guard supports risk-based authentication? I just activated it, but not sure which settings are the best for us. Could you please complete the setup? I'm sure you will do fine.
>
> Regards,
> Jo

Now you open the setup...

Configuration

Page as shown in Figure 2

Usability Questionnaire

For the assessment of the configuration system you just used, please select your agreement/disagreement with the following statements.
Please select the answer choice that most closely matches how you feel about the following statements:

SUS1 I think that I would like to use this system frequently.
○ Strongly disagree ○ Disagree ○ Neither agree or disagree
○ Agree ○ Strongly agree

SUS2 I found the system unnecessarily complex.
○ Strongly disagree ○ Disagree ○ Neither agree or disagree
○ Agree ○ Strongly agree

SUS3 I thought the system was easy to use.
○ Strongly disagree ○ Disagree ○ Neither agree or disagree
○ Agree ○ Strongly agree

SUS4 I think that I would need the support of a technical person to be able to use this system.
○ Strongly disagree ○ Disagree ○ Neither agree or disagree
○ Agree ○ Strongly agree

SUS5 I found the various functions in this system were well integrated.
○ Strongly disagree ○ Disagree ○ Neither agree or disagree
○ Agree ○ Strongly agree

SUS6 I thought there was too much inconsistency in this system.
○ Strongly disagree ○ Disagree ○ Neither agree or disagree
○ Agree ○ Strongly agree

AC Please select 'Agree' as the answer to this question.
○ Strongly disagree ○ Disagree ○ Neither agree or disagree
○ Agree ○ Strongly agree

SUS7 I would imagine that most people would learn to use this system very quickly.
○ Strongly disagree ○ Disagree ○ Neither agree or disagree
○ Agree ○ Strongly agree

SUS8 I found the system very cumbersome to use.
○ Strongly disagree ○ Disagree ○ Neither agree or disagree
○ Agree ○ Strongly agree

SUS9 I felt very confident using the system.
○ Strongly disagree ○ Disagree ○ Neither agree or disagree
○ Agree ○ Strongly agree

SUS10 I needed to learn a lot of things before I could get going with this system.
○ Strongly disagree ○ Disagree ○ Neither agree or disagree
○ Agree ○ Strongly agree

How familiar are you with the following terms? Please choose a number between 1 and 5 where 1 represents "Not at all familiar" and 5 represents "Extremely familiar" with the item.

	Not at all familiar (1)	Slightly familiar (2)	Somewhat familiar (3)	Moderately familiar (4)	Extremely familiar (5)
Malware	○	○	○	○	○
Phishing	○	○	○	○	○
Two-factor authentication	○	○	○	○	○
One-time password	○	○	○	○	○
Personal identification number (PIN)	○	○	○	○	○
Auto-fill	○	○	○	○	○
Challenge-response	○	○	○	○	○
Brute-force attack	○	○	○	○	○
Security question	○	○	○	○	○

Demography

D1 What is your official job title?
Answer: ____________________

D2 For how many years have you been working as a system administrator?
○ 0–1 years ○ 2–3 years ○ 4–5 years ○ 6–10 years
○ 11–15 years ○ >15 years

D3 How large is the organization that you work for?
○ 1–9 employees ○ 10–49 employees ○ 50–250 employees
○ >250 employees

D4 How old are you?
○ Answer: ____________________ ○ Prefer not to answer

D5 Which of these best describes your current gender identity?
○ Woman ○ Men ○ Non-binary
○ Prefer to self-describe: ____________________
○ Prefer not to answer

D6 What is the highest degree or level of school you have completed?
○ No schooling completed ○ Some high school, no diploma
○ High school graduate, diploma, or equivalent
○ Trade, technical, or vocational training ○ Bachelor's degree
○ Master's degree ○ Doctoral degree ○ Prefer not to answer

B Study Part 2: Interview

Introduction

- Thanks again for taking part in this study.
- The interview will take about 30 minutes.
- Are you OK with me recording our interview?
- <*Start recording.*>
- There are obviously no right or wrong answers here, we are just interested in your personal perceptions and your honest opinions.
- Are there any questions from your side before we start?

Warm-up Questions

Q1 What do you like about your job as an administrator?

Q2 What are the main tasks in your job?

Behavior for the Risk Levels

We're now interested in the settings for the risk-based authentication. If we use the term "settings" in the following, we refer to the table on top of the page.

