DuoLungo: Usability Study of Duo 2FA

Renascence Tarafder Prapty
University of California Irvine
rprapty@uci.edu

Gene Tsudik
University of California Irvine
gts@ics.uci.edu

***Abstract*—Multi-Factor Authentication (MFA) enhances login security by requiring users to use multiple authentication factors. MFA adoption has surged in recent years in response to the growing frequency, diversity, and sophistication of attacks. *Duo* is among the most popular MFA providers, used by thousands of organizations worldwide, including Fortune 500 companies and large educational institutions. However, its usability has not been investigated thoroughly or recently. Although prior work addressed technical challenges and user perceptions during initial implementation phases, there was no assessment of key usability metrics, such as average task completion time and System Usability Scale (SUS) scores. Moreover, relevant prior results are outdated, having been conducted years ago when the entire MFA concept was relatively new and unfamiliar to the average user.**

Motivated by the above, we conducted a long-term and large-scale Duo usability study. This study took place at the University of California Irvine (UCI) over the course of the 2024-2025 academic year and it involved $2,559$ unique participants. Our analysis is based on a large set of authentication log files and a survey of 57 randomly selected participants. The study reveals that the average overhead of a Duo Push notification task is nearly 8 seconds, a duration described by participants as short to moderate. Several factors influence this overhead, including the time of day when the task was performed, the participant's field of study, as well as their education/student level. The rate of authentication failures due to incomplete Duo tasks is 4.35%. Furthermore, 43.86% of survey respondents reported experiencing a Duo login failure at least once. The Duo SUS score is found to be 70, corresponding to a "Good" usability level: while participants generally find Duo easy to use, they also perceive it as annoying. On a positive note, Duo increases participants' sense of security regarding their accounts. Finally, participants described commonly encountered issues and provided constructive suggestions for improvement.

I. Introduction

The increasing sophistication and diversity of attacks, coupled with the limitations of traditional single-factor (password-based) authentication, motivate the widespread adoption of so-called Multi-Factor Authentication (MFA) techniques [1]. Though the trend towards MFA adoption is not new, it has accelerated in the last 5-10 years. While conventional password-based authentication remains vulnerable to phishing, key-logging, and brute-force attacks [2], MFA appreciably reduces the efficacy of such attacks. As of 2025, $\approx$ 98% of organizations worldwide support MFA, with large enterprises having an MFA adoption rate of 87% [3]. Also, nearly two-thirds of individual users now use MFA [4], which confirms its growing importance.

The concept of MFA is quite broad. The only unifying factor is that it must involve two or more modalities or factors, including: what you know (passwords or PINs), what you recognize (answers to personal questions with multiple-choice answers or pictures), what you are (biometrics), and what you have (various types of tokens). Although there are many combinations of these factors, most civilian/consumer settings use Two-Factor Authentication (2FA), with the primary (first) factor based on traditional passwords. The second factor is usually either biometrics or tokens.

Among 2FA providers, Duo is one of the most popular, recognized for its reliability, ease of use, and strong security features. Duo ranks among the top MFA providers globally, and is considered by some metrics to be the best 2FA app [5]. Although Duo does not publicly disclose the total number of its users, it is adopted by many major corporations, including Fortune 500 companies, e.g., Microsoft, Walmart, IBM, and Cisco [6]. Similarly, many large educational entities (e.g., University of Illinois [7], University of Florida [8], University of California, Berkeley [9], University of Arizona [10], and Columbia University [11]) have integrated Duo 2FA into their user authentication frameworks.

As technology evolves and users gradually become more accustomed to MFA, it is important to assess whether Duo meets current user needs and expectations. Despite its popularity, the usability of Duo remains underexplored. To the best of our knowledge, no previous research has assessed the overhead of Duo in terms of task completion time or determined Duo's System Usability Scale (SUS) score. Prior studies, e.g., [12] and [13], focused on technical challenges and user perception during the initial rollout of Duo, when its use was newly mandated within organizations. In these initial and transitional contexts, user attitudes were potentially influenced by disruptions to established authentication routines. For example, users accustomed to single-factor authentication might have perceived Duo as an added burden, whereas new users introduced to Duo from the outset likely viewed it as a normal part of the authentication process. Other prior research on the overhead and usability of push notification methods in 2FA systems [14], [15], conducted in 2017 and 2018, respectively, is simply outdated. Back then, 2FA technology was relatively new and users had not yet integrated 2FA processes into their authentication routines.

Symposium on Usable Security and Privacy (USEC) 2026
27 February 2026, San Diego, CA, USA
ISBN 978-1-970672-07-7
https://dx.doi.org/10.14722/usec.2026.23017
www.ndss-symposium.org, https://www.usablesecurity.net/USEC/

Given these limitations, we aim to gain a fresh perspective by examining Duo in a context where users are familiar with 2FA and encounter it on a regular basis. Moreover, by focusing on routine use rather than initial adoption, we can more accurately evaluate Duo's usability without any biasing effects introduced during transitional phases. To this end, we analyzed authentication logs collected over nine months, involving $2,559$ unique users, most of whom were students at the University of California, Irvine (UCI). These logs contain timing events that allow us to compute the overhead of Duo push notification tasks and the authentication failure rate. We also conducted a post-study survey – with randomly chosen 57 participants – to gain insights into user experiences and perceptions.

Our work addresses the following research questions:

RQ1. What is the overhead of the Duo push notification task? This overhead is determined by the task completion time. Completion means a user clicking the "Approve" or "Deny" button in a push notification or in the Duo app. On average, users take 7.82 seconds, which is generally perceived as short-to-moderate time.

RQ2. What factors influence the overhead of a Duo push notification task? Perhaps unsurprisingly, this turns out to be influenced by the time of day. Additional factors include the participant's major/field of study and their student/education level.

RQ3. How likely are users to experience authentication failures when using Duo? Incomplete Duo tasks lead to a failure rate of 4.35%. Also, over 43.86% of survey respondents reported experiencing at least one failed Duo login.

RQ4. What do users think about Duo? Most users express positive or mixed perceptions. Many find it annoying, however, they acknowledge its security benefits. Users note issues, such as problems with push notifications and complications when changing phone numbers or devices. They also provide constructive suggestions for improving the overall user experience.

RQ5. What is the SUS-based usability of Duo? Duo earns a "Good" usability level, with an average SUS score of 70. While Duo tasks are considered "easy", most users also view them as "annoying".

Through the combination of authentication log analysis and post-study survey, we comprehensively assess Duo usability and user perceptions. We believe that this provides a timely and expanded understanding of the Duo ecosystem.

Outline: Section II provides some background on MFA, and SUS. Section III describes the methodology of the user study. Section IV presents the results and their analysis. Section V compares our results to previous user studies. Section VI discusses the limitations of this study and provides recommendations to improve Duo MFA usability. Finally, Section VII concludes the paper.

II. Background

A. MFA

MFA is a security mechanism that mandates two or more verification factors for user authentication:

- **Knowledge**: Something the user knows, such as a password, PIN, or security question.
- **Cognition**: Something the user recognizes, such as familiar faces or answers to personal questions.
- **Possession**: Something the user has, such as a smartphone, security token, or smart card.
- **Inherence**: Something the user is (aka biometrics) based on, e.g., fingerprints.

The most common form of MFA is 2FA, which requires two factors.

One of the earliest forms of 2FA is the use of Automated Teller Machines (ATMs), which combines the possession of an ATM card and the Knowledge of a secret PIN [16]. More recent 2FA instantiations involve combining a username/password with either (1) one-time PINs or codes sent via SMS, email, or generated by specialized authentication apps, or (2) biometric verification, such as fingerprint or facial recognition.

Growing adoption of 2FA is driven by regulatory requirements and industry standards that mandate strong(er) authentication protocols. For example, the Payment Card Industry Data Security Standard (PCI DSS) and the General Data Protection Regulation (GDPR) emphasize the importance of robust authentication measures.

Despite its advantages, 2FA is not without challenges. Usability concerns, such as the complexity of managing multiple authentication factors and potential accessibility issues for users with disabilities, must be addressed to ensure widespread adoption [17]. Also, the security of MFA systems can be compromised if not properly implemented, since attackers naturally target the weakest link in the authentication chain.

B. SUS

The System Usability Scale (SUS) [18] is a time-tried, widely recognized, and efficient survey method for evaluating the usability of various products and systems. It consists of a standard set of ten questions (shown in Table I) where odd-numbered questions are positively phrased, and even-numbered questions are negatively phrased.

