How Enforcement Shapes 2FA Usability and Perceived Workload: A Comparative Study Across Populations

Zhiyu He[1][0009−0004−8352−3452], Francisco Vega[2][0009−0008−2441−4458], and Yuan Cheng[1][0000−0001−7176−3951]

[1] University of Nottingham Ningbo China, Ningbo, China
{scxzh2,yuan.cheng}@nottingham.edu.cn
[2] River Islands High School, Lathrop CA 95330, USA
fvega@riacademies.net

Abstract. Two-factor authentication (2FA) is widely used to enhance online security, yet its impact on users' perceived workload in everyday use is not well understood. We compare 2FA experiences among college students and Amazon Mechanical Turk workers based on 212 survey responses, using NASA-TLX workload ratings and open-ended questions. The results show that users who have previously experienced 2FA-related access loss report significantly higher workload across all dimensions of NASA-TLX, suggesting that poor user experience is a major factor contributing to the perceived burden of using 2FA. Additionally, users who are required to use 2FA report higher levels of time pressure, effort, and frustration than those who adopt 2FA voluntarily or through incentives. In contrast, the frequency with which users pay attention to security cues, such as IP addresses, devices, and locations, shows little effect on their subjective workload, indicating that security prompts themselves are not a major source of burden. Open-ended responses also highlight issues such as time pressure from CAPTCHA countdowns, delays caused by multi-step authentication processes, missed events due to login friction, and anxiety during account recovery. Overall, this study suggests that improving 2FA should focus on enhancing user experience, particularly through clearer guidance and more reliable account recovery processes, rather than reducing security cues or relying solely on stricter verification.

Keywords: Two-factor authentication · Usability · User workload.

1 Introduction

With the rising number of password breaches and sophisticated phishing attacks, protecting user account security has become a major challenge for cybersecurity. Over the years, organizations have tried to combat account compromises by enforcing various password policies, such as requiring complex password requirements and mandatory periodic changes. However, despite these efforts, such

policies have not been fully effective in preventing account breaches [25]. Moreover, users tend to reuse the same or similar passwords across multiple services, which exacerbates the consequences of password breaches by enabling credential stuffing and other advanced password guessing attacks [8, 15, 19].

To overcome the limitations of password-based authentication, many organizations now use two-factor authentication (2FA) as an additional security measure. 2FA requires users to provide another form of authentication, such as a one-time verification code, a biometric, or a hardware security key, in addition to a password. This extra step can make it considerably harder for attackers to gain unauthorized access. However, each method of 2FA comes with its own usability and security challenges. Inconsistent implementation across organizations further complicates the user experience, often leading to confusion and frustration. As a result, despite its clear security benefits, many users remain hesitant or skeptical about adopting 2FA. An early study reported that only 6.4% of sampled Google users used 2FA [21]. While adoption has grown substantially since then, rising to roughly 89% of consumers who report having at least one account that requires 2FA as of 2023 [7], there is still a lack of research focusing on the factors that influence users' decisions to adopt or reject 2FA technology. Understanding these factors can help promote 2FA adoption from the users' perspective.

Universities provide a useful setting for studying these questions because many institutions have recently deployed 2FA, often with mandatory enrollment. Prior work has leveraged campus rollouts to examine user opinions about 2FA use and adoption [1, 6, 23]. For example, Abbott et al. conducted a series of surveys at a university during its phased 2FA rollout, comparing perceptions between mandatory users and voluntary users [1]. Their findings revealed that mandatory policies led to decreased user satisfaction. Prior work has primarily focused on acceptance and satisfaction during campus rollouts (e.g., [1, 6]). Less is known about how enforcement shapes multidimensional workload, whether these patterns generalize beyond a single institutional context, and whether specific user experiences, such as 2FA-related account access loss and attention to contextual security cues, explain perceived burden.

Building on this line of work, we conduct a comparative usability analysis of 2FA across adoption models and user populations. We address the following research questions:

- **RQ1:** How does the adoption context (mandatory, voluntary, or incentivized) affect users' perceived workload and attitudes toward the usability of 2FA?
- **RQ2:** In what ways do university users differ from general Internet users in their perceptions of workload and usability when controlling for adoption mode and other factors?
- **RQ3:** To what extent do users report noticing or consulting contextual security information (e.g., IP address, device, location, and time) during authentication?

We collected data through two online surveys administered to two distinct cohorts: Amazon Mechanical Turk (MTurk) workers and students recruited from

a Computer Science major at a U.S. public university. The university survey was conducted approximately one year after 2FA became mandatory for campus accounts. The MTurk survey included a screening question to exclude participants without any 2FA-enabled accounts. After removing invalid or low-quality responses, the final analytic sample comprised 212 participants. Using responses from both cohorts and analyzing NASA-TLX workload measures, we report the following main findings:

- **Adoption Mode (RQ1):** Users subject to mandatory 2FA report significantly higher levels of time pressure, effort, and frustration than those who adopted 2FA voluntarily. Users who adopted 2FA through incentives fell between these two groups.
- **User Group (RQ2):** After accounting for age, adoption mode, and prior 2FA-related account access loss, we find only marginal differences in overall workload between university users and general Internet users.
- **Security Attention (RQ3):** Users who pay closer attention to contextual security information (e.g., device, location, login alerts) do not report increased workload, indicating that security awareness does not inherently add cognitive burden.

These findings indicate that 2FA burden mainly comes from mandatory enforcement and negative user experiences (particularly account access loss), rather than from security cues themselves. Accordingly, we recommend that developers and policy makers prioritize usability improvements that reduce 2FA-related account access loss risk, make recovery processes more convenient, and provide clearer interface guidance, rather than focusing solely on stricter verification.

The rest of this paper is organized as follows. Section 2 reviews related work on the usability and adoption of 2FA. Section 3 describes the survey design, measures, participant recruitment, and analysis procedures. Section 4 reports the quantitative results (NASA-TLX workload comparisons and regression analyses) and summarizes themes from open-ended responses. Section 5 discusses implications for 2FA design and deployment, as well as limitations and directions for future work. Section 6 concludes the paper.

2 Related Work

2FA Adoption and User Attitudes

Previous research has examined how users adopt and respond to 2FA, with many studies focusing on academic settings where institutions have increasingly enforced its use.

Colnago et al. [6] conducted a longitudinal study during the rollout of Duo 2FA at Carnegie Mellon University, where 2FA was mandatory for employees and voluntary for students. They found that although users initially expressed concerns about burden, their actual experiences were more positive than expected. Over time, the perceived hassle diminished, and learning costs were relatively

low. Similarly, Abbott and Patil [1] studied the phased 2FA deployment at Indiana University using both user surveys and system logs. They observed that while users mandated to adopt 2FA were slightly less satisfied than voluntary adopters, the difference was smaller than anticipated. Overall, the transition to mandatory 2FA did not significantly reduce user acceptance, and users generally found Duo to be annoying but manageable and secure. These findings align with insights from Golla et al. [14] and the Cybercrime Support Network(CSN) [7], which concluded that user motivation and perceived time pressure strongly influence 2FA acceptance.