Q3 How did you go about choosing the settings?

If not already covered by Q3

Q4 Explain the reasons for the chosen settings.

If not already covered by Q3

Q5 Explain the reasons for the chosen settings for notifying the users.

Q6 Which difficulties or problems did you have when configuring the settings?

If log showed that info page was visited

Q7 You have used the Wiki which contained more information about the settings: why did you click on the link? Was the information helpful?

Q8 Have you used any other help, e.g., Google? If yes, why?

Wording of the Notifications

We're now interested in the notifications and their settings, i.e., the text fields on the bottom of the page.

Q9 How did you go about when choosing the wording of the notifications?

If not already covered by Q9

Q10 Explain the reasons for the way you worded the notifications.

Q11 Which difficulties or problems did you have when choosing the wording of the notifications?

If log showed that info page was visited

Q12 You have used the info page which contained more information about the configuration of notifications: why did you click on the link? Was the information helpful?

Q13 Have you used any other help, e.g., Google? If yes, why?

Risk-based Authentication

Q14 How did you incorporate the scenario when making the configurations?

Q15 Have you ever received such a notification? If yes, have you thought about this experience when making the configurations?

Q16 Have you ever worked with risk-based authentication before? If yes, how did you experience the system you used compared to this one.

Potential Improvements

We're now interested in the system as a whole, i.e., both the table on the top and the text fields on the bottom of the page.

Q17 How do you rate the current level of detail in the settings options?

Q18 Please explain anything that hindered you from the risk-based authentication in the way you wanted.

Q19 What did you notice or remember most negatively about the system?

Q20 What did you notice or remember most positively about the system?

Q21 If you could change the system in any way you want: how would the perfect system look like?

Debriefing

- Research goal: Analyze the usability of an exemplary systems for the configuration of risk-based authentication, identify good and bad aspects to be able to make recommendations on how to improve such a system.
- Do you have any questions about the interview or the study?
- <*Stop recording.*>

C Additional Tables

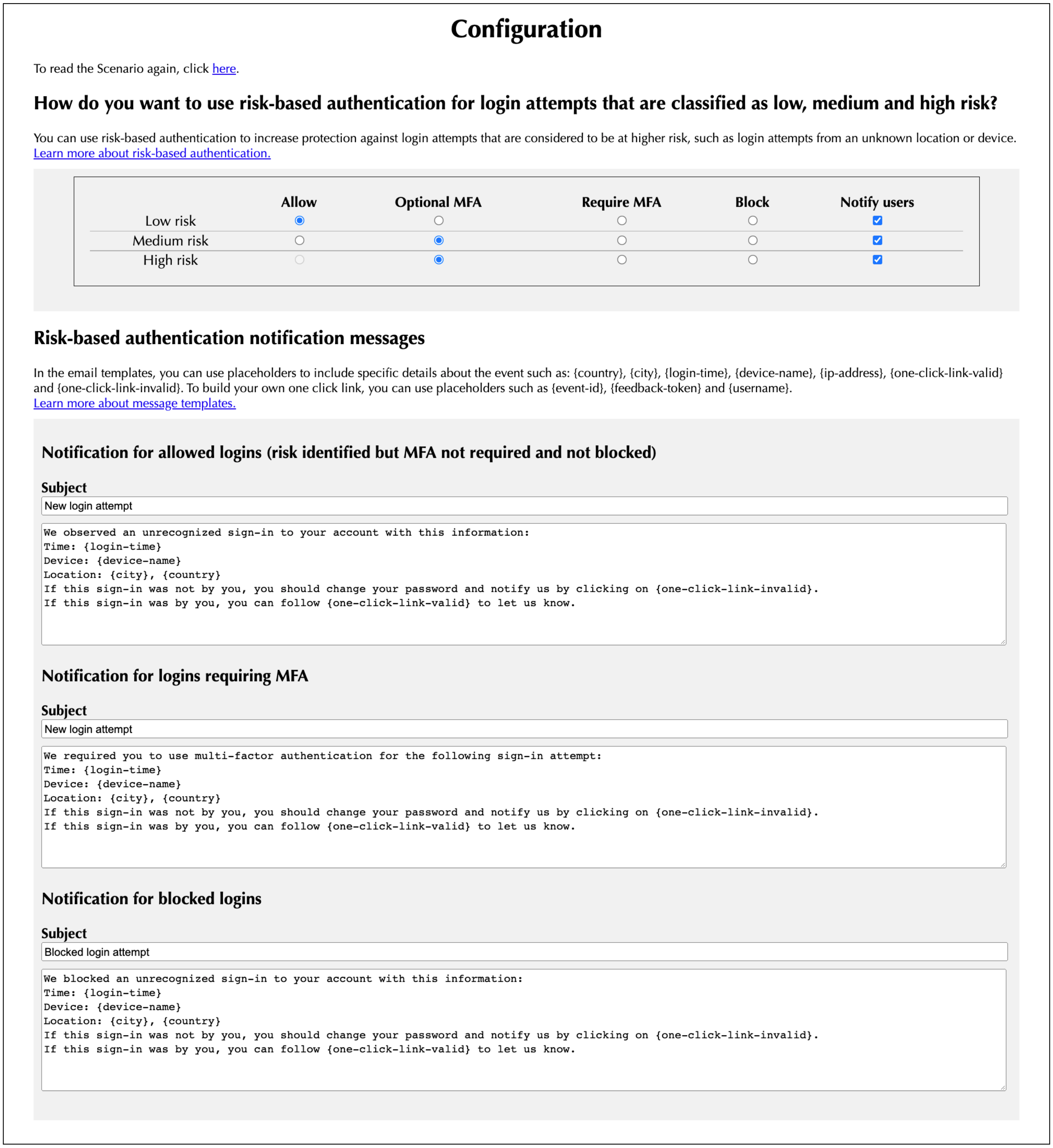
Table 4: General security knowledge of the participants determined by rating the familiarity with 9 security-related items. The items are in the order of appearance in the questionnaire.