TABLE I: Standard SUS Questionnaire

No	Question
1	I think that I would like to use this system frequently.
2	I found the system unnecessarily complex.
3	I thought the system was easy to use.
4	I think that I would need the support of a technical person to use this system.
5	I found the various functions in this system were well integrated.
6	I thought there was too much inconsistency in this system.
7	I would imagine that most people would learn to use this system very quickly.
8	I found the system very cumbersome to use.
9	I felt very confident using the system.
10	I needed to learn a lot of things before I could get going with this system.

Each question offers five response options: Strongly Disagree, Disagree, Neutral, Agree, and Strongly Agree. A numerical SUS score is derived from an individual's responses to these questions, with the score ranging from 0 to 100. The procedure for computing a SUS score includes the following steps:

TABLE II: Adjective Ratings of SUS Scores

Adjective	Mean SUS Score	SUS Score Range
Worst Imaginable	12.5	0 - 25
Poor	20.3	25.1 - 39
OK	35.7	39.1 – 51.6
Good	71.4	51.7 - 73.5
Excellent	85.5	73.6 - 85
Best Imaginable	90.9	85.1 - 100

- Assign numerical values to responses: Strongly Disagree = 1, Disagree = 2, Neutral = 3, Agree = 4, and Strongly Agree = 5
- For odd-numbered questions, subtract 1 from the corresponding response value
- For even-numbered questions, subtract the response value from 5
- Sum up all resulting scores and multiply the total by 2.5 to obtain the final SUS score

The average SUS score is computed by averaging individual SUS scores of all participants. Adjective scaling [19] is then applied to interpret the absolute usability level of the product from the average SUS score. This scaling includes seven usability levels, ranging from "worst imaginable" to "best imaginable" usability. Table II shows the relationship between SUS score and usability levels.

III. Methodology

A. Setting

The study was conducted over nine months: from August 2024 to April 2025. It took place at UCI, in the Donald Bren School of Information and Computer Sciences. The school's IT department manages a Linux server cluster used by faculty, students, and staff. Access to this cluster is required for many courses. Students from all schools on campus who enroll in these courses have access to it. Duo serves as the second factor for user authentication to access the Linux cluster. User authentication proceeds as follows:

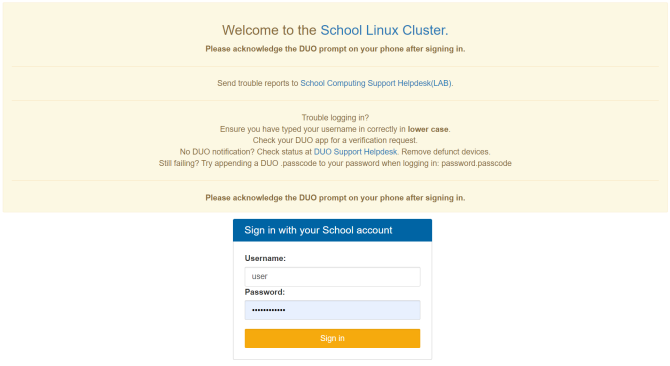


Fig. 1: Sign-In Page for Linux Cluster

1) **First Factor:** Entering one's userid and password (sign-in in Figure 1). Successful verification of the userid and password prompts Duo to send a push notification to the mobile device registered for the supplied userid.
2) **Second Factor:** Pressing the `Approve` or `Deny` button in the Duo app push notification. Figure 2 shows an example.

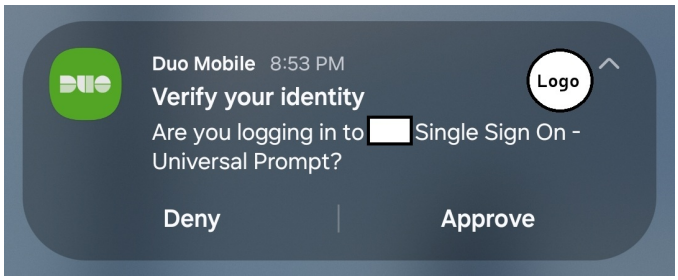


Fig. 2: Duo Push Notification

In some cases, a user cannot complete the Duo task through the push notification, e.g., if they clear the notification or the button on the push notification does not work. In such cases, the user opens the Duo app and presses the desired button.

Users without a smartphone use physical tokens. They can obtain a one-time passcode from the physical token and append it to the password in the "password.passcode" format. In certain situations, users with smartphones must also provide a Duo one-time passcode, instead of, or in addition to, pressing the approve button on the Duo push notification. Figure 3 shows an example of a one-time passcode generated in the Duo app.

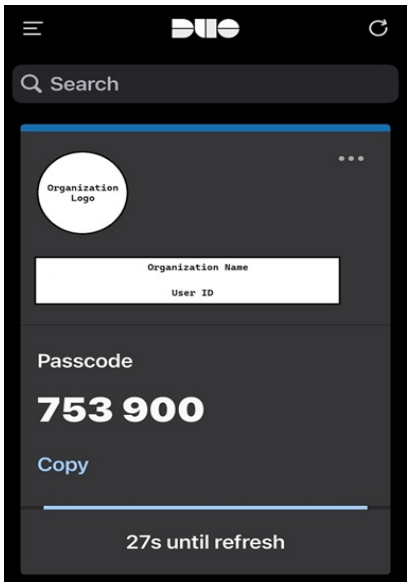


Fig. 3: Passcode generated inside Duo App

Duo log files were duly obtained from the school's IT staff. Each sign-in instance has either one or two log entries. For each failed primary factor (userid/password-based) authentication attempt, there is a single timestamped log entry. If the primary factor authentication succeeds, there are two timestamped log entries: (1) for the primary factor, and (2) for either a successful or failed second factor. Each log entry

contains a userid, which, along with the timestamp, links (1) and (2).

Although log files contained entries for two types of second factor authentication (Duo push notification and passcode), we can only reliably measure the overhead of push notification. It reflects how long a user takes to click on the Approve or Deny button. Because a user appends the passcode to the password, measuring the time spent retrieving the passcode from the user's device is infeasible. Log entries for primary and secondary authentication occur within milliseconds of each other and reflect the server's passcode verification time.

1) Website: Faculty, staff, and students use an internal website to access the online Linux cluster. Access to that website is restricted to individuals using the campus network or a Virtual Private Network (VPN). Figure 1 shows a snapshot of the sign-in page for this website, which has been slightly modified to preserve the anonymity of the authors.

2) Directory Crawler: UCI has a public directory that includes information on students, faculty, and staff. This directory contains non-sensitive details, such as names, userids, and email addresses for all university personnel. For students, the directory also provides major and student level information. For staff and faculty, it includes titles. To gather relevant information about study participants, we used a crawler specifically designed to collect all this information. This crawler uses a userid extracted from a log file as a search key.

B. Logistics and Data Cleaning

We analyzed a total of $96,048$ log entries from which we extracted $2,804$ unique userids. Some of these extracted userids were invalid as a result of user typos. (Naturally, they all corresponded to failed Duo authentication attempts.) To validate userids, we used the university's public directory – a userid was considered valid if it was found in the directory. This process yielded $2,561$ valid userids. 552 log entries corresponding to 243 invalid userids were removed.

Furthermore, over half of the log entries ($49,942$) lacked a userid and were thus excluded from the Duo overhead calculation, since we could not create a primary authentication log entry, Duo authentication log entry pair from them. However, we included these entries in our analysis to determine the rate of authentication failures caused by Duo. This is because some log entries with a "Primary authentication successful" message did not have any userid. However, they should still be included in the number of successful primary authentications.

Among the $2,561$ valid userids, three were removed. One belonged to one of the authors and the other two were from participants who opted out. 13 log entries corresponding to these three userids were also excluded from the analysis. In the end, the analysis of Duo overhead focused on $45,541$ log entries corresponding to $2,558$ userids.

C. Survey

We invited 800 randomly selected participants to complete a post-study survey, of whom 57 responded and completed the survey. On average, participants took about 5 minutes to finish the survey, and each was compensated with a $\$5$ Amazon gift card. The survey's purpose was to gather comprehensive feedback on several aspects of Duo use. It included questions about participants' overall perceptions of Duo, specifically their opinions about the time needed to complete Duo-related tasks and whether they were perceived as easy, difficult, fun, or annoying. Furthermore, the survey sought to understand whether participants believed that Duo improves the security of their accounts. Participants were also asked about their experiences of Duo authentication failures. To assess Duo usability, the survey incorporated SUS questions. The complete Survey questionnaire is included in the Appendix.