Attitudes toward 2FA varied among different roles within the university context. At Brigham Young University, Dutson et al. [11] found that students and faculty generally report more negative attitudes toward Duo compared to staff, with nearly half of respondents reporting account access loss due to 2FA. Many users indicated that they would opt out of 2FA if given the choice. However, other studies consistently show that while mandatory policies may spur some user resistance, they do not necessarily worsen users' actual usability ratings [1, 6].

User education and familiarity play crucial roles in mitigating perceived burdens associated with 2FA. Effective training can reduce user resistance, and as users become accustomed to the process, their subjective burden diminishes, allowing them to gradually adopt 2FA as a habit. Weidman and Grossklags [26] noted the role of habit formation in fostering long-term acceptance of authentication mechanisms. Arnold et al. [5] studied the emotional reactions of college students to mandatory MFA and reported strong negative feelings such as frustration and stress. Over half of the participants reported that MFA caused them to miss deadlines, and only 40% thought it was worth the inconvenience. Their work is particularly relevant to our investigation into user perceptions of subjective load and responses to safety cues.

Usability of 2FA Mechanisms

Marky et al. [17] conducted interviews with participants from diverse backgrounds to explore the usability of various 2FA methods. Their findings indicate that users prefer solutions that seamlessly integrate into their daily lives rather than those focused solely on enhancing technical security. Many existing 2FA methods fail to meet user satisfaction. Das et al. [10] analyzed over 12,500 user reviews and found that most users were resistant to 2FA because of the complexity of setup, poor cross-device support, and unclear recovery processes. Consequently, SMS-based codes remain the most popular method despite their weaker security, while professional authentication apps and hardware tokens continue to suffer from low adoption. Issues related to account recovery and a lack of guidance were also major sources of frustration. Other studies [12, 18, 20, 24] revealed that when 2FA mechanisms become inconvenient or poorly integrated into users' daily life, resistance increases.

Security Notifications, Recovery, and Advanced 2FA Schemes

Security notifications are an important yet often overlooked aspect of the 2FA experience. Markert et al. [16] found that users typically rely on device and location information to determine the legitimacy of login notifications. However, confusing or ambiguous information can lead to misunderstanding, and frequent alerts may cause alarm fatigue. Only 22% of users in their study responded appropriately after receiving a suspicious notification. To mitigate such burden, Wiefling et al. [27] proposed risk-based authentication, which prompts additional verification only when anomalies are detected, striking a balance between security and convenience.

Furthermore, inconsistencies in 2FA interaction design across different platforms contribute to cognitive burden. Lyastani et al. [13] found that a lack of unified design patterns can lead to user confusion, increased cognitive burden, and elevated security risks. These inconsistencies, along with unclear instructions and complex recovery processes, contribute significantly to user frustration.

3 Methodology

This section describes the survey design, participant recruitment, data collection and pre-processing procedures, and the key variables used in our analysis. Our study aims to investigate user perceptions of 2FA, with a focus on usability, perceived workload, and attention to security cues.

3.1 Survey Design

We designed two online surveys and distributed them separately to two participant cohorts: (1) general Internet users recruited via Amazon Mechanical Turk (MTurk) and (2) university students from a U.S. public university. Both surveys shared a common structure, with minor adaptations for the campus context. In particular, the university survey included four additional questions about the campus rollout of Duo 2FA, which had been mandatory for approximately one year before data collection. The MTurk version contained 23 questions, while the university version contained 27 questions. Participants were allowed up to one hour to complete the survey; MTurk participants completed it in 23 minutes on average.

The survey comprised four parts:

1. **Demographics and 2FA background (Q1–Q9):** Participants provided information on age, gender, education, tech-savviness, the number of accounts with 2FA enabled, reasons for adopting 2FA (including adoption mode), and usage frequency and duration.
2. **2FA context and methods (Q10–Q11):** Participants indicated where they use 2FA (e.g., school, work, personal) and which authentication methods they employed (e.g., SMS, push notification, authenticator apps, security keys).

3. **Usability and workload (Q12–Q17):** Based on the NASA-TLX framework, six Likert-scale statements (1 = strongly disagree to 5 = strongly agree) assessed users' mental and physical demand, time pressure, effort, annoyance, and perceived difficulty during 2FA login. Follow-up prompts collected brief explanations when applicable.
4. **Security attention and recovery (Q18–Q27):** Participants were asked about their attention to security cues (e.g., device, IP address, location, and time information presented during authentication), their experience with 2FA-related access loss, and what they disliked about their current 2FA systems.

3.2 Data Collection and Demographics

We initially received 133 MTurk submissions and 119 university survey responses from participants based in the United States. MTurk participants were required to have a HIT approval rating of at least 98% and a minimum of 10,000 approved HITs. A screening question excluded participants without any 2FA-enabled accounts. Each MTurk participant received $1.00 in compensation. University participants were recruited through Computer Science courses and were required to be at least 18 years old. All procedures were approved by the institution's Institutional Review Board.

To ensure data quality, survey questions were mandatory where appropriate, and open-ended responses were manually reviewed. We removed responses with missing key fields or clearly irrelevant or nonsensical text. After cleaning, the final dataset included 212 valid responses (129 MTurk and 83 university), corresponding to the exclusion of 40 submissions (15.9%).

Table 1 summarizes the sample demographics (N = 212). Gender information was provided by 199 respondents (6.1% missing): 144 identified as male (72.4% of valid responses), 53 as female (26.6%), and 1 as non-binary/third gender and 1 preferred not to say (both 0.5%). Age information was provided by 197 respondents (7.1% missing), ranging from 19 to 71 years, with a mean age of 33.6, a median of 31, and an interquartile range of 24–39.

Table 1. Demographic summary of survey participants (N = 212)

Characteristic	**Category**	**n (%)**
User group	MTurk	129 (60.8%)
	Sac State	83 (39.2%)
Gender	Male	144 (67.9%)
	Female	53 (25.0%)
	Non-binary / third gender	1 (0.5%)
	Prefer not to say	1 (0.5%)
	Missing	13 (6.1%)
Age		19–71 (M = 33.6, SD = 11.1; n = 197)

3.3 Measures

To examine how enforcement context and user characteristics relate to perceived workload and usability of 2FA, we define a set of key variables for analyzing survey responses.

Adoption Mode and User Group The primary independent variables in our comparative analysis are:

- Adoption Mode: Participants reported how they initially enabled 2FA, categorized as *Mandated*, *Incentivized*, and *Voluntary*. Our comparisons mainly focus on the *Mandated* and *Voluntary* groups to assess how enforcement impacts usability perceptions.
- User Group: A binary variable distinguishing between university users and general Internet users (i.e., MTurk participants).

