2018 IEEE Symposium on Security and Privacy

A Tale of Two Studies: The Best and Worst of YubiKey Usability

Joshua Reynolds[†*], Trevor Smith[*], Ken Reese[*], Luke Dickinson[*], Scott Ruoti[‡], Kent Seamons[*]
[†]University of Illinois at Urbana-Champaign, [*]Brigham Young University, [‡]MIT Lincoln Laboratory
joshuar3@illinois.edu, {tsmith, ken.reese, luke}@isrl.byu.edu, scott@ruoti.org, seamons@cs.byu.edu

***Abstract*—Two-factor authentication (2FA) significantly improves the security of password-based authentication. Recently, there has been increased interest in Universal 2nd Factor (U2F) security keys—small hardware devices that require users to press a button on the security key to authenticate. To examine the usability of security keys in non-enterprise usage, we conducted two user studies of the YubiKey, a popular line of U2F security keys. The first study tasked 31 participants with configuring a Windows, Google, and Facebook account to authenticate using a YubiKey. This study revealed problems with setup instructions and workflow including users locking themselves out of their operating system or thinking they had successfully enabled 2FA when they had not. In contrast, the second study had 25 participants use a YubiKey in their daily lives over a period of four weeks, revealing that participants generally enjoyed the experience. Conducting both a laboratory and longitudinal study yielded insights into the usability of security keys that would not have been evident from either study in isolation. Based on our analysis, we recommend standardizing the setup process, enabling verification of success, allowing shared accounts, integrating with operating systems, and preventing lockouts.**

***Index Terms*—two-factor authentication, hardware tokens, usability, longitudinal study, YubiKey**

I. Introduction

Passwords are the most widespread form of user authentication on the web today [1]. Although there are numerous proposals to replace passwords, none have matched their deployability nor their usability [2]. Still, passwords come with significant security and usability problems. For example, users struggle to create and remember strong passwords [3], [4], developers fail to properly secure stored passwords [5], [6], and password phishing is regularly successful [7].

To address the limitations of password-based authentication, many have advocated for a switch to two-factor authentication (2FA) [2], [8]. 2FA improves upon password-based authentication by requiring that in addition to proving knowledge of their password, a user must also prove possession of a trusted hardware device (the second factor). Even if an attacker steals a user's password, they will be unable to use it by itself to impersonate the user.

At the forefront of the push for 2FA is the FIDO Alliance and their Universal 2nd Factor protocol (U2F) [9]. The U2F protocol—which already has broad support and is available in systems used by 1.5 billion people [10]—is currently implemented by *security keys*. A security key is a hardware device that authenticates the user after the user presses a button on the security key [8]. The button tap is a test of user presence and prevents malware on the host device from using the security key surreptitiously. Most commonly, security keys are designed to be plugged into a USB port, though they can also communicate with other devices using wireless protocols (e.g., NFC, Bluetooth).

U2F security keys are designed to be easy-to-adopt and use in day-to-day life, while protecting users against phishing and man-in-the-middle attacks [8]. Lang et al. [8] previously demonstrated that security keys are highly effective in enterprise environments. In this paper, we conduct the first studies exploring whether security keys are sufficiently usable for day-to-day usage by *non-enterprise* users. While there are many different security keys (e.g., NitroKey, OnlyKey, U2F Zero, AdaFruit), we used three of Yubico's security keys for our studies (YubiKey 4, YubiKey NEO, and YubiKey Nano). These devices most closely resemble the security keys used in Lang et al.'s study, and Yubico's security keys have received a significant amount of media and industry attention. Still, our results are largely applicable to U2F, and to a lesser extent 2FA.

We first conducted a laboratory study with 31 participants, investigating whether unassisted novice users could configure Google, Facebook, and Windows 10 accounts to authenticate using a YubiKey. This study revealed that there are significant impediments for novice users when setting up their accounts with 2FA in general and YubiKeys in particular. In general, the participants viewed YubiKeys as largely unusable.

We then conducted a longitudinal study with 25 participants, exploring the usability of YubiKeys for day-to-day authentication to the participants personal Google, Facebook, and Windows 10 accounts over a period of four weeks. Importantly, study coordinators helped participants set up the YubiKeys in order to avoid having the poor setup process bias participants' views of the daily usability of a YubiKey. In contrast to the first study, users were very positive regarding YubiKeys and 2FA in general.

The contributions of our work are as follows:

(1) **First user studies of security keys for unassisted, non-enterprise users.** Our work shows that the non-enterprise setup experience was significantly worse than the experience described by Lang et al. [8]. Still, day-to-day usability seems to be comparable in the

DOI 10.1109/SP.2018.00067

enterprise and non-enterprise settings. Taken together, these results demonstrate that security keys could be a viable 2FA option for non-enterprise users if the U2F setup experience can be sufficiently improved.

(2) **First to separately study 2FA setup and daily-use.** Prior work has either explored only one of the two phases or has examined both phases in a single study. Our work revealed that each phase has unique usability challenges and both phases must be studied in order to fully understand the usability of 2FA systems.

(3) **Recommendations for usability improvements.** Both the laboratory and longitudinal studies revealed several areas where users either struggled to set up and use a YubiKey, or where the general user experience could be improved. Based on our observations and participant feedback, we offer several concrete suggestions for improving the usability of security keys and 2FA in general. For example, many of our participants were hindered by outdated and overly dense documentation. Instead, we recommend that systems provide more active forms of documentation, such as the 2FA wizard approach that Google provides.

II. Related Work

Prior work on the usability of two-factor authentication includes laboratory studies, longitudinal studies, and surveys.

A. Laboratory studies

Piazzalunga et al. [11] compared the usability of smart cards and two different kinds of USB security tokens. Ten participants were tasked to install drivers, install software, and send an encrypted email using them. The basic USB security token only included the cryptographic material, while the advanced USB security token also included digital instructions, drivers, and relevant software on the device. Both USB security tokens were faster to use, required less simulated customer service requests, and resulted in fewer errors than traditional smart cards.

Weir et al. [12] conducted a within-subject comparison of three hardware code generators under evaluation by a bank in the UK. The first system generated a code with the push of a button, the second required inserting the user's bank card into the code-generating device, and the third required inserting the user's bank card and entering their PIN using a scroll wheel mechanism. Two-thirds of the participants preferred using the push-based system, despite considering the other two systems to be more secure. In a similar study, Weir et al. [13] conducted an in-lab study of three authentication systems including SMS-based 2FA and the push-button hardware code generator used in their previous study. In this study, participants preferred the SMS-based system. While these studies included hardware-based 2FA for non-enterprise users, they did not include security keys.

Gunson et al. [14] conducted a laboratory study of 2FA in the context of automated telephone banking. Although users felt more secure while using 2FA, they reported that the system had low usability. Participants in this study were asked to authenticate to a simulated telephony banking system using a hardware code generator. Users disliked having to carry a physical token with them and were unsure of how the token improved the security of traditional knowledge-based questions.

Trewin et al. [15] conducted a user study to compare various forms of smartphone-based biometric authentication methods including voice, facial recognition, gesture, and combinations of these methods. In general, participants responded negatively to voice-based authentication and combinations of biometrics. On the other hand, participants were positive towards facial recognition and gesture-based authentication.

Karapanos et al. [16] proposed a new form of reduced interaction 2FA, called Sound-Proof, which authenticates users based on proximity to a mobile device. The system authenticates a user by matching the ambient sound recorded near a computer and a mobile phone. User study results of this system found Sound-Proof to be more usable than the Google Authenticator app.

B. Longitudinal studies

Lang et al. [8] describe the design and implementation of hardware security keys that are the precursor to U2F security keys and are nearly identical to the devices used in our study. Using the attribute-based evaluation scheme described by Bonneau et al. [2], they argue that security keys ought to have usability similar to that of passwords while offering additional security benefits. Additionally, they conduct several internal user tests with Google employees, finding that the total amount of time spent authenticating decreased markedly compared to their current system based on one-time passcodes (OTP). Furthermore, they report a reduced number of support incidents over the long-term after deploying the security keys. Our study differs from the Lang et al. study because we were primarily concerned with how users without an IT support department would view security keys. We also report qualitative data from participants' experiences and calculate a System Usability Scale (SUS) score that can be used to compare against other authentication techniques [17].

Krol et al. [18] conducted a longitudinal study with 21 individuals already using 2FA as part of the authentication process for several UK banks. Participants used a variety of 2FA systems, including card readers, hardware code generators, SMS, phone calls, and smartphone authenticator applications. The participants began an 11-day study with a preliminary interview and ended the study with a follow-up interview. As a part of the study, participants were asked to create an entry in an authentication diary each time they needed to authenticate. These entries included the time, the reason, and any problems they encountered when they needed to authenticate. Many participants had a strong dislike for specialized hardware code generators required for 2FA, and in some cases reported they accessed their information less frequently or even changed banks because of the increased effort of using the devices. Additionally, several participants mentioned getting confused because of the inconsistent terminology used by online banks.