D. Ethical Considerations

This study received approval from UCI Institutional Review Board (IRB). No additional steps were added to the normal user workflow due to our study, i.e., no extra burden was imposed on the participants. All information collected using the directory crawler is publicly available and does not compromise participants' privacy. The IRB granted permission to use the authentication log files for research purposes without requiring explicit consent from participants. Nevertheless, to err on the side of safety and ethics, we informed all participants about the intended use of their authentication log data and provided them with the opportunity to express any concerns. Out of $2,561$ notified participants, only 4 responded: 2 expressed curiosity about the study and were provided with additional information, while another 2 requested to opt out, and we removed their data from the analysis. For the second phase of the study, which included the survey, participants provided explicit consent by completing an online consent form. Access to the survey questionnaire was granted only after consent was obtained.

IV. Results and Analysis

A. Demographics

Not surprisingly, almost all participants were students. The exact breakdown was: $2,521$ students, 9 faculty members, and 28 staff. Most students were undergraduates: $1,265$ seniors, 801 juniors, 248 sophomores, and 7 freshmen. There were also 200 graduate students.

Students represented 63 academic majors. However, the highest numbers corresponded to five majors related to Computer Science: $1,417$ in Computer Science (CmptSci), 251 in Software Engineering (SW Engr), 208 in Computer Science and Engineering (CSE), 125 in Data Science (DataSci), and 111 in Informatics (IN4MATX).

Although we lack specific information regarding participants' age, gender, and highest level of degree attained (unavailable in the public directory), reasonable assumptions can be made based on their academic status. Given that most participants are undergraduates, their ages very likely range between 18 and 22. All participants have at least a high school diploma. Faculty participants are likely to hold PhD degrees, while staff and graduate students typically have at least a bachelor's degree. Age distribution is estimated to be

approximately 90.73% aged $18-22$ and 9.27% over 22. In terms of highest degree attained, the breakdown is estimated at: 90.74% – high school, 8.91% – bachelors, and 0.35% – PhD.

Fortunately, we could obtain detailed demographic information from participants who completed the survey. Figure 4 shows the demographics. Gender-wise, two-thirds were male. Also, two-thirds had a high school diploma as the highest educational degree attained, indicating current undergraduate status. In terms of age, all participants were 19 to 29, with the majority (41) between 19 and 21. Participants were also asked about their frequency of using Duo: most use it $1-2$ times per day.

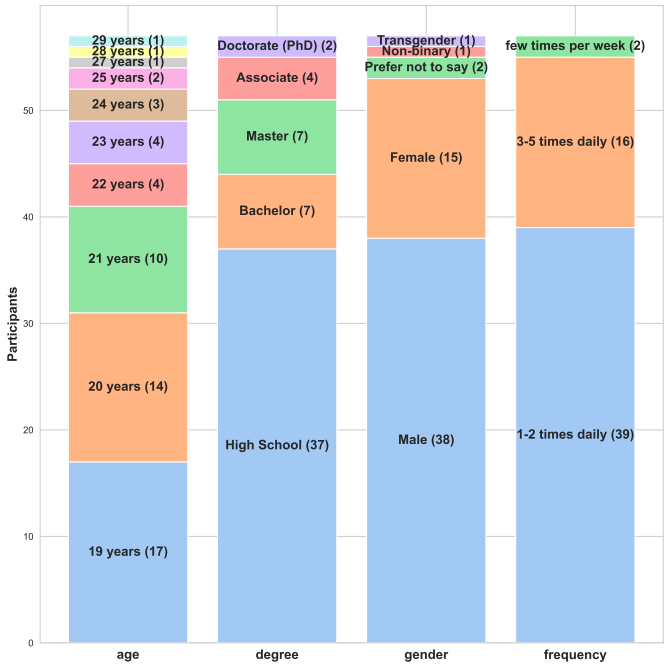


Fig. 4: Demography of Post Survey Participants

B. Qualitative Analysis of Open-Ended Responses

We analyzed all open-ended survey responses using thematic analysis. Two researchers independently performed inductive coding on the responses, iteratively grouping codes into higher-level themes. Disagreements were resolved through discussion. The themes reported in subsequent sections reflect the most salient issues raised by participants.

C. Duo Push Overhead

We define Duo push overhead as the elapsed time between a log entry for successful primary authentication and the subsequent log entry for successful Duo authentication. Each log entry provides detailed information, i.e., fields, including: userid, timestamp, and message. The userid helps us isolate log entries associated with individual users, subsequently sorting these entries in chronological order based on timestamps. The message field allows us to identify pairs of log entries where a successful primary authentication entry is immediately followed by a successful Duo authentication entry. The elapsed time between two such entries is the Duo overhead for a specific sign-in instance.

This process yielded data for $16,851$ instances. The average Duo push overhead across all instances is 9.29 seconds. The median is 6.89 seconds, with the first quartile (Q1) at 4.82 seconds, the third quartile (Q3) at 10.47 seconds, and an interquartile range (IQR) of 5.65 seconds. However, the distribution exhibits a pronounced long tail, as described in Section IV-C1. To mitigate the effect of outliers, we discarded values above the 95th percentile (i.e., over 24.71 seconds). After trimming, the average overhead decreases to 7.82 seconds, the median to 6.67 seconds, Q1 to 4.73 seconds, Q3 to 9.61 seconds, and the IQR to 4.00 seconds. We hereafter refer to the revised average as the average Duo push overhead.

The comparison between values before and after trimming demonstrates that trimming substantially reduces the mean and narrows the spread, while leaving the median and quartiles largely intact. Thus, trimming improves the validity of the reported average by reducing the impact of outlier values, without altering the central tendency of the distribution.

1) Distribution: Figure 5 illustrates the distribution of push overhead across all instances. The minimum and maximum are 1.15 and 59.4 seconds, respectively. Most overhead values fall between 3 and 15 seconds. Recall that this overhead corresponds only to successful sign-ins and is limited by the Duo push notification expiration time, set to 1 minute in our experimental setting.

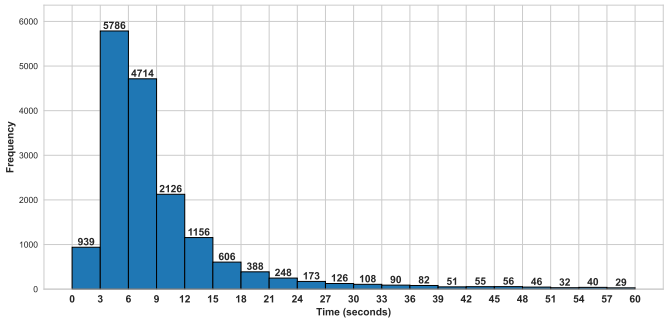


Fig. 5: Histogram of Duo Push Overhead

2) Statistical Tests: Since multiple Duo authentication attempts can originate from the same participant, the resulting push overheads are not independent across all observations. They are instead clustered within individuals. This motivates the use of a linear mixed-effects model that explicitly accounts for repeated measurements from the same participant. We partition push overheads into distinct groups based on participants' majors, student levels, and times of day when the Duo tasks were performed. To ensure that model estimates are based on sufficiently representative data, we restrict each analysis to groups that include at least 20 unique userids; groups that do not meet this criterion are excluded from the corresponding comparisons.

For each grouping factor (major, student level, time of day), we fit a linear mixed-effects model with Duo push overhead as the dependent variable, the grouping factor as a fixed effect, and a random intercept for each participant to capture within-subject correlation. We then use this model to test whether there are statistically significant differences among the levels of the factor. When an overall effect is detected, we perform post-hoc pairwise comparisons between group levels using contrasts of the model-estimated marginal means. To control for inflated Type I error due to multiple pairwise tests, we apply the Benjamini–Hochberg correction to the resulting p-values before determining which pairs are statistically significant. Only those pairs with corrected p-values below the chosen significance threshold of 0.05 are reported as exhibiting significant differences.

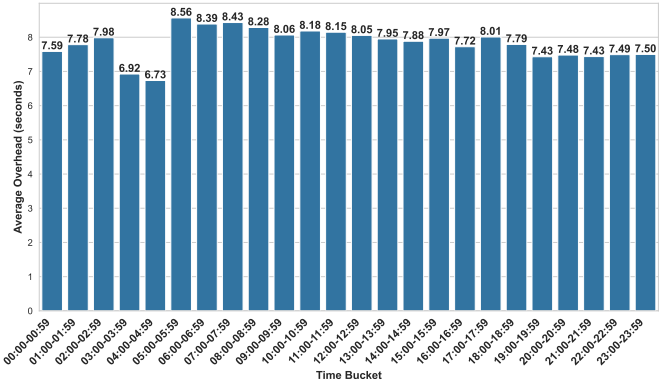


Fig. 6: Average Duo Push Overheads for Different Times of the Day

3) Time of Day: We partitioned sign-in instances into 24 groups, corresponding to each hour of the day. We then computed the average Duo push overhead for each group, as shown in Figure 6. Timestamps in log files were initially recorded as UTC and we converted them into the local time zone for this analysis. Two notable patterns emerged: (1) an overall decreasing trend throughout the day: from 5 a.m. to 5 p.m., and (2) an overall increasing trend from midnight to 3 a.m. Curiously, the two lowest average timings occurred late at night (3 a.m.–5 a.m.), followed immediately by the four highest averages (5 a.m.–9 a.m.).