NASA-TLX Workload We adopted the NASA Task Load Index (NASA-TLX) to assess the subjective workload of users during the use of 2FA. It is a widely used instrument in human-computer interaction and usability studies. Originally developed by NASA to assess operator workload, the NASA-TLX captures subjective workload across six dimensions: psychological demand, physical demand, time pressure, operational effort, frustration, and task performance. Unlike single satisfaction or ease of use score, it can reveal the source of usability problems in more detail. In addition, the scale results are easy to be compared across groups and statistically analyzed, which can be used as a quantitative basis for usability differences between different adoption patterns and user groups. Each response is converted into a numerical score for analysis.

Security-Attention Variables During 2FA login, many services display contextual security information (e.g., IP address, device type, and location). A key goal of this study is to understand whether users notice and leverage these cues. To measure users' attention to contextual security information based on their responses, the survey included four questions. Questions Q18_1 to Q18_3 asked participants about their engagement with IP- and device-related information in authentication contexts, while Q20 asked how often they pay attention to displayed contextual details such as IP address, time, or location.

All four items used a five-point frequency scale (*Never*, *Rarely*, *Sometimes*, *Often*, *Always*). For quantitative analysis, we mapped these responses to numerical values from 1 to 5, respectively. In the analyses reported below, Q20 is used to capture how often participants attended to contextual details displayed during authentication, such as IP address, time, or location, while the Q18 items are used for descriptive summaries and correlation analysis.

3.4 Data Preparation

Our dataset includes responses from two cohorts: university users and general Internet users. During data cleaning, we retained only completed responses and removed submissions with missing key fields or clearly invalid/irrelevant open-ended answers.

For analysis, five-point Likert-scale items were mapped from 1 (*strongly disagree/never*) to 5 (*strongly agree/always*). Technical proficiency (Q6), 2FA usage duration (Q7), and login frequency (Q9) were recoded into ordered numeric scales (Table 2). For workload measures, each NASA-TLX item was coded on a 1–5 scale for the six dimensions (mental demand, physical demand, temporal demand, effort, frustration, and perceived difficulty). We also computed an overall workload index, `TLX_total`, as the mean of the six subscales.

Table 2. Recoding of key ordinal variables used in the analysis

Question	Coding scheme
Technical proficiency (Q6)	"Somewhat", "Moderately", "Very", "Extremely" → 1–4
2FA usage time (Q7)	"Less than a year", "1–2 years", "More than 2 years" → 1–3
Login frequency (Q9)	"A few times a year", "A few times per month", "A few times per week", "1 or more times a day" → 1–4

3.5 Analytical Strategy

Quantitative Analysis Our quantitative analysis consists of three steps. First, we computed descriptive statistics to summarize the sample and to identify potential differences between two cohorts (e.g., age, gender, education level, 2FA methods, and the distribution of compulsory and voluntary adoption). Second, we conducted cross-sectional group comparisons of perceived workload and usability results. We compared TLX total and subscale scores (e.g., perceived difficulty and frustration) across adoption modes (mandatory vs. voluntary) and across user groups (university vs. MTurk). We selected statistical tests based on the scale and distribution of each variable. Third, we used regression models to estimate the association between adoption mode and workload while adjusting for potential confounders. For continuous results, we used linear regression with adoption mode and user group as primary predictors, included their interaction, and controlled for covariates such as technical proficiency, years of use, login frequency, age, gender, and 2FA type. For binary results, we used logistic regression with the same set of predictors and controls. These models allow us to isolate the effect of mandatory adoption beyond simple mean comparisons and to test whether mandate effects differ by population.

Qualitative Analysis To perform a qualitative analysis, we analyzed open-ended survey responses to capture users' authentic experiences with 2FA. The questions asked participants to elaborate on why they found 2FA demanding (mentally or physically), why they felt rushed, what difficulties they encountered during login, why they did not enable 2FA on all accounts, and any additional comments about 2FA. After removing incomplete or clearly irrelevant responses, we preprocessed the remaining answers by lowercasing the text, removing punctuation, and filtering stopwords. We then performed manual coding by assigning descriptive labels and consolidating them into higher-level themes. Finally, we used the resulting themes to interpret the quantitative findings. For example, topics related to countdown timers and session timeouts helped explain higher temporal demand scores, while topics related to lost equipment and missing codes shed light on frustration scores.

3.6 Ethical Considerations

All participants were informed of the study's purpose, the nature of the questionnaire, the intended use of the data, and their rights as participants prior to data collection. Participation was voluntary, and informed consent was obtained before respondents proceeded with the survey. Participants were free to withdraw at any time without penalty, and only individuals aged 18 or older were eligible to participate. The study posed minimal risk to participants and collected no personally identifiable information. Data were limited to self-reported two-factor authentication (2FA) usage experiences, subjective workload measures (NASA-TLX), and relevant background variables. All responses were anonymized and associated only with randomly assigned identifiers. Ethics approval was obtained for both the original data collection and the secondary analysis at the respective institutions. All data were used exclusively for academic research purposes.

4 Results

This section reports the results of our analysis. We first summarize the overall distribution of NASA-TLX workload scores for 2FA use. We then compare perceived workload across adoption modes and user subgroups. Next, we examine how negative experiences relate to perceived workload. Finally, we perform correlation analyses among different variables.

4.1 Overall Workload Distribution

Figure 1 shows that the overall 2FA workload in our sample is low to moderate. Most TLX total scores fall between approximately 2.3 and 3.1, with a median slightly below 2.7. While 2FA is not perceived as highly burdensome by most respondents, the presence of outliers indicates substantial variation across individuals.

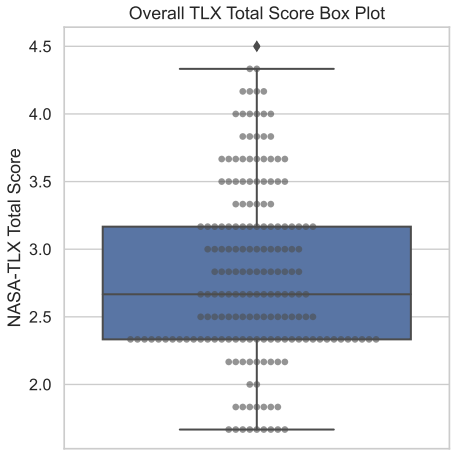

Fig. 1. TLX total score distribution.

4.2 Group Comparisons

We compared perceived workload across three grouping types: adoption mode (voluntary, incentivized, mandatory), user group (university vs. MTurk), and gender.

Figure 2 summarizes average NASA-TLX subscale scores by adoption mode. Mandated users report higher *temporal demand*, *frustration* and *subjective effort* than voluntary users, suggesting that enforcement is associated with feeling more "rushed" and "troubled" during authentication. The incentivized group generally falls between the mandated and voluntary groups.