C. Surveys

De Cristofaro et al. [19] conducted a Mechanical Turk survey of online users already using 2FA. They analyzed their findings by grouping users into overlapping groups who use hardware tokens, SMS/email, and authentication apps. SMS was the most common for personal accounts; hardware code generators were the most common for business. The survey included a standard System Usability Scale (SUS) questionnaire [20], and in contrast to many previous studies of 2FA, participants gave a high rating to all the systems in the study. Even though USB tokens were included in the questionnaire, all hardware tokens are grouped together for the data analysis and reporting. This means the results are not fine-grained enough to be conclusive regarding the usability of USB tokens, which differ from smartcards and commercial code-generating tokens in the same category.

Strouble et al. [21] administered a 40-question survey inquiring about the usability and productivity impact of using a Common Access Card (CAC) as a 2FA method to military and civilian members of the United States Air Force. The survey showed that those who were required to use a CAC and CAC reader to read email remotely gave lower usability scores than those who were not required to do so. Also, more than half of the participants stated that they had left their CAC plugged into their work computer at least once. Leaving the CAC at work is especially problematic because a CAC also doubles as an official military ID that is required to gain entry to a military base. They estimate a loss of 261 work-years per year in productivity because of forgotten CACs.

D. Methodology

Egelman et al. [22] interviewed 28 smartphone users to investigate phone locking practices and motivations. They inductively coded these interviews to identify users' rationale. Utilizing these results and those of two following studies, they found that while users appeared to be making rational choices, those choices were founded on faulty assumptions. They then gave suggestions to better communicate the risks involved with not locking a phone. Our study employed a similar methodology by using more than one study, analyzing qualitative results through inductive coding, and offering relevant recommendations.

III. System Overview

U2F security keys are hardware tokens that are used to authenticate users after they press a button on the security key.[1] Security keys come in a variety of form factors with most plugging into a device's USB port, though the U2F protocol also supports wireless communication with the host device (e.g., NFC, Bluetooth).

Security keys improve on the security of alternative authentication methods in the following ways:

(1) **Password-based authentication.** Security keys are a form of 2FA, preventing an attacker who steals a user's passwords from gaining access to that user's accounts.

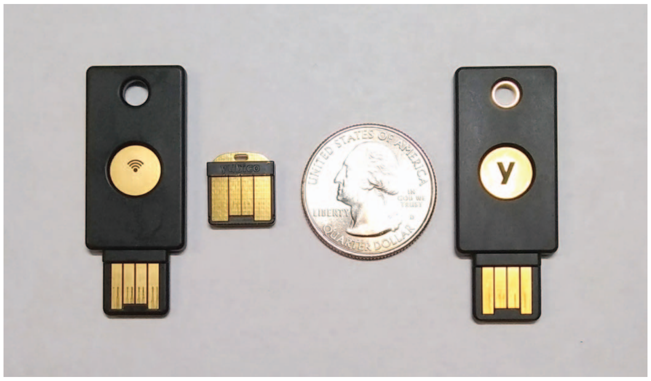


Fig. 1. From left to right: a YubiKey NEO, a YubiKey Nano, a US quarter dollar coin, and a YubiKey 4

(2) **One-time passcodes (OTP).** OTP provides two-factor authentication, but is susceptible to man-in-the-middle attacks that intercept and replay a user's passcodes [23]. Security keys use a challenge-response scheme that is resistant to these types of man-in-the-middle attacks [8].

(3) **Smart cards.** Smart cards provide many of the benefits of security keys, but are vulnerable to usage by malware as long as they are plugged into the system. In contrast, security keys require that the user acknowledges each authentication attempt, protecting the security key from surreptitious usage and helping the user detect possible attacks.

An increasing number of services currently support security keys, and our studies focus on two of the largest services—Google and Facebook. We also wanted to test 2FA on a major desktop platform, but neither Windows 10 or MacOS has native support for U2F security keys. Still, Yubico does provide software that allows the user to use a YubiKey to authenticate to the OS, and we chose to test the Windows 10 versions of this software. In the remainder of this section, we describe how a YubiKey worked with each of these services at the time of our studies.

In our studies of security keys, we had participants use three different devices: a YubiKey 4, a YubiKey NEO, and a YubiKey Nano (see Figure 1). Each device has a capacitive sensor that users tap to authenticate with the device—the gold disc in the center of the YubiKey 4 and YubiKey NEO, and the gold nub on the end of the YubiKey Nano. The YubiKey 4 and YubiKey NEO are largely identical, except that the NEO also supports NFC communication with devices, though we didn't explore this capability in our studies. The YubiKey Nano is designed to sit within the host device, with only the capacitive nub extending outside the USB port.

A. Google

Google provides a guided wizard to assist users in enabling 2FA for their accounts and setting up their security keys. This wizard is launched from the "Sign In and Security" page found in the user's account settings. Before setting up a security

[1] This button is not required to be a biometric sensor.

key, the wizard requires that users first set up 2FA using a phone number; this is used as an account recovery option if the security key is unavailable. After registering their phone number, the wizard presents users with the option to add a variety of 2FA devices, including a button to "Add a security key." After clicking this button, the wizard then helps users register their security key with Google. At the end of the process, the wizard informs users that their security key is now ready for use.

In the future, when the user authenticates to Google, they will first enter their username and password as usual. After credential verification, the user is then prompted to insert their security key and tap the button on the security key. By default, Google will then remember the device for several days and not require the YubiKey during authentication.

B. Facebook

Facebook provides a guided wizard to help users set up their security keys, but in contrast to Google's wizard it does not assist users in enabling 2FA on their account. To launch the security keys wizard, users would first navigate to the "Two-Factor Authentication" menu within the "Facebook Settings menu," and then click the link for adding a security key.

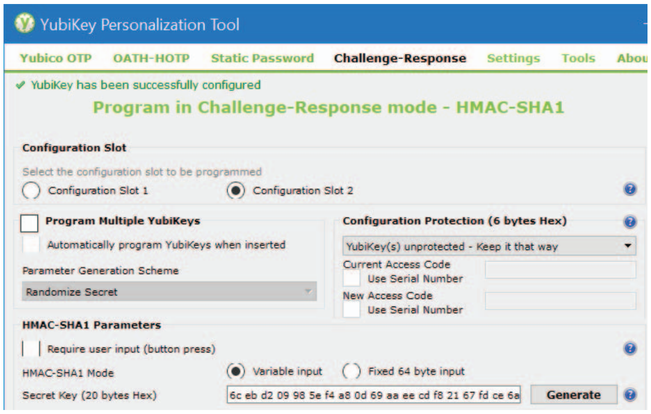
The lack of guidance regarding enabling 2FA led to two significant differences between Google and Facebook's setup process. First, users were not automatically prompted to set up a recovery 2FA scheme (SMS-based or the Facebook authenticator app); without enabling these recovery schemes, users cannot enable 2FA for their Facebook account. Second, users were not made aware that in addition to registering their security key, they would also need to separately enable 2FA on their accounts before Facebook would prompt them to use their security key.

After correctly setting up their security key and enabling 2FA, the authentication experience on Facebook was largely similar to that of Google.

C. Windows 10

Windows 10 does not natively support authentication using security keys, but Yubico has built two applications that add this functionality: the *Windows Logon Authorization Tool* and *YubiKey for Windows Hello*.

To configure their account with the *Windows Logon Authorization Tool* the user takes the following steps: First, the user needs to ensure that they have a local Windows account, not a Microsoft account[2] (the default account type); if the user has a Microsoft account, they need to first revert their account to a local Windows account. Second, the user needs to install the *YubiKey Personalization Tool* and use it to add an `HMAC-SHA1` challenge-response secret; this field is somewhat difficult to locate within the personalization tool (see Figure 2). Third, the user needs to ensure that they have installed the legacy .NET 3.5 framework, after which they then install the *Windows Logon Authorization Tool*. Fourth, and finally, the user registers a YubiKey with the *Windows Logon Authorization Tool* before restarting their computer. If the user accidentally restarts their computer between the third and fourth steps, it is possible that they will become locked out of their accounts since authentication now requires a YubiKey, but they have not yet registered one with the tool.