The analysis revealed statistically significant differences in overheads based on the time of day. We identified 43 different hourly pairings with significant differences.

4) Student Level: We examined the effect of student level on overhead (Table III). Average overheads increase from freshmen (6.88 seconds) to graduate students (8.64 seconds). Undergraduate medians are similar (6.47–6.81 seconds), while graduate students have a higher median (7.46 seconds). Quartiles show a similar pattern: graduate students have higher Q1 (5.53 seconds) and Q3 (10.41 seconds), compared to undergraduates (Q1 $\approx$ 4.5–4.7 seconds, Q3 $\approx$ 9.5 seconds). Variability also differs: freshmen have the smallest IQR (3.74 seconds), sophomores the largest (5.11 seconds), and graduates remain relatively wide (4.89 seconds). We observed statistically significant differences in overheads between graduate and (sophomore, junior, and senior) undergraduate students.

TABLE III: Duo Push Overheads per Student Level (seconds)

Student Level	Mean	Median	Q1	Q3	IQR
Freshmen	6.88	6.81	4.49	8.23	3.74
Sophomore	7.71	6.47	4.55	9.66	5.11
Junior	7.74	6.60	4.65	9.51	4.86
Senior	7.76	6.62	4.72	9.47	4.75
Graduate	8.64	7.46	5.53	10.41	4.89

5) Effect of Major: Duo push overhead is influenced by participants' academic major (field of study), as shown in Figure 7. We identified 8 pairs of majors that exhibit statistically significant differences in push overheads:

1) Master's in Computer Science (MCS), Computer Science (CmptSci)
2) Master's in Computer Science (MCS), Computer Science and Engineering (CSE)
3) Master's in Computer Science (MCS), Data Science (DataSci)
4) Master's in Computer Science (MCS), Software Engineering (SW Engr)
5) Undeclared (Undclrd), Data Science (DataSci)
6) Undeclared (Undclrd), Computer Science (CmptSci)
7) Undeclared (Undclrd), Computer Science and Engineering (CSE)
8) Master's in Computer Science (MCS), Applied Math (AppMath)

In particular, MCS appears in the top four pairs. This major consists exclusively of professional master's students. In contrast, CSE, DataSci, BIM, Undclrd, and AppMath consist entirely of undergraduate students. Among participants with the SW Engr major, 96.41% are undergraduates, closely followed by CmptSci, with 94.7% undergraduates.

These findings further support the earlier observation that overhead is influenced by the student level.

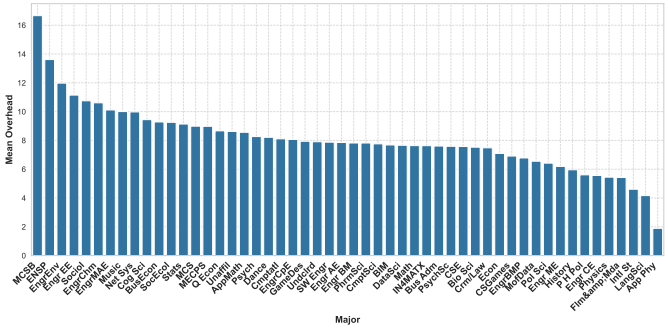


Fig. 7: Average Duo Push Overheads for Different Majors

D. Perceptions of Using Duo

We surveyed participants to gather their perceptions of Duo, using both open-ended and multiple-choice questions. The

former were posed first to minimize potential bias introduced by multiple-choice options.

Participants were asked separately about Duo push notification tasks and code input tasks, since the latter is generally more complex and time-consuming than the former. Considering that only some participants experienced the code input task, everyone was initially asked whether they had ever encountered it during the sign-in process. Only those who answered affirmatively were subsequently presented with questions about the code input task.

1) Multiple-Choice Responses: Multiple-choice questions for both the push notification task and the code input task included identical options. Figure 8 presents the number of participants who selected each option. For the push notification task, 73.68% of participants indicated they were either fine or okay with minor inconveniences. Only 7.01% expressed a desire to have it removed, whereas the remaining 19.31% acknowledged that it is necessary for security reasons, despite not liking it. For the code input task, 64% reported being fine or okay with it, 8% wanted it removed, and 28% did not like it, though they recognized its security benefits.

Most participants showed acceptance of both push notification and code input tasks, acknowledging their essential security benefits, despite finding them somewhat inconvenient.

2) Open Ended Responses for Duo Push Notification Task: From the open-ended responses regarding perceptions of the Duo push notification task, we identified four main themes: 1) Positive Sentiment, 2) Neutral Sentiment, 3) Negative Sentiment, 4) Suggestions/Improvements.

Each response was categorized with one or more themes. Figure 9 shows the division of responses according to associated themes. Among 57 responses, the largest category consisted of responses with positive sentiment. Interestingly, the second largest category was responses expressing both positive and negative sentiments. 12 responses included at least one suggestion for improvement.

One example of a response with both positive and negative sentiment is:

> "It works well and works seamlessly with my Apple Watch. It can be a pain when the service is poor."

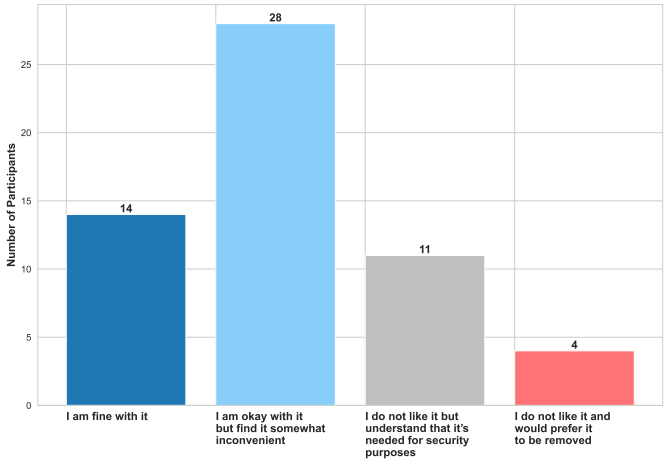
Another example response, with both neutral and negative sentiments:

> "It's fine, it serves its purpose. It's annoying that I have to sign in multiple times per day, but it's easy enough to accept from the push notification."

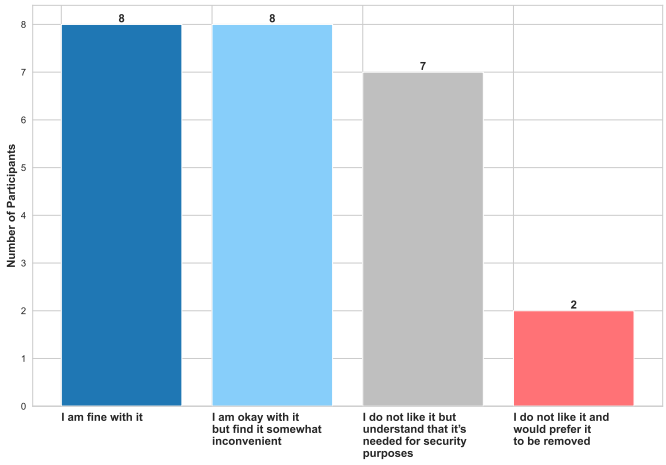
Suggestions for improvements included:

- Implementing a "Remember Device" feature
- Allowing acceptance directly from push notifications without opening the app
- Enabling fingerprint approval for Duo on desktop
- Making Duo functional without internet access

3) Open-Ended Responses for Duo Code Input Task: As before, responses fall into: 1) Positive Sentiment, 2) Neutral Sentiment, 3) Negative Sentiment. Figure 10 presents the response breakdown along these sentiments. The largest percentage of the responses conveys negative sentiment. Among



(a) Duo Push Notification Task



(b) Duo Code Input Task

Fig. 8: Multiple-Choice Responses: Perceptions of Duo

those with positive sentiment, two participants indicated that, although they were okay with the task, they preferred push notification. As expected, the negative sentiment for the code input task is higher than that for push notification.

E. Duo-induced Authentication Failure

We analyzed authentication log files and participants' survey responses in order to understand the percentage of authentication failures caused by Duo. We also investigated failure reasons via open-ended and multiple-choice questions.