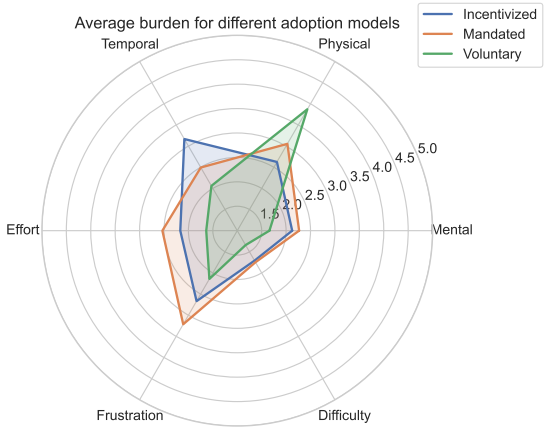

Fig. 2. Average burden for different adoption models.

Figure 3 shows the distribution of total TLX scores across adoption modes. The mandated group has a higher median and a wider interquartile range than the voluntary and incentivized groups, indicating both higher typical workload and greater variability under mandatory enforcement. In contrast, the voluntary group is more tightly clustered around lower TLX scores, suggesting a generally lighter perceived burden.

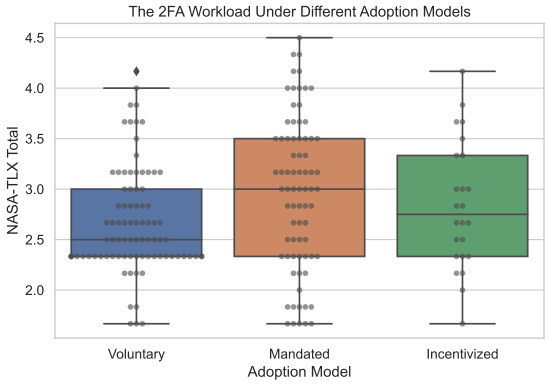

Fig. 3. The 2FA workload under different adoption models.

We next compared university and MTurk respondents. As shown in Figure 4, TLX total scores are broadly similar across the two cohorts, although the student group shows a slightly higher median and a wider spread. In contrast, the MTurk group's scores are more tightly clustered, with fewer extreme high or low values, indicating a more consistent overall experience without particularly severe or unusally easy cases.

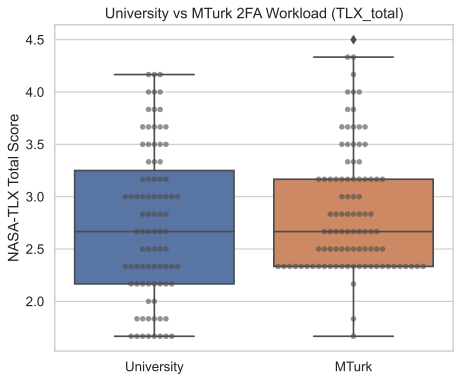

Fig. 4. Workload for students vs non-students (TLX_total).

Figure 5 further compares average scores across the six NASA-TLX dimensions. University students report higher average *mental demand*, *temporal pressure*, *effort*, *frustration* and *difficulty*, whereas the MTurk group reports higher *physical demand*, indicating that their burden was more often associated with physical actions, such as switching devices or retrieving authentication hardware.

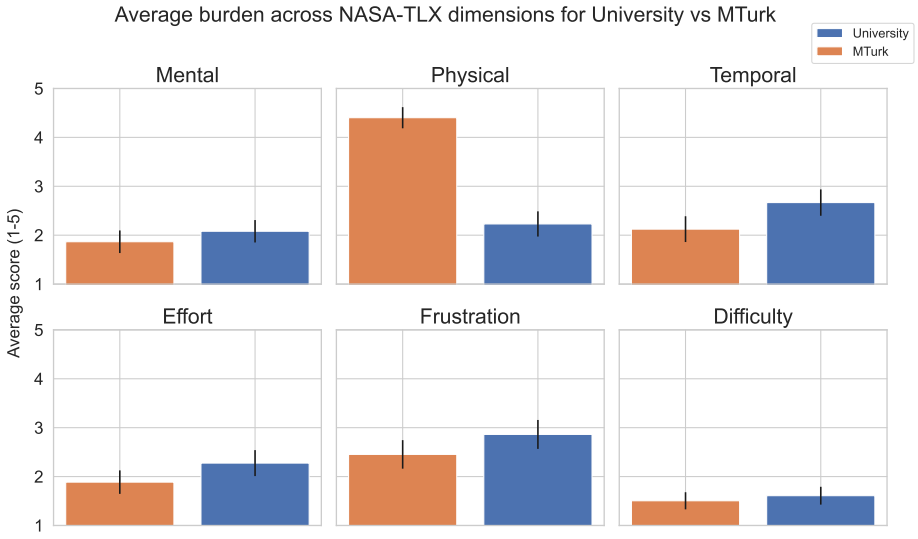

Fig. 5. Average burden across NASA-TLX dimensions for students vs non-students.

4.3 Impact of 2FA-related access loss Experiences

We next examined whether negative experiences with 2FA are associated with greater perceived workload. Figure 6 compares TLX total scores between respondents who reported access loss experience due to 2FA and those who had not. The "Yes" group has a noticeably higher median and a wider score distribution, indicating both elevated perceived workload and greater variability in individual experiences. In contrast, users in the "No" group has lower and more tightly clustered TLX scores, suggesting that without prior 2FA-related access loss incidents, 2FA tends to be perceived as only a light to moderate burden. This finding underscores the lasting impact of negative security experiences on user sentiment.

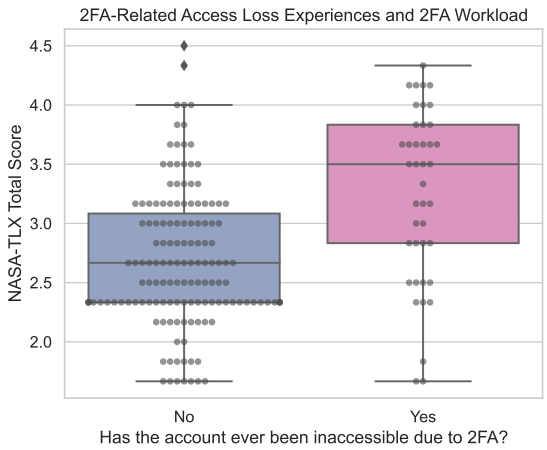

Fig. 6. 2FA-related access loss experiences and 2FA workload.

4.4 Correlation Analysis

Understanding the relationships between user attributes and perceived workload is a critical part of our analysis. Since variables such as age, technical proficiency, 2FA usage duration, 2FA usage frequency, and the NASA–TLX scores are ordinal or approximately ordinal in nature, we used Spearman's rank correlation to examine their associations.