✓ YubiKey has been successfully configured

Program in Challenge-Response mode - HMAC-SHA1

Configuration Slot

Select the configuration slot to be programmed

○ Configuration Slot 1 ● Configuration Slot 2

□ Program Multiple YubiKeys

□ Automatically program YubiKeys when inserted

Parameter Generation Scheme: Randomize Secret

Configuration Protection (6 bytes Hex): YubiKey(s) unprotected - Keep it that way

Current Access Code □ Use Serial Number

New Access Code □ Use Serial Number

HMAC-SHA1 Parameters

□ Require user input (button press)

HMAC-SHA1 Mode ● Variable input ○ Fixed 64 byte input

Secret Key (20 bytes Hex): 6c eb d2 09 98 5e f4 a8 0d 69 aa ee cd f8 21 67 fd ce 6a

Fig. 2. Users were required to configure the YubiKey to use `HMAC-SHA1` with the *YubiKey Personalization Tool* before it could be used with the *Windows Logon Authorization Tool*

After the full setup, the user will first select their account and enter their password as usual. When these credentials are verified, if the user's YubiKey is already inserted, they will be immediately logged in without any further action on their part. If the YubiKey is not already inserted, the user is prompted to insert it. The prompt is incredibly vague, only displaying the words "Status:" under the password bar. Authentication succeeds once the user inserts the YubiKey. Importantly, Windows 10 YubiKey authentication is not U2F compliant, as the user does not need to press the button on the YubiKey, and more closely resembles traditional smart card 2FA.

The second tool, *YubiKey for Windows Hello*, requires far fewer steps to set up than the *Windows Logon Authorization Tool*. All the user must do is install the application, create a PIN for their Windows account, and register a YubiKey with the application. Following the set up of *YubiKey for Windows Hello*, users still enter their standard Windows credentials to log into Windows after their computer first starts up, but can subsequently insert the YubiKey (no tap required) or enter their PIN in place of entering their password to re-authenticate at the lock screen—for example, after the computer wakes from sleep. Unlike the *Windows Logon Authorization Tool*, *YubiKey for Windows Hello* is still single-factor authentication, replacing something the user knows (password or PIN) with something they have (YubiKey).

IV. Measuring Setup Usability—Methodology

There are two dimensions to the usability of security keys: setting up the security key and using the security key in daily life. In this section, we describe our laboratory study to evaluate

[2] https://account.microsoft.com/account/faq

the ability of non-expert users to set up a security key without any assistance from a study coordinator. In Section VI we explore the day-to-day usability of security keys. Both of these studies were approved by our institution's IRB and the data gathered from each is available at https://isrl.byu.edu/data/sp2018/.

A. Study Design

The laboratory study ran for two weeks—beginning June 15, 2017, and ending June 26, 2017. In total, 31 participants completed the study and each was compensated $15 USD. The study ran between 25 and 70 minutes.

At the beginning of each study, participants were provided with a YubiKey 4 in its original packaging. Participants were then directed to use the next 5 minutes to learn about the YubiKey by accessing the Internet on a laboratory desktop computer. This was intended to compensate for the fact that most participants had no prior experience with a YubiKey and that in real life, users would have read about YubiKeys before purchasing one and receiving it in the mail.

Participants were then given three tasks: set up the YubiKey to be used as part of the login process for Google, Facebook, and Windows 10. An enumeration of the six possible task orderings was created and shuffled. Each participant was sequentially assigned to one of the task orderings (i.e., P1–P6 each used a different ordering, P7 used the same ordering as P1, etc.) Approximately 15 minutes were allocated for each task. Participants completed these tasks using account credentials provided by the study coordinator; the provided Windows 10 account was a local Windows account not linked to a Microsoft account. To mimic the user's normal computing environment, all major browsers (Google Chrome, Opera, Mozilla Firefox, Microsoft Edge, and Internet Explorer) were available, even though only Google Chrome and Opera supported U2F at that time. Participants had open access to the Internet, but at no time did they receive assistance from the study coordinator in setting up the YubiKey.

During the study, coordinators took notes on any sources the participant used both during the orientation period and while attempting to complete each task. We also captured audio and screen recordings for each participant. Participants were instructed to inform the coordinator when they completed the task. Coordinators noted specifically whether the participant was successful at configuring the YubiKey for each system. If the task took overly long (20 minutes), participants were asked to abandon the current task and move onto the next task. Participants were also allowed to move on if they decided they could not complete the task. If the participant was not successful, then the coordinator also noted the reason for the failure (for instance, the participant ran out of time, followed incorrect documentation, etc.)

After the participants had completed all three tasks, they answered a three-part survey. The first part of the survey was the standard System Usability Scale (SUS) [20], [24] questionnaire, consisting of a ten-item Likert scale. We chose to calculate a single SUS score at the end of the three tasks rather than after each because we wanted to transcend the current interfaces of each service to understand the overall usability of YubiKeys. We also wanted to ensure sufficient time for the participant to try all the systems and to avoid survey fatigue.

The second portion of the study's survey contained six additional Likert items inquiring about attitudes toward YubiKeys. The survey concluded with a few free-response questions and some basic demographic questions. These questions were coded by two researchers to report the quantities we present in our results.

B. Recruitment

We recruited 32 participants using posters (available in the appendix) in dozens of locations across the Brigham Young University campus. We requested that participants have prior experience with Facebook, Gmail, and Windows 10 to help ensure that the usability results reflected the experience of setting up the YubiKey and did not result from a lack of experience with any of the three systems. Of the initial 32 participants, one arrived so late that they did not have sufficient time to properly attempt all three tasks. Because of this substantially differing treatment, this participant's data was ultimately excluded from our results, leaving a total of 31 participants, referred to hereafter as A1–A31.

C. Demographics

Participants in our study skewed male: male (22; 70%), female (9; 29%). Nearly all were young adults: 18–25 years (25; 80%), 26–35 years (4; 12%), 46–55 years (2; 6%). Two-thirds of the participants had completed some college but did not yet have a degree: some college with no degree (20; 64%), associate degree (2; 6%), bachelor's degree (5; 16%), post-graduate degree (4; 12%). Four in ten were currently using a 2FA solution: currently using 2FA (13; 41%), previously used 2FA (6; 19%), have never used 2FA (12; 38%). Most participants reported having an intermediate level of computer expertise: beginner (9; 29%), intermediate (18; 58%), advanced (4; 12%).

D. Limitations

Since the study was conducted in a laboratory, participants may not behave the same as they would in the real world. The study population was drawn from a college campus, so the results are not generalizable. Future research could replicate this study with different populations to gain additional insights into the usability of security keys. Participants were not using their personal accounts or machines, which may have affected the user's ability to properly set up the YubiKey. Still, this was necessary to prevent users from accidentally locking themselves out of their accounts.

V. Measuring Setup Usability—Results

Table I reports how many users were able to set up the YubiKey with each service (Google, Facebook, Windows 10). We begin by reporting what resources our participants consulted in their self-orientation. Then, based on study coordinator notes and audio/video recordings we describe the common problems

	N=31	%
Google		
Success	**26**	**83%**
Correctly identified completion	22	70%
Failure	**5**	**16%**
Facebook		
Success	**10**	**32%**
Correctly identified completion	6	19%
Failure	**21**	**67%**
Registered YubiKey without enabling 2FA	12	38%
Windows 10		
Success	**12**	**38%**
Set up the *Windows Logon Authorization Tool*	5	16%
Set up *YubiKey for Windows Hello*	7	22%
Failure	**19**	**61%**
Failed to set up the *Windows Logon Authorization Tool*	9	29%
Failed to set up *YubiKey for Windows Hello*	5	16%
Locked out of the computer	6	19%

TABLE I
LABORATORY STUDY SUCCESS RATES

encountered by participants. Finally, we report on the overall SUS score, the responses to the post-study survey, and other qualitative feedback from participants.

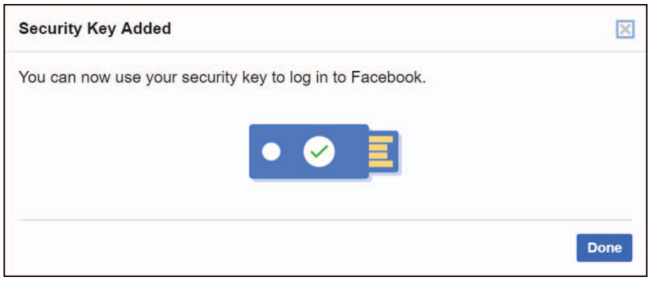
A. Self-Orientation

Even though we instructed the participants to learn about YubiKeys on their own without any specific direction from us, they were fairly consistent in where they went to learn about YubiKeys during the self-orientation. Participants rarely accessed the large printed URL on the shipping envelope, but rather used the Google or Bing search engines to locate YubiKey information. Most of them navigated to Wikipedia's entry on YubiKeys[3] and to various internal pages/content on Yubico's web site.[4] Several participants watched one of two videos[5,6] on the Yubico website. A couple of the participants found other resources and one participant installed the *YubiKey for Windows Hello* app.

B. Google

Most participants (26; 83%) successfully configured the YubiKey to work with the Google account. Four of the participants (4; 12%) reported being unsure whether they had finished setting up the YubiKey, while twenty-two participants (22; 70%) correctly finished and did not move ahead with uncertainty. Two of these participants logged out and back in several times after they had configured the YubiKey to test whether or not it was working; unfortunately, neither participant noticed that "don't ask for the security key again on this computer" was automatically selected on their first login, leaving them confused as to why subsequent logins didn't require the YubiKey. The third participant tried registering additional 2FA systems with their account, and the fourth tried restarting their computer. Despite their uncertainty, all participants had in fact set up the YubiKey correctly for Google.