1) Failure Statistics: Log files: The message field in each log entry allows us to calculate the percentage of Duo authentication failures. As mentioned earlier, log files contained $43,939$ entries for successful primary authentication and $1,913$ entries for Duo failure events. This denotes that 4.35% of authentication attempts failed due to Duo. Since this occurred after a user provided the correct userid and password, the failure rate appears to be significant.

To better understand the underlying causes of Duo authentication failures, we conducted a root cause analysis using the log files. The results are summarized in Table IV. The most

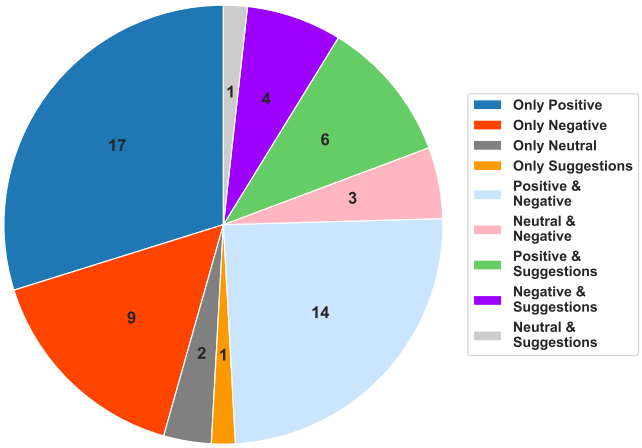


Fig. 9: Open Ended Responses: Perceptions of Duo Push Notification Task

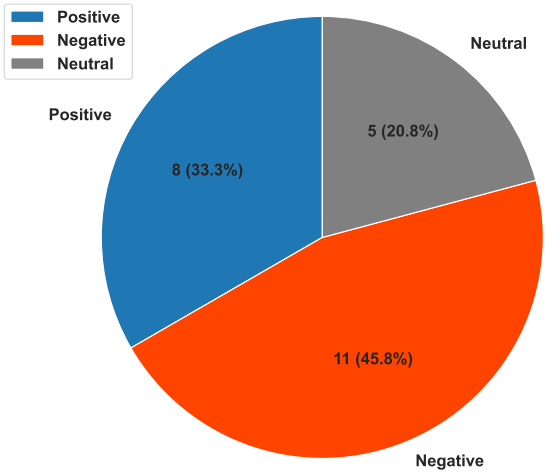


Fig. 10: Open Ended Responses: Perceptions of Duo Code Input Task

prevalent issue, accounting for 84.89% of all failures, is the timeout of Duo Push Notification tasks. This indicates that users either did not respond to the push request in time, or that the notification was not successfully delivered.

The second most common failure type, responsible for 7.73% of cases, stems from user accounts being disabled or lacking the necessary permissions to access the Linux cluster. These failures were identified both during the authentication and pre-authentication phases.

The third major contributor, comprising 4.91% of failures, involves users who are either not enrolled in Duo or have misconfigured Duo settings, resulting in no usable authentication factors.

Other less frequent failure types include incorrect passcode entries and users actively denying the authentication request via the Duo app.

TABLE IV: Duo Failure Types

Type	Count
Duo authentication was rejected - Login timed out.	1624
Duo preauth result was Your account is disabled and cannot access this application. Please contact your administrator.	131
Duo authentication was rejected - Your account is disabled and cannot access this application. Please contact your administrator.	17
User has no Duo factors usable with this configuration	89
Duo preauth result was Access is not allowed because you are not enrolled in Duo. Please contact your organization's IT help desk.	5
Duo authentication was rejected - Login request denied.	42
Duo authentication was rejected - Incorrect passcode. Please try again.	5

2) Failure Statistics: Survey Responses: We asked survey participants whether they were ever unable to sign in because they could not complete Duo 2FA tasks. Out of 57 participants, 25 responded affirmatively to experiencing failures, which corresponds to 43.86%. This highlights that, while Duo enhances security, it also causes occasional denial of service.

3) Failure Reasons: For participants who answered "yes" to experiencing Duo sign-in failures, we followed up by asking about exactly what happened, offering four options. Table V summarizes responses, with the majority indicating that they did not have access to their phone associated with their Duo account. This aligns closely with our log file analysis, which identified Duo Push Notification timeouts as the predominant failure type. Without access to their registered devices, users were unable to respond to the push notification before it expired, resulting in failed sign-in attempts. Those who chose

TABLE V: Duo Failure Reasons

Option	# Responses
Didn't have physical access to phone with Duo app	16
Duo task did not appear on my phone	5
Duo task timed out before I could complete them	1
Others, please specify in the following answer	3

the fourth option, "Other", were given an open-ended follow-up question asking them to specify the failure reason. From those open-ended responses, we identified four additional issues:

- **Device Performance Issues**: Problems related to the phone's performance affecting app functionality (e.g., app unloading due to slow processing) – reported by one participant.
- **Authentication Process Challenges**: Difficulties in completing the authentication process due to app unloading or other technical issues – reported by one participant.
- **Notification Functionality Issues**: Problems with the notification system, such as the inability to approve/deny directly from notifications – reported by one participant.
- **Connectivity Issues**: Problems related to a lack of access to cellular data or Wi-Fi – reported by two participants.

Each participant provided multiple reasons for their sign-in failures. Notably, one reported experiencing all the issues mentioned in the previous multiple-choice question as well as additional connectivity problems.

F. Multiple Code Inputs in a Single Authentication Instance

We asked 25 participants who experienced the code input task whether they ever had to input the code multiple times. Ten participants answered yes. In addition, we asked them participants to indicate the maximum number of codes they had to input in a single sign-in instance. Five participants selected 2, three chose 3, and two selected 6 or more. This suggests that participants are sometimes trapped in a frustrating cycle of repeatedly having to enter 6-digit codes.

G. Perception of Duo Overhead

We surveyed participants regarding their opinions about the time to complete Duo push notification and code input tasks using multiple-choice questions. The goal was to investigate the participants' perception of the Duo overhead. Results are presented in Figure 11. The majority of participants felt that Duo push notification tasks do not require much time. Conversely, responses about the Duo code input task were more evenly split between those who felt that (1) it does not take much time, or (2) it requires a moderate amount of time. Importantly, very few participants felt that either type of Duo task took a long time. Therefore, we can conclude that participants are generally not bothered by the overhead of a Duo task.

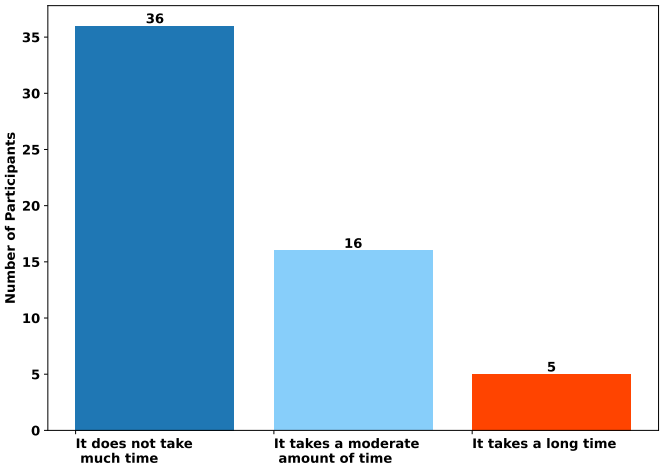
H. Usage of 2FA systems outside of school

We inquired whether participants used any other 2FA or MFA systems. Surprisingly, 38 participants (66.67%) responded "Yes". Therefore, many participants are likely accustomed to periodic MFA/2FA tasks, potentially making them more efficient and positively inclined towards completing Duo tasks, as compared to those who do not use 2FA outside of school.

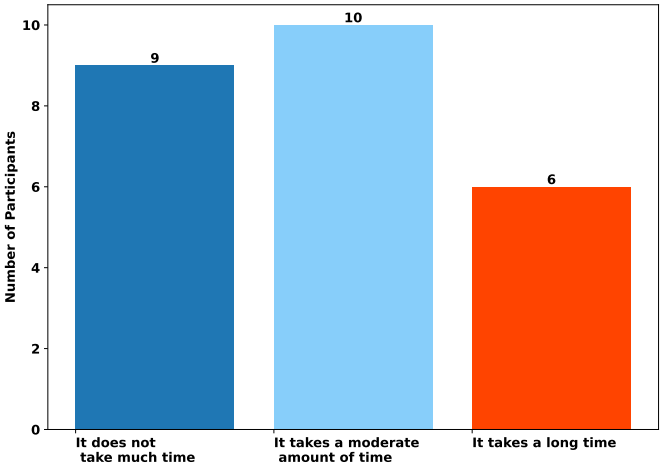
We also asked participants whether using Duo 2FA at school motivated them to adopt MFA/2FA for any of their non-school accounts. 40 (70.18%) replied "No". This suggests that – while most users tolerate MFA/2FA when it is mandatory – they are not inclined to adopt it voluntarily for security purposes. However, it is interesting that $\approx$ 30% were incentivized to voluntarily adopt 2FA for other accounts.