Figure 7 presents the correlation matrix between these user characteristics and the six TLX dimensions, as well as the overall TLX total score. As expected, all six TLX subscales are positively and significantly correlated with each other and with `TLX_total`. Among them, *effort*, *frustration* and *temporal* exhibit the strongest correlations with the total score (coefficients ranging from 0.64 to 0.76), suggesting that these factors are the primary drivers of perceived workload during 2FA. *Mental demand* also shows a strong correlation with the total score, while *physical demand* and *task difficulty* are slightly less influential.

Regarding background variables, age shows a moderate positive correlation with physical demand (approximately 0.51), indicating that older users were more likely to perceive physical inconvenience when using 2FA. At the same time, age is negatively correlated with the frequency of 2FA usage (`frequency_2fa`) (approximately -0.44), indicating that young users engage in 2FA-required logins more frequently. Notably, usage frequency itself is negatively correlated with physical demand (approximately -0.43), implying that more frequent users may experience less physical burden. This is likely due to increased familiarity and operational fluency. By contrast, both `tech_level` and `duration_2fa` have weak or negligible correlations with TLX scores, indicating a limited effect on perceived workload in our sample.

We also examined whether responses to contextual security information is associated with workload. Figure 8 shows correlations between `TLX_total` and

four survey-based attention measures capturing how often participants attend to security cues in different contexts. The four attention measures are moderately to strongly correlated with one another (approximately 0.5 to 0.75), suggesting similar response patterns across contexts. In contrast, their correlations with `TLX_total` are weak (approximately −0.11 to 0.03), indicating no meaningful association between self-reported attention to security cues and overall workload in our sample.

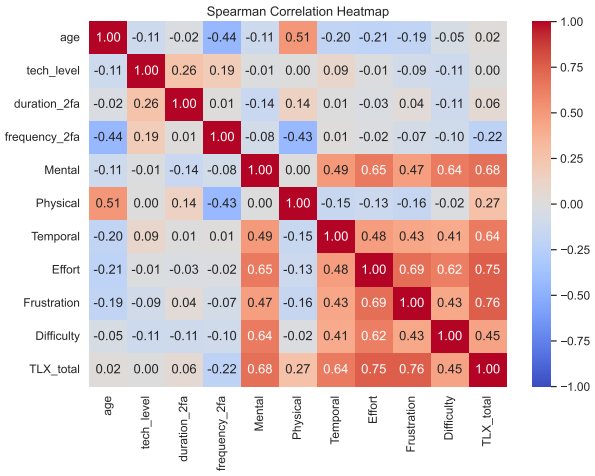

Fig. 7. Spearman correlation heatmap among ordinal variables and TLX scores.

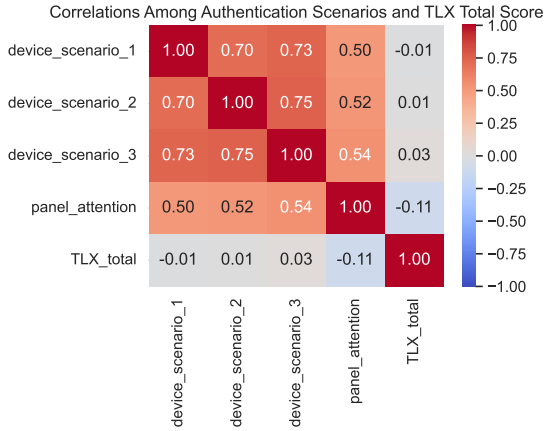

Fig. 8. Correlations Among Authentication Scenarios and TLX Total Score.

4.5 Effect of Notification Attention Levels

We examined whether users who more frequently inspect security details in 2FA push notifications report higher perceived workload. As shown in Figure 9, participants were divided into a lower-attention group ($\leq$ *Rarely*) and a higher-attention group ($\geq$ *Sometimes*) based on how often they viewed the notification details panel. The resulting TLX total score distributions largely overlap in median and interquartile range, with both groups concentrated in the low-to-moderate workload range. This pattern is consistent with the correlation analysis above, which shows weak associations between security push notification attention measures and `TLX_total`.

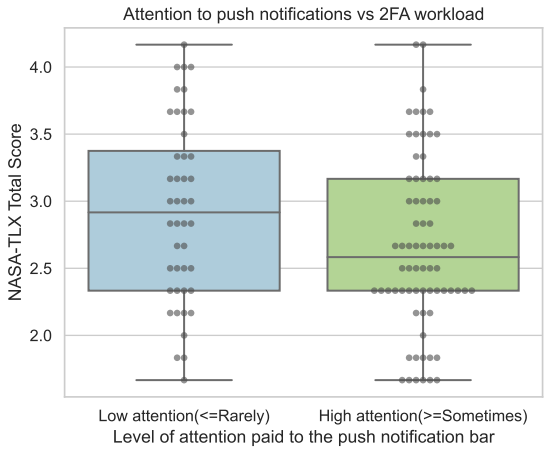

Fig. 9. Attention to push notifications vs 2FA workload.

4.6 Regression Analysis

We fit a multiple linear regression model with `TLX_total` as the dependent variable to estimate the association between key predictors and overall perceived 2FA workload while controlling for other factors (Figure 10). Prior 2FA-related access loss is the strongest predictor. Respondents who reported a access loss have substantially higher TLX scores, even when other factors are held constant. Adoption mode is the next most influential factor. Mandatory 2FA use is associated with higher perceived workload than voluntary or incentivized adoption. This supports our earlier findings that enforced usage introduces more friction and user burden. In contrast, the coefficient for `Student` is negative, indicating that after controlling for age, adoption mode, and access loss experience, students report slightly lower overall workload than non-students. Age has a near-zero coefficient, suggesting limited association with workload in this model.

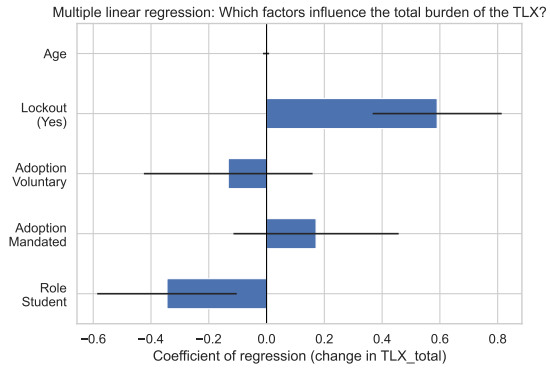

Fig. 10. Multiple linear regression results.

4.7 Qualitative Analysis of Open-Ended Responses

To gain deeper insight into users' actual experiences with 2FA, we conducted a qualitative analysis of participants' open-ended responses. These responses illuminated the cognitive, temporal, and emotional challenges users face while using 2FA, which complement the quantitative findings from the NASA-TLX survey. We analyzed user responses to several open-ended prompts related to (1) mental and temporal workload, (2) sources of effort and frustration, (3) reasons for not enabling 2FA on all accounts, and (4) general feedback or suggestions. We applied TF-IDF and word frequency techniques to extract key terms, and then grouped them semantically similar responses into distinct themes.