Of the five participants (5; 16%) who failed to complete this task, four enabled the phone number-based 2FA but failed to notice the "Add a security key" link. Of these four participants, only one thought that they had succeeded after setting up phone number-based 2FA.


Security Key Added

You can now use your security key to log in to Facebook.

Fig. 3. The Facebook popup displayed to users after registering their security key. However, users are not able to use their security key at this point unless they had also enabled 2FA in their account settings.

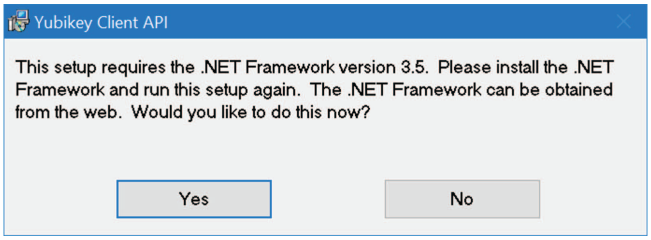
C. Facebook

Only a third of the participants (10; 32%) successfully configured the YubiKey to work with their Facebook account. Even successful participants had difficulty navigating Facebook's website to find the options to enable security keys. While there are instructions on Yubico's website for setting up a YubiKey with Facebook, they are out-of-date and only caused more confusion for users.

Over half of the participants that failed to complete this task (12; 38%) successfully registered the YubiKey with Facebook but did not properly configure 2FA on the account. There were two primary causes for this disconnect. First, Facebook requires that users set up phone number-based or code generator-based 2FA before allowing users to use a security key. Second, even if users properly registered 2FA for their account, it would be inactive until the participant also "enabled" 2FA in the account settings. Both of these problems were especially hard to detect as the Facebook interface told users that they were ready to use their security key when they were done registering it (see Figure 3), regardless of whether they had completed the other two steps to correctly set up 2FA for the account.

Participants also struggled to test whether the YubiKey was set up correctly. If users attempted to log out and back in several times to test the YubiKey, they would only be prompted to use it on the first attempt. After that, Facebook would store a cookie in the browser that indicates that security key login was no longer needed on the device. This is similar to Google's technique, but without the option to opt-out of having the device remembered. Ultimately, this prevented four participants (4; 12%) from being sure whether they had properly set up the

[3] https://en.wikipedia.org/wiki/YubiKey
[4] https://www.yubico.com
[5] https://player.vimeo.com/video/201088517
[6] https://player.vimeo.com/video/137100978



Yubikey Client API

This setup requires the .NET Framework version 3.5. Please install the .NET Framework and run this setup again. The .NET Framework can be obtained from the web. Would you like to do this now?

Fig. 4. .NET 3.5 was required to install the YubiKey software

YubiKey. As with Google, each uncertain user had configured the key correctly.

D. Windows 10

In total, twelve participants (12; 38%) successfully set up the YubiKey to work with their Windows accounts. Twenty-one participants (21; 67%) attempted to set up the YubiKey using the *Windows Logon Authorization Tool* and five succeeded, giving a success rate of 24% for this tool. In contrast, twelve participants (12; 38%) attempted to use *YubiKey for Windows Hello* and seven succeeded, giving a success rate of 58% for this tool.[7]

The differences in success rates can be attributed to the high complexity of setting up the *Windows Logon Authorization Tool* (see Section III). Especially problematic was the requirement that users set up .NET 3.5. When the *Windows Logon Authorization Tool* installer detected that .NET 3.5 was not installed, it would prompt users to install it (see Figure 4). Unfortunately, clicking on "Yes" would take users to a dead link, requiring the users to manually discover where to obtain the .NET 3.5 installer.[8]

Finally, six participants (6; 19%) locked themselves out of their computers while attempting to setup the *Windows Logon Authorization Tool* (more than successfully enabled YubiKey using that tool). This occurs because the *Windows Logon Authorization Tool* allows users to enable YubiKey authentication without requiring the user to register a YubiKey. This was especially likely to happen as after installation the *Windows Logon Authorization Tool* would immediately prompt users to enable YubiKey authentication: "YubiKey Logon is not enabled. Do you want to enable it? (Yes/No)." After selecting "Yes," users were then shown a dialog informing them "YubiKey Logon enabled, please reboot the computer for settings to take effect." Participants who immediately obeyed this message and rebooted the computer were locked out of their accounts, having failed to notice the GUI element that would have allowed them to register the YubiKey before rebooting. This experience was especially annoying to participants, with one participant exclaiming, **A6:** *"Now I can't get onto my own computer—and I'm out $50."*[9]

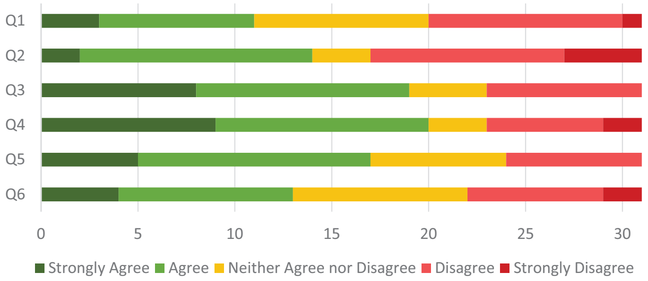


Fig. 5. Participant responses to statements about the YubiKey

While the study did not require participants to restore access to their Windows accounts, it is possible to do so through safe mode, as long as YubiKey authentication has not also been enabled for safe mode (not the default). If YubiKey authentication has been enabled for safe mode, then the user would have to either re-install Windows altogether or attempt to change a Windows registry key via a recovery terminal. This is especially problematic as installing the *Windows Logon Authorization Tool* requires that users disconnect their Microsoft account, disconnecting any cloud storage of settings that may have initially been stored there.

E. SUS

After completing all three tasks, participants evaluated the YubiKey 4 using the System Usability Scale (SUS). It received a mean score of 49.7 and a standard deviation of 16.8. Based on comparisons to other systems and contextual descriptions provided by Bangor et al. and Sauro et al. [25], [26], [27], [28] (see Figure 8 available in the appendix), this SUS score falls between the 0th and 15th percentiles, is considered "not acceptable," and receives a F grade.

F. Likert Items

Participants were also asked to rate the following six statements about the YubiKey using a 5-point Likert item:

Q1 I would like to keep using a YubiKey.
Q2 I would not always be able to find my YubiKey when I needed to log in.
Q3 I could easily keep YubiKey around me to use whenever I log into all my accounts.
Q4 Using the YubiKey got easier to use by the second or third time I set it up.
Q5 It would be worth the extra work to use the YubiKey to protect my personal accounts.
Q6 I would rather get a code texted to my phone than use a YubiKey.

Responses are summarized in Figure 5. Interestingly, half of the participants (17; 54%) indicated that they would be willing to accept YubiKey's poor usability if it really could better protect their accounts (Q5). This likely explains why even though YubiKey received such poor SUS scores a third of the participants (11; 35%) were willing to continue using it

[7] Three participants attempted to setup both the *Windows Logon Authorization Tool* and *YubiKey for Windows Hello* and one participant attempted neither.

[8] This problem has since been fixed.

[9] The price to purchase a YubiKey 4.

(Q1) and just over a third (13; 41%) preferred it to SMS-based 2FA (Q6).

Most participants (20; 64%) also felt that setting up the YubiKey got easier after the first setup experience (Q4). Additionally, most participants (19; 61%) felt they would have little difficulty carrying the YubiKey around with them (Q3), though almost half (14; 45%) noted that they might not always be able to find it when it was needed (Q2).

G. Open-Response Questions

To conclude the study, participants were asked three open-ended questions. We used these written responses to generate unique codebooks for each question through inductive analysis. Then, two researchers independently coded responses using the appropriate codebooks. We calculated Cohen's kappa for each of the nine codes to measure the inter-rater reliability. These kappa values had a mean of 0.848 and median of 0.832 with a range from 0.592 to 1.000. Differences were reconciled by the coders and the final results are described below.

(1) *What did you like about the experience of setting up the YubiKey?* About a quarter of the participants (7; 22%) were enthusiastic about the security of 2FA and security keys in particular:

> **A14:** *"I liked learning about a product I was unaware of. It seems like a good idea to implement into accounts that have extremely sensitive information. I would like to see them around more if they were easier to use."*

Four participants (4; 12%) also emphasized that while the initial setup was difficult, it was easy to use after that:

> **A11:** *"I liked how this could easily protect me and was very simple and easy to use after I got everything set up. I thought the design was very practical. I also could easily see a great advantage for a company to use this type of system."*
>
> **A20:** *"I liked that even though it took several steps to set up the YubiKey, the process of using the YubiKey afterwards was very simple."*

Additionally, eight participants (8; 25%) felt that the instructions and documentation they found were mostly good and easy-to-follow. Finally, two participants (2; 6%) indicated that there was nothing positive about their experience.