I. Perception of Security

We asked participants two similar multiple-choice questions to assess how Duo affects their perception of account security. Figure 12 provides a breakdown of responses. Participants responded very consistently: 77.19% selected either "Agree" or "Strongly Agree" for both questions. However, we found that 3 participants chose opposing options, e.g., agreeing in the first question and strongly disagreeing in the second. We believe that this is probably caused by inattention. Despite this minor inconsistency, most participants believe that Duo enhances the security of their accounts.



(a) Duo Push Notification Task



(b) Duo Code Input Task

Fig. 11: Participants' thoughts on Overheads of Duo

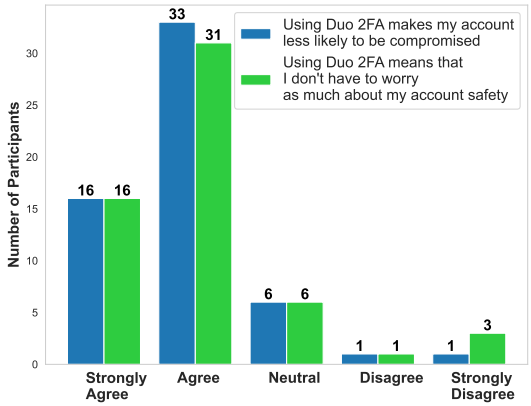


Fig. 12: Security Perceptions of Duo

J. Perception of Usability

We asked participants to rate four statements to assess their perception of Duo usability. Results are presented in Figures 13 and 14.

Most find Duo easy to use, yet annoying. Participants

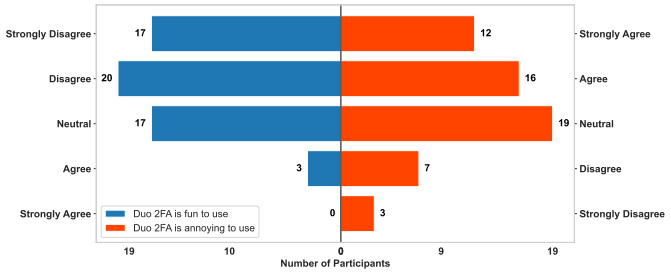


Fig. 13: Duo 2FA is fun to use, Duo 2FA is annoying to use

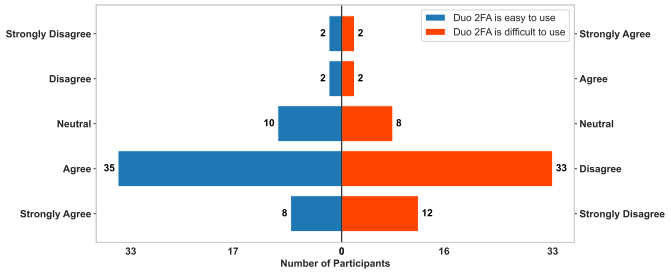


Fig. 14: Duo 2FA is easy to use, Duo 2FA is difficult to use

responded to these statements consistently, with the number of those who responded positively to Duo being "easy" closely matching the number of those who responded negatively to Duo being "difficult". Responses regarding Duo being "fun" versus "annoying" reflected a similar consistency. We also checked each participant's responses to the "easy vs. difficult" and "fun vs. annoying" statements for inconsistencies and found that all responses were consistent. This increases our confidence in the quality of the participants' responses, suggesting that they did not just select options randomly.

Usability Perception vs Experienced Push Overhead: We investigated whether Duo push overhead influenced participants' responses to these four statements. For example, we examined whether participants who selected "Agree/Strongly Agree" to the statement "Duo is annoying to use" experienced higher average Duo push overhead compared to those who selected "Disagree/Strongly Disagree". The results are presented in Table VI.

TABLE VI: Average Push Overhead (seconds) by Responses to Duo 2FA Usability Statements

Statement	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Duo 2FA is annoying to use	7.98	8.34	7.36	8.038	7.33
Duo 2FA is difficult to use	7.07	7.89	8.04	10.41	8.55
Duo 2FA is easy to use	8.55	6.01	8.08	7.85	7.17
Duo 2FA is fun to use	7.41	7.85	8.26	6.82	NA

Our analysis revealed no discernible relationship between average push overhead and participants' responses. Furthermore, we found no statistically significant differences in push overhead values when participants were grouped by their responses to the usability statements.

K. Other Issues with Duo

Finally, we included an optional open-ended question in the survey asking participants to describe any other issues or difficulties they had encountered with Duo. It is important to note that Duo is also used for second-factor authentication in the single sign-on system of the university where this study was conducted. Consequently, some issues reported by participants involved challenges with the central system. We received ten responses, and the following themes were identified from these responses:

1) **Ease of Use**: Mentions of Duo being easy to use.
2) **Frequency of Use**: Annoyance with having to use Duo multiple times a day.
3) **Single Sign-On (SSO) Issues**: Frustrations with the university's single sign-on not working as expected across systems.
4) **Manual Login Issues**: Problems with having to manually enter login information in apps.
5) **Device Change Issues**: Difficulties encountered when changing phones or phone numbers.
6) **Phone Dependency**: Inability to log in without the phone.
7) **Notification Issues**: Problems with receiving push notifications, including delays or failures.
8) **App Compatibility**: Mention of app availability on devices like the Galaxy Watch.
9) **Network Issues**: Problems related to network issues affecting Duo's functionality.
10) **Overall Satisfaction**: General satisfaction or lack of difficulties with using Duo.

Three responses mentioned notification issues, and two – highlighted device change issues (i.e., difficulties in changing a registered phone number or device). Other themes were reported by one response each.

L. SUS Score

We used SUS to quantitatively assess Duo usability. The SUS score computed from the survey is 70, indicating a "Good" usability level. This score aligns with participants' perceptions that Duo tasks are easy and quick to complete. However, the perception of Duo being annoying likely contributes to it not achieving an "Excellent" usability level, despite the tasks being straightforward and time-efficient. We also examined the potential correlation between participants' experienced overheads and their reported SUS scores. No significant relationship was observed between these two measures.

V. Comparison with Related Work

Our research focuses on two primary areas: (1) quantifying the overhead associated with Duo, and (2) examining the usability and user perception of Duo. Consequently, we compare our study with prior research within these two domains. Table VII presents an overview of the comparison.

TABLE VII: Comparison with Related Work

	#Participants	Year	Push Overhead (seconds)	Failure Rate (log, survey)	Push SUS Score	User Perception
this	**2559**	**2024-2025**	**7.82**	**4.35%, 43.86%**	**70**	**easy, annoying, increases security**
[12]	>77K	2018	NA	$> 5\%$, NA	NA	NA
[14]	27	2017	58	5%, NA	85	NA
[15]	72	2018	11.8	NA	81	increases security
[20]	2047	2017	NA	NA	NA	easy, annoying, increases security
[21]	4275	2018	NA	NA, 50%	NA	easy, increases security
[22]	565	2016-2018	NA	10%, NA	NA	NA
[23]	465	2021	NA	NA	NA	increases stress, anxiety, frustration
[24]	NA, 12500 user reviews	2019	NA	NA	NA	setup difficulty, backup and migration issue, forced to use, demand for device compatibility
[25]	42	2021	NA	NA	NA	requirements: excellent usability, location-independence, functional without internet, independence from energy source, secure interaction

A. MFA/2FA Overhead

Prior studies reported 2FA overheads ranging from 11 seconds to 58 seconds [14], [15], [26], notably higher than our calculated overhead of 7.82 seconds. Acemyan et al. [14] reported a 58 seconds average for Gmail login using Google Prompt and SMS, though this included the entire login sequence. Reese et al. [15] utilized a simulated banking environment to find a median push overhead of 11.8 seconds. Recent work by Kruzikova et al. [26] examined hardware-based 2FA (NFC and payment cards), finding overheads between 11 seconds and 35 seconds; they concluded that higher "time satisfaction" directly correlates with a higher perceived security.

Reynolds et al. [12] used university log data to mathematically estimate annual time-costs rather than direct measurement, yet their findings align with ours regarding technical friction, noting a failure rate exceeding 5%. Our significantly lower overhead of 7.82 seconds likely reflects a combination of improved UI designs and a more technologically adept user base compared to studies conducted 5–7 years ago.