Table 3 summarizes the most frequent keywords from each question. For example, participants frequently mentioned *time pressure*, *countdown*, and *slow or unstable network* in relation to mental and temporal workload (Q10, Q12), particularly in high-pressure situations such as exams or assignments. Responses regarding effort and frustration (Q14, Q16) often cited concerns about the *multi-step login process* and the inconvenience of *switching back and forth between different devices*. Reasons for not enabling 2FA on all accounts (Q24) included *inconvenience*, *complexity*, and *unclear instructions*. In responses to the open comment question (Q25), many users expressed a desire for *clearer guidance*, *better recovery options*, and *fewer access loss risks*.

From this analysis, we identified five recurring themes, summarized in Table 4. These themes help interpret the common sources of user burden:

1. **Time Pressure and Countdown Pressure**. Users often felt rushed due to short expiration windows and session timeouts, especially during high-stakes tasks like tests or assignments. Some expressed anxiety that authentication would "expire before I even finished typing."
2. **Process Complexity and Multi-step Friction**. Participants described the 2FA process as unnecessarily complicated, citing repeated steps, app-

Table 3. Top keywords extracted from open-ended responses

Question	Top Keywords (TF-IDF / frequency)
Q10: Mental demand	time pressure, countdown, stressful, switching devices
Q12: Temporal demand	timeout, code expires, rush, delay, slow network
Q14: Effortfulness	multiple steps, switching apps, typing code, verification loop
Q16: Frustration sources	lost device, wrong code, waiting, repeated attempts
Q24: Why not enable 2FA	inconvenience, too many steps, confusing setup, device dependency
Q25: Additional comments	improve recovery, clearer instructions, reduce access loss

switching, and juggling between devices, which was particularly frustrating when repeated multiple times a day.

3. **Access loss and Recovery Anxiety**. Users who has access loss experience expressed fear and anxiety about using 2FA again. Many found the account recovery process to be slow, unclear, or overly restrictive.
4. **Device Dependency and Portability Issues**. Many respondents noted that 2FA becomes a major obstacle when a trusted device is unavailable, dead, or replaced, effectively locking them out of important accounts and services.
5. **Need for Clearer Guidance and More Forgiving Design**. Participants expressed frustration with technical jargon in security prompts and called for more intuitive guidance and flexible fallback mechanisms.

5 Discussion

5.1 Comparison with Prior 2FA Rollout Studies

Our analyses suggest that the overall perceived workload of 2FA, as measured by NASA-TLX, falls within a low to moderate range for most users. This is broadly consistent with prior 2FA rollout studies, which generally report that 2FA is perceived as manageable after initial adoption. For example, Colnago et al. [6] found that although CMU users initially anticipated 2FA to be burdensome, their actual experiences were often more positive than expected, with perceived burden decreasing over time. Similarly, Abbott and Patil [1] observed that while mandatory users were somewhat less satisfied than voluntary adopters, the transition to mandatory 2FA did not substantially reduce overall user acceptance. These findings and ours suggest that 2FA is often experienced as manageable rather than overwhelmingly burdensome.

At the same time, our results extend this literature in several important ways. First, rather than focusing primarily on satisfaction or acceptance, we examine 2FA through the lens of multidimensional workload. This reveals that the burden associated with mandatory adoption is concentrated in specific dimensions,

Table 4. Themes identified from qualitative coding of open-ended responses

Theme	Description	Example quote
Time pressure and countdown stress	Users describe feeling rushed by expiring codes, session timeouts, and strict time limits during critical tasks (e.g., tests, exams, deadlines).	"The code times out so fast that I feel rushed every time I log in."
Process complexity and multi-step friction	Burden caused by multi-step flows, switching between apps or devices, and repeating the same verification steps many times a day.	"There are too many steps just to log in, especially when I have to do it several times a day."
Access loss and recovery anxiety	Negative experiences of access loss, fear of losing access again, and frustration with slow or confusing recovery procedures.	"I was locked out once, and now I get anxious every time I see a 2FA prompt."
Device dependence and portability issues	Difficulties when the trusted device is not at hand, broken, or replaced, leading to disruption in everyday activities.	"If I forget my phone or it dies, I basically can't do anything that needs 2FA."
Need for clearer guidance and more forgiving design	Users ask for clearer instructions, less technical wording, and more forgiving mechanisms (e.g., easier fallback or recovery).	"The security messages are too technical; I just want clear steps telling me what to do next."

particularly temporal demand, effort, and frustration, rather than being uniformly distributed across all aspects of the user experience. Second, our findings show that negative experiences are not driven by enforcement alone. Many users described 2FA as merely an "extra step" or "a bit of a hassle" unless they had experienced 2FA-related access loss or recovery failure. Once such incidents occurred, however, the perceived burden increased substantially. Some respondents reported severe consequences, including missed exams and extended periods of inaccessibility. These perspectives help refine prior interpretation of 2FA rollout studies. While earlier work emphasizes the role of mandatory adoption in shaping user attitudes, our results indicate that the interaction between enforcement and failure scenarios is critical. Mandatory 2FA becomes particularly burdensome when combined with punitive or poorly supported recovery processes.

Our findings are broadly consistent with prior campus-based studies but provide additional explanatory details. Like Abbott and Patil [1], we observe that mandatory adoption is associated with more negative user experience than voluntary adoption. However, our data further shows that a stronger indicator of

elevated workload is prior 2FA-related access loss rather than enforcement status alone. This complements the findings of Dutson et al. [11], who reported frequent access loss and more negative attitudes among students and faculty than staff, and Arnold et al. [5], who documented frustration, stress, and missed deadlines among university students under mandatory MFA. In our study, students reported slightly higher TLX scores than general Internet users, but qualitative responses suggest that this difference is largely driven by situational factors. In particular, students often encountered 2FA in high-pressure situations such as deadlines and exams, which amplified perceived burden. In summary, our main contribution is to show that 2FA workload is best understood not as a simple consequence of mandatory adoption, but as the result of interactions among enforcement context, recovery mechanisms, and situational factors.

5.2 Design Implications for 2FA Deployment

Rather than modeling 2FA adoption as a binary question of whether it should be mandatory, our findings emphasize the complex influence of multiple factors. Mandatory 2FA policies do increase users' perceived time pressure, effort, and frustration, but not every mandated user experiences a high burden. Thus, the more important design question is not simply whether 2FA should be required, but how enforcement policies can be implemented in a way that is fair, supportive, and user-friendly.

Based on users' self-reported behavior, our results show that attention to security-related notifications is not associated with higher perceived workload. This challenges the common assumption that increasing the visibility of security information inevitably disrupts the user experience. Instead, users' attention to contextual security information may reflect habitual vigilance rather than additional cognitive burden. While excessive or frequent prompts can still lead to notification fatigue, as noted in previous studies [2, 3, 14, 22], our findings imply that designers can incorporate clearer and more informative contextual cues without necessarily increase user burden.