(2) *What would you improve about the experience of setting up the YubiKey?* Two-thirds of the participants (23; 74%) requested improved setup instructions and documentation. Of those, twelve (12; 38%) indicated that the instructions could be improved in clarity and brevity, with five (5; 16%) suggesting that the instructions could be improved through the addition of video tutorials. Finally, six participants (6; 19%) noted that the links and instructions they had found were out-of-date and that they should be updated.

(3) *How does a YubiKey make an account safer?* All but one participant (30; 96%) demonstrated a basic understanding that the YubiKey made their account safer:

> **A11:** *"It requires both a password and a physical key to get into an account. Someone could hack your basic password but without the key they still wouldn't be able to get in."*
>
> **A21:** *"A YubiKey is an extra layer of protection that cannot be accessed via hacking because it is an external device that generates a unique code each time you log in."*

VI. Measuring Day-to-Day Usability—Methodology

While the laboratory study identified clear problems with setting up a U2F security key, it was insufficient to explore the usability of security keys in day-to-day use. To better explore the daily usability of security keys, we conducted an IRB-approved longitudinal study tasking participants with using a YubiKey to authenticate to their personal Google, Facebook and Windows 10 accounts.

A. Study Design

The longitudinal study ran over the course of two months—beginning July 12, 2017, and ending September 6, 2017. Participation involved a four-week commitment to use a YubiKey with the participant's personal Google, Facebook, and Windows 10 accounts and ended with a post-study interview. In total, 25 participants completed the study and were each compensated $75 USD.[10]

Each study began with an in-person meeting between the study coordinator and the participant. During this meeting, the study coordinator explained what a YubiKey was and how they would be using it over the next four weeks. Participants were assigned to use either a YubiKey NEO or a YubiKey Nano (see Figure 1) to use for the duration of the study. The NEO is physically identical to the YubiKey 4 used in the first study and the Nano is thumbnail sized.[11] Eleven participants (11; 44%) used the YubiKey NEO and fourteen (14; 56%) used the YubiKey Nano. Participants used the same device for the entire study and every participant returned the YubiKey.

After selecting a YubiKey form factor, the study coordinator helped participants to configure it with their Google, Facebook, and Windows 10 accounts (using *YubiKey for Windows Hello*). In five cases, the study coordinator was unable to set up *YubiKey for Windows Hello* on the participant's machine. These participants completed the study using the YubiKey for only their Google and Facebook accounts.

After setting up the participants' YubiKey, the study coordinator walked participants through using the YubiKey to authenticate to each account and confirmed that participants understood how to use the YubiKey. We were not testing authentication on mobile devices and coordinators helped participants learn to authenticate via SMS on their smartphones. Participants were then given an authentication journal and

[10] Participants were free to withdraw from the study at any time and receive prorated compensation, though no participant choose to withdraw.

[11] The difference between the YubiKey 4 and the NEO is the image on the button and the fact that the NEO can also communicate via NFC.

instructed to write down notable experiences they had while using the YubiKey [18]. Unlike Hayashi and Hong's [29] password study, we did not request that participants record every login attempt in their journal. At this point, participants were free to leave and begin using the YubiKey.

At the completion of the four-week study, participants returned for a semi-structured post-study exit interview. In the interview, participants were asked to share their experiences using the YubiKey, both positive and negative. The study coordinator conducting the interviews referenced the journals to guide his discussion of their experiences. They were also asked how they thought the support for security keys and the YubiKey itself could be improved. During this exit interview, participants also provided demographic information and completed the SUS questionnaire. Finally, the coordinator helped participants remove YubiKey-based authentication from each account and collected the YubiKeys.

B. Data Analysis

Similar to the approach used by Egelman et al. [22], we conducted an inductive analysis of the final interviews. Two researchers independently listened to audio recordings of each interview and took detailed notes. We then identified the most salient themes and compiled a codebook of 26 codes based on these themes.

Three researchers then independently listened to the audio recording of each interview and coded the interview. We used Fleiss' kappa scores to measure the inter-rater reliability between the three coders. Kappa values ranged from 0.56 to 1 with a mean of 0.70 and a median of 0.68.[12] This mean falls into the category of ”substantial agreement” as described by Landis and Koch [30]. The three coders addressed coding discrepancies by majority vote and by consulting the audio recordings for clarification. Our results are based on these reconciled codes.

C. Participant Safety

Because participants were using their personal accounts, we took significant precautions to minimize potential harm to participants. Our IRB provided external validation that our efforts to protect our participants were sufficient and complete. First, participants were given a sealed set of instructions detailing how to disable the YubiKey from their accounts should they want to exit the study early. Second, we enabled both the SMS and backup codes 2FA options for Google and Facebook so participants could access their account if the YubiKey was not functioning or unavailable. Third, participants were provided with a phone number that they could call at any time to receive technical support. Fourth, all participants were given the telephone numbers of several study coordinators in case they required emergency assistance to access their accounts.

Finally, to minimize harm in Windows 10 we required that participants use *YubiKey for Windows Hello*. While the two-factor authentication of the *Windows Logon Authorization Tool* would have been preferable and had even been approved by our IRB, we decided to amend our protocol for two reasons. First, the *Windows Logon Authorization Tool* requires that users unlink their Windows account from their Microsoft account. This had an unacceptable risk of causing harm to users by disrupting Windows features and applications that relied on a linked Microsoft account—for example, automatic file backup using OneDrive. Second, the *Windows Logon Authorization Tool* allows users to permanently lock themselves out of their computer if configured incorrectly, which was deemed to be too risky for the study. Additionally, *YubiKey for Windows Hello* allowed users to continue using their password as a backup means to authenticate to their device in case the YubiKey was not functioning or unavailable.

Although backup authentication modes were available for all accounts, participants were instructed to use the YubiKey to authenticate whenever possible.

D. Recruitment

We recruited 25 participants using posters (available in the appendix) on the Brigham Young University campus. We specified that participants, referred to hereafter as B1–B25, must have Google and Facebook accounts and own a Windows 10 laptop.

E. Demographics

Participants in our second study skewed slightly male: male (14; 56%), female (11; 44%). All participants were young adult: 18–25 years (21; 84%), 26–35 years (4; 16%). Four-fifths of the participants had completed some college but did not yet have a degree: some college with no degree (20; 80%), bachelor degree (2; 8%), post-graduate degree (3; 12%). Most participants reported having an intermediate level of computer expertise: beginner (1; 4%), intermediate (17; 68%), advanced (7; 28%).

F. Limitations

The study population was drawn from a college campus, so the results are not generalizable. Future research could replicate this study with different populations to gain additional insights into the usability of security keys in day-to-day usage.

Windows 10 authentication using the YubiKey was only single-factor authentication. While this was necessary to minimize harm to participants, future work needs to be done to better explore the usability of security keys for day-to-day authentication to the operating system.

VII. Measuring Day-to-Day Usability—Results

During the study, only two participants phoned study coordinators requesting assistance. In the first case, the participant asked how to use the YubiKey to log into Gmail on their phone. As we were not exploring this interaction in this study, the study coordinator told the participant how to log into Gmail using the backup 2FA options. In the second case, the participant reported that the YubiKey no longer worked with

[12] We dropped six results where the kappa score was low (<0.5).

Windows 10. The study coordinators were unable to solve this problem; this participant was instructed to continue using the YubiKey with only their Google and Facebook accounts.

During the study, participants encountered a variety of errors when attempting to authenticate using the YubiKey. Rarely, users were unable to authenticate to Google or Facebook with the YubiKey, even though they were using it correctly. However, these problems were quickly resolved by refreshing the page, reinserting the YubiKey, or restarting the browser. In contrast, nearly every user reported having a problem at some point using the YubiKey to log into Windows 10, with many reporting that the problems occurred frequently. In these cases, participants would temporarily use their PIN or restart their computers to regain access to their accounts.

In this section, we first report the SUS scores for YubiKey in this study. The remainder of the section reports on observations and comments made by participants in the exit interview.

A. SUS

Eleven participants (11; 44%) rated the YubiKey NEO and gave it a mean SUS score of 76.4 with a standard deviation of 15.7. Fourteen participants (14; 56%) rated the YubiKey Nano and gave it a mean SUS score of 71.9 with a standard deviation of 9.6. In total, YubiKey received a mean SUS score of 73.9 with a standard deviation of 12.5 in this second study. Based on comparisons to other systems and contextual descriptions provided by Bangor et al. and Sauro et al. [25], [26], [27], [28] (see Figure 8), this SUS score is considered "acceptable," and receives a B grade.