Item	Mean	SD
Malware	4.6	0.6
Phishing	4.8	0.5
Multi-Factor Authentication	4.8	0.4
One-Time Password	4.7	0.5
Personal Identification Number (PIN)	4.8	0.4
Auto-Fill	4.5	0.6
Challenge-Response	3.9	1.1
Brute-Force Attack	4.5	0.9
Security Question	4.6	0.6
Composite score	4.6	0.7
Cronbach's α	0.80	

Table 5: Configuration for the behavior of the risk levels (✓: allow, ?: optional MFA, !: require MFA, ⊖: block), notifying users (🔔: notify, 🔕: do not notify), and changes to the notification (✎: changed, –: unchanged).

		Risk Level Behavior			Notify Users			Changed
	Participant	Low	Medium	High	Low	Medium	High	Notification
	Default	✓	?	?	🔔	🔔	🔔	–
Neutral	N-P1	✓	!	!	🔕	🔕	🔔	✎
	N-P2	✓	!	⊖	🔔	🔔	🔔	✎
	N-P3	✓	!	⊖	🔔	🔔	🔔	–
	N-P4	?	!	!	🔔	🔔	🔔	–
	N-P5	!	!	!	🔔	🔔	🔔	✎
	N-P6	✓	?	?	🔔	🔔	🔔	–
	N-P7	!	!	!	🔔	🔔	🔔	–
Security	S-P1	✓	?	!	🔔	🔔	🔔	–
	S-P2	✓	?	!	🔔	🔔	🔔	✎
	S-P3	?	!	⊖	🔔	🔔	🔔	✎
	S-P4	!	!	⊖	🔔	🔔	🔔	✎
	S-P5	?	!	⊖	🔔	🔔	🔔	✎
	S-P6	?	!	!	🔔	🔔	🔔	–
	S-P7	?	!	⊖	🔕	🔔	🔔	✎
Usability	U-P1	✓	!	⊖	🔔	🔔	🔔	–
	U-P2	?	!	!	🔔	🔔	🔔	–
	U-P3	?	!	!	🔕	🔕	🔔	–
	U-P4	?	!	⊖	🔔	🔔	🔕	✎
	U-P5	?	!	!	🔔	🔔	🔔	✎
	U-P6	✓	?	!	🔔	🔔	🔔	✎
	U-P7	?	!	!	🔕	🔔	🔔	✎
Neutral (in-house)	NI-P1	?	!	!	🔔	🔔	🔔	✎
	NI-P2	✓	?	!	🔕	🔔	🔔	✎
	NI-P3	?	!	⊖	🔕	🔔	🔔	✎
	NI-P4	!	!	!	🔔	🔔	🔔	✎
	NI-P5	!	!	⊖	🔔	🔔	🔔	–
	NI-P6	?	!	!	🔔	🔔	🔔	–
	NI-P7	!	!	!	🔕	🔔	🔔	–

D RBA Configuration Interface



Configuration

To read the Scenario again, click here.