At the same time, the top themes derived from the qualitative analysis, which are countdown pressure, multi-step verification, device dependency, and complex recovery procedures, suggest that future efforts should prioritize creating more forgiving authentication systems. Prior work has likewise emphasized the importance of recovery support, clearer guidance, and reducing friction in multi-factor authentication [4, 9, 13, 17]. Instead of focusing solely on strengthening defenses, designers should also consider how to reduce friction, make recovery easier, and support users through challenging scenarios. In particular, reducing avoidable access loss and providing clearer fallback paths may do more to improve long-term acceptance than simply adding more factors or more frequent prompts.

In summary, our study reveals that the perceived burden of 2FA not only from its enforcement but also from how it is implemented and experienced. Well-designed interfaces, intuitive guidance, and robust recovery mechanisms may do more to improve long-term acceptance than simply adding more factors or controls. It is entirely possible that 2FA can evolve into an effective, reliable,

and unobtrusive protection method rather than a punitive mechanism users are forced to endure.

5.3 Limitations

This study has several limitations. First, it is based on cross-sectional, self-reported survey data. Thus, the results capture perceived workload and reported experiences rather than directly observed behavior, and do not support causal claims. Second, our sample is limited to two populations: MTurk workers and students from a single U.S. public university, which may limit generalizability to other user groups. Third, our measures of attention to contextual security cues should be interpreted with caution. Not all 2FA systems expose the same information (e.g., IP address, location, or device details), so lower reported attention may reflect limited cue availability rather than reduced user vigilance. Finally, while we applied response-quality screening and manually reviewed open-ended responses, low-effort or inattentive responding cannot be entirely ruled out.

6 Conclusion and Future Work

Our analysis of survey data from 212 participants reveals that 2FA-related access loss are the strongest contributor to increased workload, while mandatory 2FA adoption is associated with higher time pressure, effort, and frustration compared to voluntary or incentivized adoption. Differences between university users and general Internet users are relatively small after controlling for relevant factors, and attention to security cues does not meaningfully increase perceived workload. Overall, the findings indicate that 2FA burden stems primarily from enforcement policies and negative usage experiences rather than from security prompts themselves, suggesting that improving recovery processes, reducing access loss, and providing clearer verification guidance are more effective ways to improving usability.

In future work, we plan to move beyond self-reported survey data by examining real-world 2FA usage logs within a large organization. In particular, we aim to analyze behavioral indicators such as authentication completion time, repeated login attempts, device switching, recovery requests, and unsuccessful access episodes, enabling a more direct assessment of the relationship between perceived workload and actual authentication behavior. We also plan to conduct a longitudinal study to track how users' workload, attitudes, and coping strategies evolve over time, especially before and after mandatory 2FA deployment and after negative experiences such as 2FA-related access loss or recovery failure. Finally, building on our qualitative findings, we will design and evaluate targeted improvements to lockout and recovery user experience, including clearer fallback guidance, more transparent recovery workflows, and less disruptive authentication prompts. We will assess whether these changes can reduce frustration, effort, and time pressure during authentication.

References

1. Abbott, J., Patil, S.: How mandatory second factor affects the authentication user experience. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. pp. 1–13 (2020)
2. Akanda, M.M.R.R., Lacy, A., Saxena, N.: SoK: Inaccessible & insecure: An exposition of authentication challenges faced by blind and visually impaired users in state-of-the-art academic proposals. In: 34th USENIX Security Symposium (USENIX Security 25). pp. 1393–1413 (2025)
3. Akanda, M.M.R.R., Mahdad, A.T., Saxena, N.: Broken access: On the challenges of screen reader assisted two-factor and passwordless authentication. In: Proceedings of the ACM on Web Conference 2025. pp. 1116–1128 (2025)
4. Amft, S., Ali, M., Preuveneers, D., Albayram, Y., Volkamer, M., Acar, Y.: "We've Disabled MFA for You": An evaluation of the security and usability of multi-factor authentication recovery deployments. In: Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security (CCS '23). pp. 1380–1394. ACM, Copenhagen, Denmark (2023)
5. Arnold, D., Blackmon, B., Gibson, B., Moncivais, A.G., Powell, G.B., Skeen, M., Thorson, M.K., Wade, N.B.: The emotional impact of multi-factor authentication for university students. In: CHI Conference on Human Factors in Computing Systems Extended Abstracts. pp. 1–4 (2022)
6. Colnago, J., Devlin, S., Oates, M., Swoopes, C., Bauer, L., Cranor, L., Christin, N.: "It's not actually that horrible" exploring adoption of two-factor authentication at a university. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. pp. 1–11 (2018)
7. Cybercrime Support Network: MFA Consumer Sentiment Report. https://fightcybercrime.org/mfa-consumer-sentiment-report/ (2023), accessed: 2026-04-19
8. Das, A., Bonneau, J., Caesar, M., Borisov, N., Wang, X.: The tangled web of password reuse. In: NDSS. vol. 14, pp. 23–26 (2014)
9. Das, S., Kim, A., Jelen, B., Huber, L., Camp, L.J.: Non-inclusive online security: Older adults' experience with two-factor authentication. In: Proceedings of the 54th Hawaii International Conference on System Sciences (HICSS). pp. 5392–5401 (2021)
10. Das, S., Wang, B., Tingle, Z., Camp, L.J.: Evaluating user perception of multi-factor authentication: A systematic review. In: Proceedings of the Thirteenth International Symposium on Human Aspects of Information Security & Assurance (HAISA 2019) (2019)
11. Dutson, J., Allen, D., Eggett, D., Seamons, K.: Don't punish all of us: measuring user attitudes about two-factor authentication. In: 2019 IEEE European symposium on security and privacy workshops (EuroS&PW). pp. 119–128. IEEE (2019)
12. Farke, F.M., Lorenz, L., Schnitzler, T., Markert, P., Dürmuth, M.: "You still use the password after all"–exploring FIDO2 security keys in a small company. In: Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020). pp. 19–35 (2020)
13. Ghorbani Lyastani, S., Bugiel, S., Backes, M.: A systematic study of the consistency of two-factor authentication user journeys on top-ranked websites. In: Network and Distributed System Security (NDSS) Symposium 2023 (2023)
14. Golla, M., Ho, G., Lohmus, M., Pulluri, M., Redmiles, E.M.: Driving 2FA adoption at scale: Optimizing two-factor authentication notification design patterns. In: 30th USENIX Security Symposium (USENIX Security 21). pp. 109–126 (2021)