B. Trade-offs Between Security and Convenience

Five participants (5; 20%) mentioned that they had previously been warned about unauthorized attempts to access their Google or Facebook accounts, or had known someone whose account had been compromised. These participants were confident that adding 2FA to their accounts would secure them against future attacks:

> **B11:** *"The best part was just the extra security that it offered in addition to my password; I have had people try to hack into [my Google] account."*
>
> **B18:** *"I've known a lot of people that have had their accounts hacked; it is so much more hard to do that with [the YubiKey] because you have a physical copy of something."*

In contrast to the YubiKey experience with Google and Facebook, *YubiKey for Windows Hello* increased the convenience of authentication but reduced its security—as long as the participant's YubiKey was plugged into their device, users would not need to do anything to access their Windows 10 account. While this affects the YubiKey NEO experience, it was much more pronounced for the YubiKey Nano, which is intended to be left in the machine. This essentially removed any protection against an attacker who gained physical access to the machine. Three participants (3; 12%) noted that this was a significant concern:

> **B10:** *"[It was] kind of a double-edged sword: I liked having my computer open up and just log me in, but that meant if anyone got my computer it was much less safe because it would just login."*

C. Desire to Continue Using the YubiKey

Fourteen participants (14; 56%) indicated that they would like to continue using a YubiKey for 2FA. Of these, six (6; 24%) inquired about purchasing a YubiKey:

> **B19:** *"Where do you get [the YubiKeys]? . . . I loved them, it was really cool using them."*
>
> **B12:** *"I actually mentioned it to my wife and said, 'You need to get one of these, and if they're cheap, we should make them a Christmas present and get them for lots of people.' I think it is a good added security, especially for Google, bank accounts, big profile things that you want to make sure you've got a second-factor security on it."*

D. No Need to Further Protect Their Accounts

Of the ten participants (10; 40%) who indicated that they did not want to continue using the YubiKey, six (6; 24%) commented that they did not have any accounts that needed the extra protections provided by 2FA. These participants felt that either their accounts were of such low value that they would not be the target of an attacker or that even if their accounts were compromised, there was nothing of value to be lost. However, all six participants did mention that they would consider using a YubiKey in the future if they eventually became responsible for more sensitive information.

> **B14:** *"If I am carrying some top secret or high security stuff, I would like to have that sense of protection. But, at this point as a student, maybe not. They can steal my essays, whatever."*
>
> **B25:** *"If I had more confidential things, then I would use it. At this point I wouldn't because of the added steps and the little bit of added complexity to just check an email that wasn't confidential."*

E. Sharing Accounts Between Users

One unexpected use case that was brought up by three participants (3; 12%) was the use of account sharing, generally with a spouse. This proved difficult as only the user that currently had possession of the device was able to access these accounts. If the other user needed access to the account, they would have to borrow the YubiKey or use one of the backup 2FA schemes:

> **B2:** *"Let's say my wife needed to grab something from my email—she would try to login from her computer, but she couldn't. I know there are workarounds, like the codes you print out at the beginning . . . or maybe having two [YubiKeys]."*

The need to share accounts also led to some surprising interactions when needing to obtain the YubiKey from the other party:

> **B19:** *"The only time I had a problem was when my wife was driving and I needed to login to something, and [the YubiKey] was on the keys in the ignition. But I was able to slip it off the key ring and use it."*

F. The YubiKey Would Occasionally Insert Gibberish Text

Six participants (6; 24%) reported that when the YubiKey was plugged into their devices, random text would occasionally appear in the application they were using. Sometimes the application would also perform an operation on the random text—for example, Chrome performing a Google search for gibberish text. This problem was caused by users accidentally touching the YubiKey's button while using their computer, prompting the YubiKey to provide an authentication token. As the YubiKey communicates to the computer using the USB Keyboard interface, this authentication token would appear as a random-looking string of text followed by a carriage return.

Most participants recognized that this behavior was caused by the YubiKey, but still found it to be annoying:

> **B16:** *"I would accidentally touch [the YubiKey] and it would enter a string of numbers and letters . . . the first few times I didn't realize that it was the key."*
>
> **B1:** *"If you . . . just leave this plugged in and you accidentally tap the gold button on the side, it just spits out this string of numbers and then presses enter. So, if you have YouTube open, it unpauses your video; if you have Chrome open in a new tab, then it will spit out that string of numbers and then Google it."*

G. The YubiKey Nano Was Not Sufficiently Portable

Participants assigned the YubiKey Nano reported issues related to its small size. The YubiKey Nano is designed to be small enough to plug into a USB port and keep it there permanently (see Figure 1). Unfortunately, this did not match the participants' needs, as they often needed to use the YubiKey with multiple devices—for example, logging into their personal email account at work—forcing them to unplug the YubiKey Nano and carry it around with them.

Eight of 14 participants with the Nano reported that it was not sufficiently portable:

> **B23:** *"It's very small, which made it very hard to carry around. I would carry it around in a little container because I was unsure if I could put it on my keys or something like that. I wish it was a little bit bigger."*
>
> **B16:** *"If there was a version that was a tiny smidge bigger, it could have a little hole into where you could put it on your key ring."*

Three of the Nano participants were also reluctant to devote one of their limited number of USB ports on their laptop entirely to the YubiKey Nano:

> **B16:** *"It was a pain having it occupying space in one of my USB ports since I don't have many. It's just real estate that is wasted."*

Getting the YubiKey Nano out of the computer could also be very difficult, with one participant needing to use needle-nose pliers to extract the YubiKey Nano.

H. Participants Were Worried About Device Loss

While three participants (3; 12%) did temporarily misplace the YubiKey, no participant permanently lost it. Still, eight of the participants (8; 32%) indicated that they were worried that they would lose the YubiKey:

> **B15:** *"There was the potential that I might lose it—there is a small little tiny hole on the end of it, but there wasnt anything I could connect easily to my keychain . . . I was constantly worried about losing it."*

Backup authentication schemes conveniently fill this void. Five participants (5; 20%) told us they would like another YubiKey for sharing or backup. Other backup schemes are usually less secure than the primary authentication using a security key (e.g., SMS). The result is a quandary where the most secure backup schemes are often unavailable, and the most available schemes are less secure. This is a problem that requires more research.

I. Comparison to Other Forms of 2FA

Thirteen participants (13; 52%) explicitly stated that they preferred using the YubiKey to other forms of 2FA:

> **B11:** *"I prefer having the YubiKey because I don't have to worry about typing in a number, I can just press it really quick."*
>
> **B18:** *"The YubiKey is just easier to use. You can't always get to your phone, or [you] have to wait and type to type something in."*
>
> **B8:** *"It was faster than SMS two-factor authentication, more convenient because when you get a string of numbers . . . then you have to input that and could get it wrong. [The YubiKey] was just faster, I would prefer this over SMS."*

Four participants (4; 16%) told us they preferred other forms of 2FA to YubiKeys.

J. Security Keys on Mobile

All participants but one used other 2FA methods to authenticate themselves on their phones. One participant (1; 4%) discovered that the YubiKey NEO supported authentication over NFC and was very positive about this experience:

> **B15:** *"I think I might get the YubiKey that comes with NFC capabilities, because I like the potential of using that with my phone as well. Being able to use that to get into apps like LastPass really appeals to me too."*

VIII. DISCUSSION

After conducting both studies and analyzing the results, several issues stood out. This section discusses these issues in greater depth and serves as the background for the recommendations we make in the next section.

A. Usability Differences Between Setup and Day-to-Day Usage

Our work differs from prior work in that we studied the setup and day-to-day usage phases of 2FA in separate studies. In our results, we found that participants' experiences for each phase were drastically different. Our laboratory study revealed significant impediments with the setup process, and participants generally viewed the usability of the YubiKey poorly. In contrast, in the longitudinal study, the majority of the participants were quite pleased with their experience, rating YubiKey's usability very highly.

The results show that it is possible to have a system that is highly unusable during setup, but highly usable in day-to-day use. Research that only examined the setup phase would dismiss security keys as wholly unusable, whereas research focusing on the day-to-day usage phase would be overly optimistic about security keys' usability. Only by examining both phases is it possible to get a holistic view.

Additionally, while it is possible to study both phases in a single study, we caution against doing so. It would be difficult to design a study that ensures the setup experience does not confound usability results regarding day-to-day usability, especially if the two usability experiences are as diametrically opposed as they were in our study. While studies should eventually consider the interaction of these two phases, such studies are best performed after building a firm understanding of each phase in isolation.

B. Need for an Improved Setup Process

Our laboratory study revealed a multitude of issues with the setup process for security keys. Comparing the different success rates, it is clear that Google's setup experience was superior to either Facebook's setup or Yubico's Windows 10 software. We believe that this largely arises from the fact that Google used a wizard that walked users through enabling 2FA and setting up their security keys. In contrast, Facebook's wizard only helped users set up their security key, but did not help them enable 2FA on their accounts. YubiKey's Windows software provided no wizards, instead requiring users to read 17 pages of dense technical documentation. The participants that struggled to complete the setup tasks due to inadequate instructions illustrate Norman's Gulf of Execution [31]. As prior research has shown [32], wizards are more effective than written documentation at helping users complete otherwise complicated tasks. In line with this, our results suggest that a wizard is essential for guiding users through all phases of enabling 2FA and their security keys.