B. MFA/2FA Usability and Perception

While prior studies of push-based 2FA report "excellent" usability with SUS scores between 81 and 85 [14], [15], our Duo-specific SUS score of 70 indicates a "good" but lower usability level. This discrepancy may be explained by the mandatory nature of Duo in university settings. Colnago et al. [20] and Dutson et al. [21] found that while users find Duo "easy", they simultaneously find it "annoying". This annoyance is exacerbated during mandatory rollouts; Abbott et al. [22] observed that user acceptance degrades significantly when 2FA moves from an optional or selective phase to a universal requirement.

Social and emotional factors also play a role. Weidman et al. [13] highlighted the tension of using personal devices for work-related 2FA, which many users view as an infringement on personal life. Arnold et al. [23] found that 2FA can induce stress and anxiety, particularly when technical failures hinder time-sensitive tasks. Large-scale sentiment analysis by Das et al. [24] supports this, revealing that Duo holds a middle-tier sentiment score (0.349) compared to other authenticators, with users frequently citing backup and device compatibility issues.

Finally, our findings on connectivity concerns mirror Marky et al. [25], who identified location-independence and offline functionality as primary user requirements. Other research continues to explore the lack of uniform design patterns across 2FA journeys [27] and the potential for lifecycle management tools to improve the setup and removal process [28], [29], [30], [31], [32], [33].

VI. Discussion

A. Limitations

While this study draws on extensive real world Duo usage and combines fine grained log analysis with user survey data, the following limitations should be taken into account when interpreting our findings.

Demographic Bias. The sample is strongly skewed toward technically proficient users: over 98% were students (mostly majoring in Computer Science or related disciplines) who use Duo frequently. They may be more able and willing to work around Duo friction than older adults, non-technical types, or those with lower digital literacy.

Survey Response Bias. The data sources are imbalanced: logs cover $2,500+$ users, but only 57 completed the survey which corresponds to 2.2% response rate. This likely introduces volunteer bias (e.g., stronger opinions, higher security awareness, or more extreme experiences), so perceived overhead, annoyance, security benefit, and SUS should be treated as indicative rather than representative.

Limited Insight into Overhead Sources. Retrospective logs measure only aggregate push overhead, not its components. The 7.82-second average delay combines user behavior,

device performance, and network/server latency. Logs lack device/OS information, network conditions, and whether approval happened via notification or by opening the app. Thus, we cannot attribute delay to inattention, notification issues, UI design, or the infrastructure.

Lack of Accessibility Evaluation. Accessibility and inclusivity were not assessed. We did not recruit/identify users with visual, motor, or cognitive impairments or observe assistive technology use. So we cannot conclude how usable Duo is for people with disabilities.

B. Recommendations for Improving Duo MFA Usability

Duo MFA achieved a "Good" usability level (SUS=70) and low average authentication overhead (7.82 seconds). However, log data and user feedback indicate recurring sources of friction. The following suggested improvements are intended to reduce user burden while preserving Duo's security guarantees.

Reducing Prompt Frequency. Repeated daily MFA prompts were a common complaint. "Remember device" and risk-based policies can be enabled to suppress prompts for trusted devices and low-risk contexts (e.g., same browser, device, and network for 8–24 hours). Although this capability is supported by Duo, it was not enabled in the studied deployment. Appropriate "remember windows" should be configured based on the institutional risk tolerance (e.g., longer intervals for staff desktops and shorter ones for shared lab machines), and the conditions under which a new prompt is required should be clearly communicated to maintain user trust.

Improving Push Reliability and Timeout Recovery. Push notification timeouts constituted the largest category of Duo-related authentication failures. Push delivery robustness can be increased through retries across multiple channels and through clear on-screen status indicators (e.g., request sent, waiting for approval, push failed). When a push is likely to time out, fallback options (e.g., passcode entry or use of a secondary device) should be offered proactively rather than requiring a full restart of the login sequence.

Supporting Offline and Low-Connectivity Scenarios. Delays and lockouts were reported in low-connectivity contexts. Offline options should be made more prominent during setup and within the authentication interface, including guidance for generating passcodes in advance and storing them safely for travel or field work. When network problems are detected, targeted guidance (e.g., recommending passcode use instead of waiting for a push) should be presented to reduce avoidable failures.

Strengthening Fallbacks When the Phone Is Unavailable. Lack of access to the registered phone was the dominant user-reported reason for being unable to complete Duo tasks, aligning with failure modes observed in the logs. Enrollment of at least one additional factor (e.g., a backup device, hardware token, or landline) should be encouraged, and periodic reminders to verify backup validity should be incorporated. Self-service recovery flows based on pre-registered secondary factors should be provided to reduce lockouts and administrative burden while maintaining security.

Lowering Interaction Cost at Approval Time. Many participants reported friction when phone unlock and app-opening were required for approval, particularly when in-notification "Approve/Deny" buttons did not work. Reliable in-notification approval actions would reduce interaction cost and better align with user expectations. Biometric confirmation on mobile or desktop platforms (e.g., Touch ID or Windows Hello) can also be integrated to streamline approvals; these features are configurable but were not utilized in the studied deployment.

Clarifying Passcode-Based Authentication Flows. Multiple passcode entries within a single session were reported by some participants. Interfaces should distinguish clearly between expired and invalid codes, and passcode prompts should be limited to explicit fallback scenarios rather than repeated without explanation. Contextual feedback should be provided to indicate next steps (e.g., generating a new code or carefully re-entering the current code), and recovery behavior should be streamlined when a code fails.

These recommendations translate empirical failure patterns and user feedback into design and deployment changes that can improve everyday usability without weakening Duo's security guarantees.

VII. Conclusion

This paper reports on a comprehensive usability analysis of Duo 2FA, focusing specifically on the overhead of Duo push notification tasks, the factors influencing this overhead, authentication failures associated with Duo, the SUS score, and user perceptions of the platform. Our research used authentication log data from UCI, involving $2,559$ unique users over a nine-month period, from August 2024 to April 2025. Also, we administered a survey to 57 randomly selected participants to gather qualitative insights. Our findings indicate that the average overhead of a Duo push notification task is 7.82 seconds, influenced by factors such as the time of day, participants' academic major, and educational levels. The system achieved a SUS score of 70, reflecting a "Good" level of usability. Notably, 43.86% of participants reported experiencing at least one authentication failure due to an incomplete Duo task. While users acknowledged that Duo tasks were generally easy to complete, they also described the tasks as annoying. Nevertheless, Duo was found to enhance users' sense of security regarding their accounts. Participants also identified common issues, such as needing to open the Duo app when a push notification fails and experiencing difficulties completing Duo tasks due to occasional lack of Internet access. They several improvements, including a "Remember Device" feature.

Acknowledgements: We thank USEC'26 reviewers for their constructive feedback. This work was supported in part by funding from the NSF Award SATC-1956393.

REFERENCES

[1] Fadi Aloul, Syed Zahidi, and Wassim El-Hajj. Two Factor Authentication Using Mobile Phones. In *2009 IEEE/ACS International Conference on Computer Systems and Applications*, pages 641–644. IEEE, 2009.

[2] Anupam Das, Joseph Bonneau, Matthew Caesar, Nikita Borisov, and XiaoFeng Wang. The Tangled Web of Password Reuse. In *NDSS*, volume 14, pages 23–26, 2014.

[3] 2025 Multi-Factor Authentication (MFA) Statistics & Trends to Know. https://jumpcloud.com/blog/multi-factor-authentication-statistics.

[4] The Secure Sign-In Trends Report 2023. https://www.okta.com/reports/the-secure-sign-in-trends-report/.

[5] The Best Two-Factor Authentication App. https://www.nytimes.com/wirecutter/reviews/best-two-factor-authentication-app/.

[6] Top Companies Using Duo Security. https://rocketreach.co/cl/companies-using-duo-security$_0$03.

[7] UI Verify: Two-Factor Authentication at the University of Illinois System. https://verify.uillinois.edu/.

[8] Multi-Factor Authentication at the University of Florida. https://it.ufl.edu/get-connected/multi-factor-authentication/.

[9] CalNet 2-Step Verification at UC Berkeley. https://calnet.berkeley.edu/calnet-2-step/2-step-devices/enroll-smartphone.

[10] Two-Factor Authentication (Duo) at the University of Arizona. https://it.arizona.edu/node/705.

[11] Multifactor Authentication (Duo) at Columbia University. https://cuit.columbia.edu/mfa.

[12] Joshua Reynolds, Nikita Samarin, Joseph Barnes, Taylor Judd, Joshua Mason, Michael Bailey, and Serge Egelman. Empirical Measurement of Systemic 2FA Usability. In *29th USENIX Security Symposium (USENIX Security 20)*, pages 127–143, 2020.

[13] Jake Weidman and Jens Grossklags. 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*, pages 212–224, 2017.