15. Han, W., Li, Z., Ni, M., Gu, G., Xu, W.: Shadow attacks based on password reuses: A quantitative empirical analysis. IEEE Transactions on Dependable and Secure Computing **15**(2), 309–320 (2016)
16. Markert, P., Lassak, L., Golla, M., Dürmuth, M.: Understanding users' interaction with login notifications. In: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems. pp. 1–17 (2024)
17. Marky, K., Ragozin, K., Chernyshov, G., Matviienko, A., Schmitz, M., Mühlhäuser, M., Eghtebas, C., Kunze, K.: "Nah, it's just annoying!" a deep dive into user perceptions of two-factor authentication. ACM Transactions on Computer-Human Interaction **29**(5), 1–32 (2022)
18. Nanda, A., Jeong, J.J., Shah, S.W.A., Nosouhi, M., Doss, R.: Examining usable security features and user perceptions of physical authentication devices. Computers & Security **139**, 103664 (2024)
19. Nosenko, A., Cheng, Y., Chen, H.: Password and passphrase guessing with recurrent neural networks. Information Systems Frontiers **25**(2), 549–565 (2023)
20. Owens, K., Anise, O., Krauss, A., Ur, B.: User perceptions of the usability and security of smartphones as FIDO2 roaming authenticators. In: Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021). pp. 57–76 (2021)
21. Petsas, T., Tsirantonakis, G., Athanasopoulos, E., Ioannidis, S.: Two-factor authentication: is the world ready? quantifying 2FA adoption. In: Proceedings of the Eighth European Workshop on System Security. pp. 1–7 (2015)
22. Stanton, B., Theofanos, M.F., Prettyman, S.S., Furman, S.: Security fatigue. IT Professional **18**(5), 26–32 (2016)
23. Tetlay, A., Treharne, H., Ascroft, T., Moschoyiannis, S.: Lessons learnt from a 2FA roll out within a higher education organisation. arXiv preprint arXiv:2011.02901 (2020)
24. Turner, M., Schmitz, M., Bierey, M.M., Khamis, M., Marky, K.: Tangible 2FA– an in-the-wild investigation of user-defined tangibles for two-factor authentication. In: Nineteenth Symposium on Usable Privacy and Security (SOUPS 2023). pp. 245–261 (2023)
25. Von Zezschwitz, E., De Luca, A., Hussmann, H.: Survival of the shortest: A retrospective analysis of influencing factors on password composition. In: IFIP Conference on Human-Computer Interaction. pp. 460–467. Springer (2013)
26. Weidman, J., Grossklags, J.: I like it, but i hate it: Employee perceptions towards an institutional transition to BYOD second-factor authentication. In: Proceedings of the 33rd Annual Computer Security Applications Conference. pp. 212–224 (2017)
27. Wiefling, S., Dürmuth, M., Lo Iacono, L.: More than just good passwords? a study on usability and security perceptions of risk-based authentication. In: Proceedings of the 36th Annual Computer Security Applications Conference. pp. 203–218 (2020)

A Survey Questionnaire

The questionnaire items used in the analysis are listed below.

Q1. How many accounts have you enabled 2FA for?
- 1
- 2
- 3
- 4 or more

Q2. What is your age?

Q3. What is your gender?
- Male
- Female
- Non-binary/third gender
- Prefer not to say

Q4. What is your highest level of education?
- Some high school
- High school graduate
- Some college, no degree
- Associate's degree
- Bachelor's degree
- Master's degree
- Doctorate degree

Q5. Are you a:
- College student
- College faculty
- College staff
- None of the above

Q6. How tech-savvy are you?
- Not at all
- Somewhat
- Moderately
- Very
- Extremely

Q7. How long have you been using Two-Factor Authentication (2FA)?
- Less than 1 year
- 1–2 years
- More than 2 years

Q8. When you first enabled 2FA, why did you enable it?
- Mandated
- Incentivized
- Voluntary

Q9. How frequently do you log in with your 2FA authenticator?
- 1 or more times a day
- A few times per week
- A few times per month
- A few times per year

Q10.(1-3) Which accounts have you enabled 2FA for?

	Enabled	Not enabled	N/A
School account	☐	☐	☐
Workplace account	☐	☐	☐
Personal account	☐	☐	☐

Q11.(1-3) When you initially enabled 2FA, which two-factor authenticator(s) did you enable?

	Text message	Push	Phone Call	App Code	Security Key	N/A
School account	☐	☐	☐	☐	☐	☐
Workplace account	☐	☐	☐	☐	☐	☐
Personal account	☐	☐	☐	☐	☐	☐

Q12.(1-6) Do you agree or disagree with the following statements?

(a) Using my 2FA authenticator to login is mentally demanding (e.g. thinking, searching, deciding)

SD SWD N SWA SA
[ ] [ ] [ ] [ ] [ ]

(b) Using my 2FA authenticator to login is physically demanding

SD SWD N SWA SA
[ ] [ ] [ ] [ ] [ ]

(c) I feel rushed when having to use my 2FA authenticator to log in

SD SWD N SWA SA
[ ] [ ] [ ] [ ] [ ]

(d) Using my 2FA authenticator to login to my account requires too much effort

SD SWD N SWA SA
[ ] [ ] [ ] [ ] [ ]

(e) I feel annoyed when having to use my 2FA authenticator

SD SWD N SWA SA
[ ] [ ] [ ] [ ] [ ]

(f) I have difficulty using my authenticator to log into my account

SD SWD N SWA SA
[ ] [ ] [ ] [ ] [ ]

SD = Strongly disagree; SWD = Somewhat disagree; N = Neither agree nor disagree; SWA = Somewhat agree; SA = Strongly agree.

Q13. Why is using your 2FA authenticator to login mentally demanding?

Q14. Why is using your 2FA authenticator to log in physically demanding?

Q15. Why do you feel rushed when using your 2FA authenticator?

- Having to input the code before the countdown timer expires
- Other _____

Q16. Why do you feel that having to use your 2FA authenticator to login requires too much effort?

Q17. What difficulty did you have with using your 2FA authenticator to log in?

Q18.(1-3) Can you recognize the IP address from the device you logged in from memory?

	Never	Rarely	Sometimes	Often	Always	N/A
Personal computer	☐	☐	☐	☐	☐	☐
Workplace computer	☐	☐	☐	☐	☐	☐
Personal Phone	☐	☐	☐	☐	☐	☐

Q19. Have you used the push authenticator?
- No
- Yes

Q20. Did you ever pay attention to the IP address, time, or location that is displayed?
- Never
- Rarely
- Sometimes
- Often
- Always

Q21. Have you ever been unable to access your account because of 2FA?
- No
- Yes

Q22. Why were you unable to access your account because of 2FA?

Q23. Have you ever canceled an authentication request?
- No
- Yes

Q24. Why did you cancel the authentication request?
- Did not have 2FA authenticator
- Mistake
- Other _____

Q25. Have you ever not responded to an authentication request?
- No
- Yes

Q26. Why did you not respond to the authentication request?
- Did not have 2FA authenticator
- Other _____

Q27. What do you dislike about 2FA?