We also noticed that after setting up the YubiKey, many users immediately tried to authenticate with it. This served two purposes: to confirm to the user that they had successfully set up the YubiKey and to familiarize themselves with the new authentication experience. Unfortunately, none of the services included this component of setup in their wizard or documentation. Testing out the YubiKey was made more difficult by the fact that Google and Facebook would remember the device after the first authentication attempt, preventing users from authenticating with the YubiKey. The disconnect between activating a security key and not being able to immediately experience how it works illustrates Norman's Gulf of Evaluation [31]. In all cases, users would have benefited from a mechanism for testing out the YubiKey after they had finished registering it with the system.

Finally, we note that while it would be beneficial if individual service providers improved their setup experience, it would still leave an overall fragmented experience for users. Instead, users would be well-served by a standardized setup procedure that would substantially ease the burden of setting up a YubiKey with multiple services.

C. Account Sharing

The longitudinal study revealed that users occasionally share their accounts between multiple individuals—for example, spouses might have a shared Gmail account. This caused difficulty for participants in our study as it required them to trade off possession of the security key, preventing participants from accessing services if they had lent the security key to the other account owner. Also, while not studied in the paper, we note that disabled and elderly users often have especially critical needs to allow others to assist them with their computing tasks [33]. Enabling 2FA on their accounts is especially problematic when this assistance is provided remotely, preventing them from manually sharing a 2FA device.

A simple solution to these problems is to allow for more than one 2FA device to be registered with an account. Research could also explore alternative options for supporting account sharing—for example, using 2FA for remote authorization instead of authentication.

D. Native OS Support for U2F Security Keys

The U2F protocol is not natively supported by any operating system. Operating systems *do* allow security keys to function as traditional hardware tokens (i.e., smart cards), but this removes the benefits obtained by requiring that the user be present to tap the security key's button before it authenticates. In this mode, if the security key is left attached to the host device, as is the intended use case for the YubiKey Nano, a remote attacker will be able to attack the device as if it did not support 2FA at all. This is an unfortunate limitation, as these problems are well known [21], and U2F was designed to address them.

While Yubico did provide several software solutions for enabling a YubiKey to work with Windows 10, there were several limitations. First, it was easy for users to lock themselves out of their accounts accidentally. Second, it was difficult to configure correctly to attain U2F functionality. Third, in some configurations a YubiKey could only be used to log in at session resumption, not initial account log in. While these problems may be addressable through software patches, the preferred solution is for operating systems to fully implement the U2F protocol, maximizing the chance that users would have a consistent, usable experience.

IX. Recommendations

Based on the results of the user studies and our observations, we emphasize the following recommendations for improving U2F security keys and 2FA in general:[13]

(1) **Study setup and day-to-day usability separately.** Research into 2FA should analyze setup and day-to-day usability separately, ideally exploring them in separate studies.

(2) **Standardize the setup process.** The setup process for security keys, and 2FA in general, should be standardized across services, providing a uniform experience for users. This setup process should include wizards that provide active guidance to users.

(3) **Include clear indicators of success.** There should be an easy way to verify that setup was successful. Users must have clear and correct indicators of success, or direct access to a trial run of the authentication process.

(4) **Consider shared accounts.** Users need to share account access—for example, with a spouse—and 2FA systems need to support this use case. At a minimum, this means that all services need to allow users to register more than one 2FA device.

(5) **Integrate with operating systems.** Major operating systems should begin to provide native support for security keys. To prevent unexpected text input, the key should only activate in an authentication context.

(6) **Prevent lockouts.** Users must be made aware of, and be able to revert, failed security key setup. The *Windows Logon Authorization Tool* should be updated to prevent account lockouts, and *YubiKey for Windows Hello* should be modified to allow proper U2F-style authentication.

X. Conclusion

To explore the usability of U2F security keys for non-enterprise users, we conducted a laboratory and longitudinal study of the YubiKey, a popular line of security keys. Our laboratory study revealed significant impediments to correctly setting up a YubiKey, resulting in it being perceived as unusable. In contrast, the longitudinal study examined the day-to-day usability of a YubiKey, revealing that users find it highly usable. The vastly different usability results underscore the utility and importance of separately analyzing both the setup and day-to-day use of 2FA schemes.

During the laboratory study, many participants struggled to set up a YubiKey for Windows 10 and Facebook. In contrast, participants were much more successful setting up a YubiKey for the Google account. The higher Google success rates are evidence that the setup phase can be improved.

Participants in the longitudinal study were much more positive about the YubiKey, suggesting that if the initial usability hurdle for the setup phase could be overcome, YubiKeys could be a viable 2FA option for non-enterprise users. Additionally, we found that a majority of our longitudinal participants preferred U2F-based 2FA to other forms of 2FA. Our qualitative data analysis revealed important usability issues encountered by the participants and served as a basis for our recommendations.

XI. Acknowledgments

We thank the anonymous reviewers for their suggested revisions and our shepherd, Sascha Fahl, for his guidance. We also thank Michael Bailey, Dennis Eggett, John Fontana, Joshua Mason, Deepak Kumar, Jasmine Reynolds, and Kaylia Reynolds for their feedback on the paper. This work was supported in part by the National Science Foundation under Grant No. CNS-1528022 and Sandia National Laboratories.

References

[1] J. Bonneau and S. Preibusch, "The Password Thicket: Technical and Market Failures in Human Authentication on the Web," in *The Ninth Workshop on the Economics of Information Security*, 2010.

[2] J. Bonneau, C. Herley, P. C. Van Oorschot, and F. Stajano, "The Quest to Replace Passwords: A Framework for Comparative Evaluation of Web Authentication Schemes," in *2012 IEEE Symposium on Security and Privacy*. IEEE, 2012, pp. 553–567.

[3] D. Florencio and C. Herley, "A Large-Scale Study of Web Password Habits," in *Proceedings of the 16th International Conference on World Wide Web*. ACM, 2007, pp. 657–666.

[4] B. Ur, F. Noma, J. Bees, S. M. Segreti, R. Shay, L. Bauer, N. Christin, and L. F. Cranor, "'I added '!' at the end to make it secure': Observing Password Creation in the Lab," in *Proceedings of the Eleventh Symposium On Usable Privacy and Security*, 2015.

[5] A. Naiakshina, A. Danilova, C. Tiefenau, M. Herzog, S. Dechand, and M. Smith, "Why Do Developers Get Password Storage Wrong? A Qualitative Usability Study," in *ACM SIGSAC Conference on Computer and Communications Security*, 2017.

[6] J. Atwood, "You're Probably Storing Passwords Incorrectly." [Online]. Available: https://blog.codinghorror.com/youre-probably-storing-passwords-incorrectly/

[7] B. Ives, K. R. Walsh, and H. Schneider, "The Domino Effect of Password Reuse," *Communications of the ACM*, vol. 47, no. 4, pp. 75–78, 2004.

[8] J. Lang, A. Czeskis, D. Balfanz, M. Schilder, and S. Srinivas, "Security Keys: Practical Cryptographic Second Factors for the Modern Web," in *International Conference on Financial Cryptography and Data Security (FC)*. Springer, 2016, pp. 422–440.

[9] S. Srinivas, D. Balfanz, E. Tiffany, and F. Alliance, "Universal 2nd Factor (U2F) Overview," *FIDO Alliance Proposed Standard*, pp. 1–5, 2015.

[10] "What is FIDO?" October 2017. [Online]. Available: https://fidoalliance.org/about/what-is-fido/

[11] U. Piazzalunga, P. Salvaneschi, and P. Coffetti, "The Usability of Security Devices," in *Security and Usability: Designing Secure Systems that People Can Use*. O'Reilly, 2005, pp. 221–242.

[12] C. S. Weir, G. Douglas, M. Carruthers, and M. Jack, "User Perceptions of Security, Convenience and Usability for Ebanking Authentication Tokens," *Computers & Security*, vol. 28, no. 1, pp. 47–62, 2009.

[13] C. S. Weir, G. Douglas, T. Richardson, and M. Jack, "Usable Security: User Preferences for Authentication Methods in eBanking and the Effects of Experience," *Interacting with Computers*, vol. 22, no. 3, pp. 153–164, 2010.

[14] N. Gunson, D. Marshall, H. Morton, and M. Jack, "User Perceptions of Security and Usability of Single-factor and Two-factor Authentication in Automated Telephone Banking," *Computers & Security*, vol. 30, no. 4, pp. 208–220, 2011.