[14] Claudia Ziegler Acemyan, Philip Kortum, Jeffrey Xiong, and Dan S Wallach. 2FA Might Be Secure, But It's Not Usable: A Summative Usability Assessment of Google's Two-Factor Authentication (2FA) Methods. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 62, pages 1141–1145. SAGE Publications Sage CA: Los Angeles, CA, 2018.

[15] Ken Reese, Trevor Smith, Jonathan Dutson, Jonathan Armknecht, Jacob Cameron, and Kent Seamons. A Usability Study of Five Two-Factor Authentication Methods. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*, pages 357–370, 2019.

[16] Anil K Jain, Arun Ross, and Salil Prabhakar. An Introduction to Biometric Recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):4–20, 2004.

[17] 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 *2012 IEEE Symposium on Security and Privacy*, pages 553–567. IEEE, 2012.

[18] John Brooke et al. SUS—A Quick and Dirty Usability Scale. *Usability Evaluation in Industry*, 189(194):4–7, 1996.

[19] Aaron Bangor, Philip Kortum, and James Miller. Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *Journal of Usability Studies*, 4(3):114–123, 2009.

[20] Jessica Colnago, Summer Devlin, Maggie Oates, Chelse Swoopes, Lujo Bauer, Lorrie Cranor, and Nicolas Christin. "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*, pages 1–11, 2018.

[21] Jonathan Dutson, Danny Allen, Dennis Eggett, and Kent Seamons. 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)*, pages 119–128. IEEE, 2019.

[22] Jacob Abbott and Sameer Patil. How Mandatory Second Factor Affects the Authentication User Experience. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2020.

[23] Davis Arnold, Benjamin Blackmon, Brendan Gibson, Anthony G Moncivais, Garrett B Powell, Megan Skeen, Michael Kelland Thorson, and Nathan B Wade. The Emotional Impact of Multi-Factor Authentication for University Students. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, pages 1–4, 2022.

[24] Sanchari Das, Bingxing Wang, and L Jean Camp. MFA Is a Waste of Time! Understanding Negative Connotation Towards MFA Applications via User Generated Content. *arXiv Preprint arXiv:1908.05902*, 2019.

[25] Karola Marky, Kirill Ragozin, George Chernyshov, Andrii Matviienko, Martin Schmitz, Max Mühlhäuser, Chloe Eghtebas, and Kai Kunze. "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.

[26] Agata Kruzikova, Michal Muzik, Lenka Knapova, Lenka Dedkova, David Smahel, and Vashek Matyas. Two-Factor Authentication Time: How Time-Efficiency and Time-Satisfaction Are Associated With Perceived Security and Satisfaction. *Computers & Security*, 138:103667, 2024.

[27] Sanam Ghorbani Lyastani, Sven Bugiel, and Michael Backes. A Systematic Study of the Consistency of Two-Factor Authentication User Journeys on Top-Ranked Websites. 2023.

[28] Garrett Smith, Tarun Yadav, Jonathan Dutson, Scott Ruoti, and Kent Seamons. "If I Could Do This, I Feel Anyone Could:" The Design and Evaluation of a Secondary Authentication Factor Manager. In *32nd USENIX Security Symposium (USENIX Security 23)*, pages 499–515, 2023.

[29] Sanchari Das, Bingxing Wang, Zachary Tingle, and L Jean Camp. Evaluating User Perception of Multi-Factor Authentication: A Systematic Review. *arXiv Preprint arXiv:1908.05901*, 2019.

[30] Sanchari Das, Bingxing Wang, Andrew Kim, and L Jean Camp. MFA Is a Necessary Chore!: Exploring User Mental Models of Multi-Factor Authentication Technologies. 2020.

[31] Ahmed Anu Wahab, Daqing Hou, and Stephanie Schuckers. A User Study of Keystroke Dynamics as Second Factor in Web MFA. In *Proceedings of the Thirteenth ACM Conference on Data and Application Security and Privacy*, pages 61–72, 2023.

[32] Kevin Jensen, Faiza Tazi, and Sanchari Das. Multi-Factor Authentication Application Assessment: Risk Assessment of Expert-Recommended MFA Mobile Applications. *Proceeding of the Who Are You*, 2021.

[33] Sabrina Klivan, Sandra Höltervennhoff, Nicolas Huaman, Alexander Krause, Lucy Simko, Yasemin Acar, and Sascha Fahl. "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*, pages 3138–3152, 2023.

APPENDIX

SURVEY QUESTIONNAIRE

- 2FA: Two-Factor Authentication
- MFA: Multi-Factor Authentication

Demographics

1) What is your age?
2) What is the highest level of education you attained?
   - Below High School
   - High School
   - Associate Degree
   - Bachelor Degree
   - Master Degree
   - Doctorate (PhD)
3) What is your gender?
   - Male
   - Female
   - Non-binary
   - Transgender
   - Prefer not to say

General Frequency and Opinion

1) How frequently do you interact with Duo to access UCI services?
   - a few times per month or less
   - a few times per week
   - 1 to 2 times daily
   - 3 to 5 times daily
2) What do you think of Duo 2FA system?

Detailed Experiences

1) What is your opinion about the time required to complete Duo 2FA tasks?
   - It does not take much time
   - It takes a moderate amount of time
   - It takes a long time
2) Have you used any other 2FA or MFA system before encountering Duo 2FA at UCI?
   - Yes
   - No
3) Have you been motivated by using Duo at UCI to adopt MFA/2FA for any of your non-UCI accounts?
   - Yes
   - No
4) What is your opinion about completing Duo tasks as part of UCI access?
   - I am fine with it
   - I am okay with it but find it somewhat inconvenient
   - I do not like it but understand that it's needed for security purposes
   - I do not like it and would prefer it to be removed

Specific Incidents

1) Were you ever unable to login because it wasn't possible to complete Duo 2FA tasks?
   - Yes
   - No
2) If yes, which of the following happened?
   - I did not have physical access to the phone registered with Duo
   - Duo task did not appear on my phone
   - Duo task timed out before I could complete them
   - Other, please specify

Opinion and Experience Details

Please rate the following statements about Duo 2FA:

1) Using Duo 2FA makes my account less likely to be compromised.
   - Strongly Disagree
   - Disagree
   - Neutral
   - Agree
   - Strongly Agree
2) Using Duo 2FA means that I don't have to worry as much about my account safety.
   - Strongly Disagree
   - Disagree
   - Neutral
   - Agree
   - Strongly Agree
3) Duo 2FA is fun to use.
   - Strongly Disagree
   - Disagree
   - Neutral
   - Agree
   - Strongly Agree
4) Duo 2FA is easy to use.
   - Strongly Disagree
   - Disagree
   - Neutral
   - Agree
   - Strongly Agree
5) I think that Duo 2FA is difficult to use.
   - Strongly Disagree
   - Disagree
   - Neutral
   - Agree
   - Strongly Agree
6) I think that Duo 2FA is annoying to use.
   - Strongly Disagree
   - Disagree
   - Neutral
   - Agree
   - Strongly Agree

System Usability Scale

In regard to Duo 2FA, denoted as 'system' below, please score the following 10 items:

1) I think that I would like to use this system frequently:
   a) Strongly Disagree
   b) Disagree
   c) Neutral
   d) Agree
   e) Strongly Agree
2) I found the system unnecessarily complex:
   a) Strongly Disagree
   b) Disagree
   c) Neutral
   d) Agree
   e) Strongly Agree
3) I thought the system was easy to use:
   a) Strongly Disagree
   b) Disagree
   c) Neutral
   d) Agree
   e) Strongly Agree
4) I think that I would need the support of a technical person to be able to use this system:
   a) Strongly Disagree
   b) Disagree
   c) Neutral
   d) Agree
   e) Strongly Agree

5) I found that the various functions in this system were well integrated:
   a) Strongly Disagree
   b) Disagree
   c) Neutral
   d) Agree
   e) Strongly Agree
6) I thought there was too much inconsistency in this system:
   a) Strongly Disagree
   b) Disagree
   c) Neutral
   d) Agree
   e) Strongly Agree
7) I would imagine that most people would learn to use this system very quickly:
   a) Strongly Disagree
   b) Disagree
   c) Neutral
   d) Agree
   e) Strongly Agree
8) I found the system very cumbersome to use:
   a) Strongly Disagree
   b) Disagree
   c) Neutral
   d) Agree
   e) Strongly Agree
9) I felt very confident using the system:
   a) Strongly Disagree
   b) Disagree
   c) Neutral
   d) Agree
   e) Strongly Agree
10) I needed to learn a lot of things before I could get going with this system:
   a) Strongly Disagree
   b) Disagree
   c) Neutral
   d) Agree
   e) Strongly Agree