How do you want to use risk-based authentication for login attempts that are classified as low, medium and high risk?

You can use risk-based authentication to increase protection against login attempts that are considered to be at higher risk, such as login attempts from an unknown location or device. Learn more about risk-based authentication.

	Allow	Optional MFA	Require MFA	Block	Notify users
Low risk	●	○	○	○	☑
Medium risk	○	●	○	○	☑
High risk	○	●	○	○	☑

Risk-based authentication notification messages

In the email templates, you can use placeholders to include specific details about the event such as: {country}, {city}, {login-time}, {device-name}, {ip-address}, {one-click-link-valid} and {one-click-link-invalid}. To build your own one click link, you can use placeholders such as {event-id}, {feedback-token} and {username}.
Learn more about message templates.

Notification for allowed logins (risk identified but MFA not required and not blocked)

Subject

New login attempt

```
We observed an unrecognized sign-in to your account with this information:
Time: {login-time}
Device: {device-name}
Location: {city}, {country}
If this sign-in was not by you, you should change your password and notify us by clicking on {one-click-link-invalid}.
If this sign-in was by you, you can follow {one-click-link-valid} to let us know.
```

Notification for logins requiring MFA

Subject

New login attempt

```
We required you to use multi-factor authentication for the following sign-in attempt:
Time: {login-time}
Device: {device-name}
Location: {city}, {country}
If this sign-in was not by you, you should change your password and notify us by clicking on {one-click-link-invalid}.
If this sign-in was by you, you can follow {one-click-link-valid} to let us know.
```

Notification for blocked logins

Subject

Blocked login attempt

```
We blocked an unrecognized sign-in to your account with this information:
Time: {login-time}
Device: {device-name}
Location: {city}, {country}
If this sign-in was not by you, you should change your password and notify us by clicking on {one-click-link-invalid}.
If this sign-in was by you, you can follow {one-click-link-valid} to let us know.
```

Figure 2: The interface of the central page in our study where participants configured the risk-based authentication. The layout of this interface is modeled after the risk-based authentication system of AWS Cognito (see Section 3). All aspects of the risk level and notification configuration match the Cognito interface, including texts, links, tooltips, help pages, and the overall design.

E Codebook

Table 6: Codebook for Q3–Q8 used in Section 5.1 Risk Level Configuration.

Code	Freq.	Description	Example
Q3: How did you go about choosing the settings? **/ Q4**: Explain the reasons for the chosen settings.			
MFA	14	MFA was mentioned as a central aspect.	*“As soon as it’s a risk, I want to require MFA.”* (N-P7)
Increase each level	10	Configuration was chosen such that the action increases with each level.	*“Accordingly, I then reinforced the whole thing for each risk level.”* (S-P3)
Misunderstanding	9	Answer reveals a misunderstanding of one or multiple features of the configuration.	*“A low risk for example is that I use MFA [...] a high risk would be that I just do username/password and no notification.”* (U-P7)
Example	8	An example was mentioned that triggers RBA.	*“If I log in from 400 kilometers away, because I’m on vacation, [...] I would expect that this is classified as medium and that I would have to provide a second factor.”* (N-P4)
Risk levels unclear	8	Configuration was affected by being unsure how the risk levels work.	*“To really make a liquidated decision, I would just have to understand: what does this mean?”* (U-P3)
Real world	4	A real-world RBA system was used as an orientation.	*“Microsoft, Amazon, Google, they all do it exactly the same way [...] I know it from there and have tried to set that as a goal worth striving for.”* (N-P4)
User	4	User should not be turned down by the RBA.	*“Blocking is of course extremely invasive. I mean, I would bounce our customers and we don’t want that. Maybe they go to a competitor.”* (N-P3)
Q6: Which difficulties or problems did you have when configuring the settings?			
None	13	None occurred	*“No. Pretty straightforward.”* (N-P5)
Risk levels unclear	8	Functionality of the risk levels was unclear.	*“I missed details saying how the login attempts are checked and which parameters classify low, medium, and high risk. That would definitely be a criteria for me.”* (S-P3)
Missing description	4	A description was missing.	*“I wasn’t dead sure what you meant by ’Block’, you know, was that like a block after a certain amount of time?”* (S-P6)
Missing option	3	An option was missing.	*“The system I’m used to spits out in a nice little JSON structure so you can actually have more granularity in exactly what’s going on.”* (NI-P2)
Q7: Where you looking for a specific information in the Wiki? If yes, which and did you find it?			
Info actions	6	Participant was looking for information about the actions to the risk levels.	*“I was trying to get the specifics of the allow, optional, and require MFA just to make sure it was doing what I think it was.”* (N-P5)
Info risk levels	5	Participant was looking for information about how the risk levels work.	*“I was just trying to look and see what the risk levels mean, if they were defined”* (NI-P4)
Curiosity	4	Participant was just curious and not looking for any specific information.	*“You put the link there. I was going to click on it. I’m curious.”* (U-P5)
Q8: Have you used any other help, e.g., Google? If yes, why?			
No	26	Participant did not use any other help .	*“Yeah, no. This is pretty common stuff right now.”* (N-P6)
MFA	1	Participant was looking for information about MFA.	*“I googled MFA just before this.”* (NI-P6)
RBA	1	Participant was looking for information about RBA.	*“I don’t know anything about this risk-based authentication, so I googled it last night, read the Wikipedia article and thought to myself: ’yes, of course, you’ve heard of it and it’s used regularly’.”* (N-P4)

Table 7: Codebook for Q9–Q15 used in Section 5.2 Notification Configuration.

Code	Freq.	Description	Example
Q5 Explain the reasons for the chosen settings for notifying the users.			
Inform user	13	User should be informed.	*"Keeping the user informed at every step along the way does introduce trust, and not notifying is the easiest way to lose that trust, even if you are doing everything else correctly."* (N-P6)
Low risk negligible	5	No notification for logins with a low risk because a low risk is negligible.	*"For logging attempts that are low risk, this is just normal run-of-the-mill everyday activity; I don't want to notify users about it."* (NI-P2)
Personal preference	5	Personal preference of the participant for a certain setting.	*"I know it can seem a little bit tedious, but I'd much rather know then not know that something's been signed in on. I find that important."* (S-P5)
Fatigue	2	No notification in certain cases to avoid fatigue.	*"If you get bombarded with sign-in notifications you get annoyed. [...] why would you look at the high risk notification unless you make it screaming? So I chose to only notify when there's a reason."* (N-P1)
Real world	2	Experience with RBA in the real world.	*"A good example that exists right now is Disney Plus. Disney Plus does not notify you when you've been signed onto a new device. Well, my account for Disney Plus was compromised, and as a result of that, somebody was watching all kinds of stuff, but I had no idea."* (S-P5)
Q9: How did you go about when choosing the wording of the notifications? / Q10: Explain the reasons for the way you worded the notifications.			
Add details	8	Details were added.	*"I added some more information that I feel would be nice to see. Just so I can verify that it is my IP that I'm using on either my phone or on my computer."* (S-P2)
Default OK	6	Notifications were not changed because the defaults meet the expectations.	*"I found the mail to be basically fine. Of course you can still customize it individually, but in the end, the users get the information they need."* (N-P6)
Similar	6	Notifications are similar to those used by real world services.	*"I mean most of the message that I receive are similar to this one."* (N-P5)
Wording	4	Wording was changed.	*"For the middle and bottom one, I just made it a little more urgent, saying 'hey, you have to do something', or there is an attempt we blocked that needs attention right away."* (NI-P4)
Add context	3	Context was added.	*"I added some context, that it was from dresscode.com in the subject, so it stands out a little bit more."* (N-P1)
Prevent phishing	3	Notification was changed such that phishing is prevented.	*"I would remove the link and just say: if this was not you, then you should change your password and notify us, period."* (S-P2)
Location distrust	2	Notification was changed to control for the distrust in the location parameter.	*"The location is never 100% accurate. That database changes far too often, and it can be changed arbitrarily. [...] Sometimes when I have a new IP, it goes back to somewhere in Kansas or whatever the center point of America is. So the word 'approximate' is important."* (S-P5)
Q11: Which difficulties or problems did you have when choosing the wording of the notifications?			
None	22	None occurred.	*"No, that was very straightforward."* (U-P2)
Missing description	3	A description was missing.	*"So these placeholders, there may be more of them, but I wouldn't have known which keyword to search for."* (S-P7)
Missing option	2	An option was missing.	*"A couple of the tools that we use will actually give you the view that the user will see whether they're using a PC or a mobile device."* (N-P6)
Repetitive	2	The repetitiveness of three similar notification fields.	*"It was a bit confusing, because the options were always the same, only the text was different."* (N-P3)
Q12: Where you looking for a specific information in the Wiki? If yes, which and did you find it?			
Placeholders	5	Information about the placeholders that can be used in the notifications.	*"I was checking the security template placeholders."* (N-P5)
Curiosity	4	Out of curiosity not looking for specific information.	*"I just wanted to see what information was offered."* (U-P3)
Q13: Have you used any other help, e.g., Google? If yes, why?			
No	28	Participant did not use any other help .	*"No, I'm familiar with email templates."* (S-P3)

Table 8: Codebook for Q14–Q16 used in Section 5.3 Other Influential Factors.

Code	Freq.	Description	Example
Q14: How did you incorporate the scenario when making the configurations?			
No: general approach	10	General approach was taken which is independent of the scenario.	*“Regardless of the scenario, I believe that requiring MFA for everyone is better for every organization.’* (N-P5)
Yes: generally	8	Context of the company was considered in general.	*“I considered it a little in the sense that they have 250 employees. So this is like a small business [...] They probably have never used a system like this before, and that’s why I wanted to be cautious with it when first rolling it out. If this was a large enterprise with twenty thousand users, then they’re probably used to this already.’* (NI-P2)
Yes: info missing	4	Scenario was considered but participant would ask for additional information if it was an actual task.	*“I might have asked if it was certain that it really was a hack. But let’s put it this way, if the boss says turn it on, then you turn it on.”* (S-P7)
Yes: tradeoff	4	Tradeoff between the security of the online shop / VPN and its usability.	*“When you have a web shop, you have lots of customers so it’s a balance [...] you always want to have this nice and easy experience, but at the same time you want to protect the customer.”* (S-P2)
No: current job	2	Settings are based on the background of the current job.	*“So honestly, I did it based on my job right now.”* (U-P2)
Q15: Have you ever received such a notification? If yes, have you thought about this experience when making the configurations?			
Contained info	16	The contained information should match the one present in real world notifications.	*“I actually think that Facebook does a pretty good job of these. If I remember correctly, their emails look a lot like this and include most of these things, you know, time, device, location.”* (NI-P2)
Not considered	6	Experience with notifications was not considered.	*“No. I can tell you, I personally get such notifications very rarely.”* (S-P1)
Configuration	5	The configuration should match the one used by real world services.	*“I’ve been using 2FA for my Google [account] for a few years now. Ever since, I’ve been getting notifications regularly, e.g., if a new device is used. And that’s essentially what I expect from such a system.”* (N-P4)
Phishing	5	Experience with phishing attempts based on RBA notifications.	*“The classic example is Amazon or Paypal [...]: ‘please verify your account’ or ‘click on the link’ and if you take a look at the link, then it leads to I don’t know where.”* (U-P4)
Fatigue	2	Notification fatigue should be avoided with the chosen configuration.	*“People are getting a lot of new device notification, even though they have been using their device for two years. [...] I think people are just deleting it.”* (N-P1)
Q16: Have you ever worked with risk-based authentication before? If yes, how did you experience the system you used compared to this one.			
No	16	Participant has not worked with RBA before.	*“I haven’t configured anything similar myself yet.”* (S-P3)
Yes: similar	7	Participant has worked with a similar RBA system before.	*“They all offer basically the same. They have similar commonalities and they’re all here.”* (N-P6)
Yes: different	5	Participant has worked with a different RBA system before.	*“The one I’m using is a, that’s a much more sophisticated system in a couple of ways.”* (NI-P2)

Table 9: Codebook for **Q17**–**Q21** used in Section 5.4 Using the System.

Code	Freq.	Description	Example
Q17: How do you rate the current level of detail in the settings options?			
Missing: actions	12	Actions in response to the risk levels need to be more fine-grained.	*“I would like to fine-tune that a bit, the notifications and the classification. When do we block? And I would like more options, e.g., exclude certain regions completely.”* (U-P3)
Sufficient: misc	7	Different reasons for the granularity being sufficient.	*“For me, it would have been sufficient, at least for the start, i.e., to start with this configuration, set that up, and test it.”* (S-P4)
Missing: description	6	Description need to be more detailed.	*“If I don’t know what low, medium, or high risk exactly means, there is no reason for me to distinguish between them.”* (NI-P7)
Sufficient: simple	5	Settings are simple but sufficient.	*“I would say this is a as simple as you could make it. So, I wouldn’t want there to be any less options than this.”* (S-P5)
Missing: risk levels	3	Risk levels need to be more fine-grained.	*“There should be four steps if we want to be able to choose each and every one of these options that you have. So you should have like ‘super high,’ for example, and then you use block.”* (S-P2)
Sufficient: small businesses	2	Granularity is sufficient for small business.	*“A system being used for small businesses. So the owner could decide to set up some security, but not that granular.”* (N-P5)
Missing: options	1	Options need to be more fine-grained.	*“You could go into more detail with the notify options and say: how should the user be notified? Email? SMS?”* (N-P3)
Q18: Please explain anything that hindered you from the risk-based authentication in the way you wanted			
Nothing	15	Nothing hindered the participant.	*“Based on the ask, everything was right there.”* (N-P6)
More actions	7	More actions would have been necessary.	*“I would like to have the option to select the second factor, maybe also depending on the risk level [...], e.g., for high risk, it must be a hardware token.”* (N-P4)
Missing description	7	Description which was missing or not detailed enough.	*“The documentation of the risk levels, [...], it was difficult to understand what is meant which is why I had to guess.”* (S-P7)
Q19: What did you notice or remember most negatively about the system?			
Nothing	9	No negative aspect was mentioned.	*“Well. I’m not sure if I have one.”* (U-P2)
Missing actions	7	Action which was missing.	*“Where would I define what low, medium, and high risk and stuff like that is? What does that involve? That is not clear.”* (S-P2)
Missing description	4	Description which was missing or not detailed enough.	*“What bothered me the most, there is no info about the risk calculation. I mean, that’s the core task of it.”* (U-P3)
Notification settings	4	Something about the settings for the notifications.	*“The modernization of a preview to the users for said notification [...] What’s the user going to see? Like, I literally want to see what it’s going to look like on iOS, Android, Windows?”* (N-P6)
UI issues	2	Issues with user interface.	*“It’s all very text-heavy.”* (NI-P3)
Q20: What did you notice or remember most positively about the system?			
Simplicity	14	The simplicity with which the settings can be adjusted.	*“I think the simplicity of it [...] It does what I need it to do and, you know, my time is forever compromised, right?”* (S-P5)
Feature	7	A certain feature.	*“The fact that you had an optional MFA in there rather than just allow and require.”* (S-P6)
Clarity	7	The way in which the different settings are presented.	*“I definitely like the option matrix at the top. It clearly explains: if this is a low, medium, high risk, these are my options.”* (N-P6)
Adjustability	4	The ability to adjust the settings.	*“It’s not binary. It can adjust to context.”* (N-P1)
Q21: If you could change the system in any way you want: how would the perfect system look like?			
Risk level adjustment	10	Make the risk levels adjustable.	*“I would want to differentiate the set of rules very precisely, like down to the smallest possible detail.”* (N-P7)
MFA	5	More configurability of the MFA options.	*“I would want a selection dialog for each 2FA method, to say for which risk level it’s allowed.”* (S-P7)
UI change	5	Change the presentation of the settings.	*“I would make it so that you can expand and collapse the messages below, just to make it clearer”* (U-P1)
Description	4	Change or add descriptions.	*“What does low risk mean? What does medium risk mean? Maybe you could add an explain button or something.”* (N-P3)
No changes	4	No changes necessary.	*“For me, it’s perfect. Can’t add anything to that”* (U-P4)
Reporting	4	Add reporting for the logins.	*“Reports are important, with all the details in there, not only login time, device, but everything about this case.”* (NI-P5)
Notifications	3	More configurability of the notifications and the notification channel.	*“If there was the option to make that a bit more detailed, like how the user should be notified, that would also be good.”* (N-P3)
Preview	3	Preview of the notifications.	*“I would add this kind of ‘show me a preview’ when I click like generate this”* (S-P2)