[15] S. Trewin, C. Swart, L. Koved, J. Martino, K. Singh, and S. Ben-David, "Biometric Authentication on a Mobile Device: A Study of User Effort, Error and Task Disruption," in *Proceedings of the 28th Annual Computer Security Applications Conference*. ACM, 2012, pp. 159–168.

[16] N. Karapanos, C. Marforio, C. Soriente, and S. Capkun, "Sound-Proof: Usable Two-Factor Authentication Based on Ambient Sound." in *24th USENIX Security Symposium (USENIX Security 15)*, 2015, pp. 483–498.

[17] S. Ruoti, B. Roberts, and K. Seamons, "Authentication Melee: A Usability Analysis of Seven Web Authentication Systems," in *Proceedings of the 24th International Conference on World Wide Web (WWW)*. International World Wide Web Conferences Steering Committee, 2015.

[13] Prior to submission, we made these results available to Yubico.

[18] K. Krol, E. Philippou, E. De Cristofaro, and M. A. Sasse, "'They brought in the horrible key ring thing!' Analysing the Usability of Two-Factor Authentication in UK Online Banking," *USEC Workshop on Usable Security at the Network and Distributed System Security Symposium*, 2015.

[19] E. De Cristofaro, H. Du, J. Freudiger, and G. Norcie, "A Comparative Usability Study of Two-Factor Authentication," in *USEC Workshop on Usable Security at the Network and Distributed System Security Symposium*, 2014.

[20] J. Brooke, "SUS—A Quick and Dirty Usability Scale," *Usability Evaluation in Industry*, vol. 189, no. 194, pp. 4–7, 1996.

[21] D. D. Strouble, G. Schechtman, and A. S. Alsop, "Productivity and Usability Effects of Using a Two-factor Security System," *Southern Association for Information Systems Research*, pp. 196–201, 2009.

[22] S. Egelman, S. Jain, R. S. Portnoff, K. Liao, S. Consolvo, and D. Wagner, "Are you ready to lock?" in *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, 2014.

[23] M. Boodaei, "Real-Time Phishing Takes Off," November 2010. [Online]. Available: https://securityintelligence.com/real-time-phishing-takes-off/#.VdOTBHhh1Bw

[24] J. Brooke, "SUS: A Retrospective," *Journal of Usability Studies*, vol. 8, no. 2, pp. 29–40, 2013.

[25] J. Sauro, *A Practical Guide to the System Usability Scale: Background, Benchmarks & Best Practices*. Measuring Usability, LLC, 2011.

[26] A. Bangor, P. T. Kortum, and J. T. Miller, "An Empirical Evaluation of the System Usability Scale," *International Journal of Human–Computer Interaction*, vol. 24, no. 6, pp. 574–594, 2008.

[27] A. Bangor, P. Kortum, and J. Miller, "Determining What Individual SUS Scores Mean: Adding An Adjective Rating Scale," *Journal of Usability Studies*, vol. 4, no. 3, pp. 114–123, 2009.

[28] T. S. Tullis and J. N. Stetson, "A Comparison of Questionnaires for Assessing Website Usability," in *Usability Professional Association Conference*, 2004, pp. 1–12.

[29] E. Hayashi and J. Hong, "A diary study of password usage in daily life," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2011.

[30] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, pp. 159–174, 1977.

[31] D. Norman, in *User Centered System Design: New Perspectives on Human-Computer Interaction*. Lawrence Erlbaum Associates, Publishers, 1986, ch. 3, pp. 31–61.

[32] L. Phelps, "Active documentation: Wizards as a Medium for Meeting User Needs," in *Proceedings of the 15th Annual International Conference on Computer Documentation*. ACM, 1997, pp. 207–210.

[33] B. Dosono, J. Hayes, and Y. Wang, "'I'm stuck!': A Contextual Inquiry of People with Visual Impairments in Authentication." in *Symposium On Usable Privacy and Security*, 2015.

XII. Appendix

A. Recruiting Posters

The recruiting posters for the two YubiKey studies are shown in Figures 6 and 7.

B. SUS Interpretation

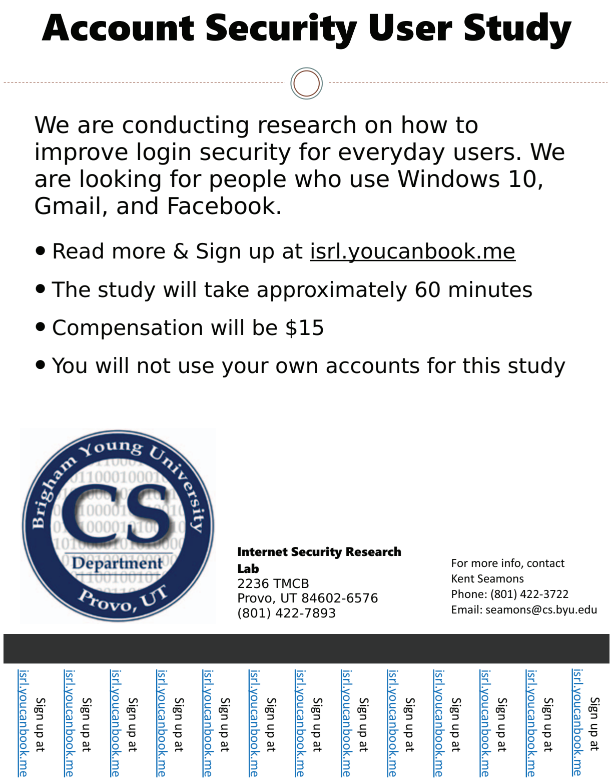
An adjective-based rating scale used for interpreting System Usability Scale (SUS) scores is shown in Figure 8.

C. Laboratory Study Survey

In the following survey, the word system refers to the YubiKey you used. All questions must be answered. If you feel you cannot answer one of the items, mark the center of the scale. Please record your initial reaction after carefully reading each question.

Strongly Disagree; Disagree; Neither Agree nor Disagree; Agree; Strongly Agree

(1) I think that I would like to use this system frequently.
(2) I found the system unnecessarily complex.



Account Security User Study

We are conducting research on how to improve login security for everyday users. We are looking for people who use Windows 10, Gmail, and Facebook.

- Read more & Sign up at isrl.youcanbook.me
- The study will take approximately 60 minutes
- Compensation will be $15
- You will not use your own accounts for this study

Internet Security Research Lab
2236 TMCB
Provo, UT 84602-6576
(801) 422-7893

For more info, contact
Kent Seamons
Phone: (801) 422-3722
Email: seamons@cs.byu.edu

Sign up at isrl.youcanbook.me
Sign up at isrl.youcanbook.me
Sign up at isrl.youcanbook.me
Sign up at isrl.youcanbook.me
Sign up at isrl.youcanbook.me
Sign up at isrl.youcanbook.me
Sign up at isrl.youcanbook.me
Sign up at isrl.youcanbook.me
Sign up at isrl.youcanbook.me
Sign up at isrl.youcanbook.me
Sign up at isrl.youcanbook.me
Sign up at isrl.youcanbook.me
Sign up at isrl.youcanbook.me

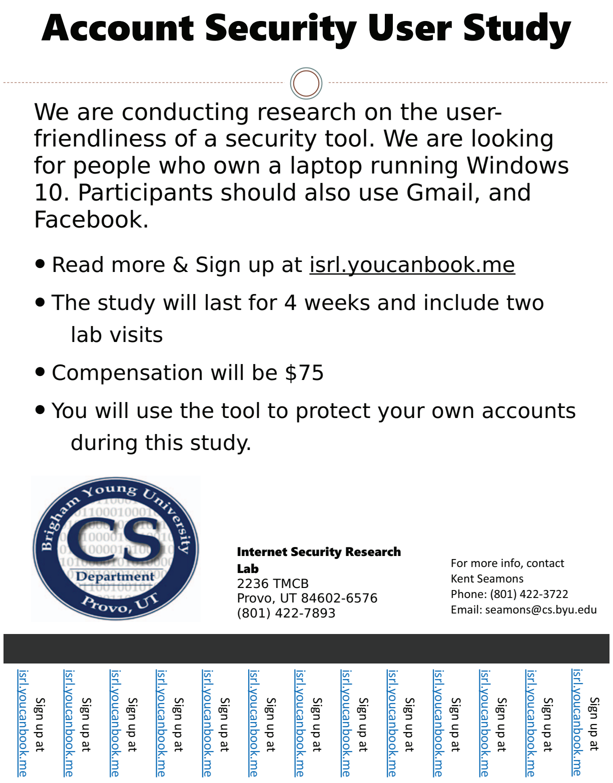
Fig. 6. Recruitment poster used to find participants for the laboratory study

(3) I thought the system was easy to use.
(4) I think that I would need the support of a technical person to be able 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.

Answer the following:

Strongly Disagree; Disagree; Neither Agree nor Disagree; Agree; Strongly Agree

(1) I would like to keep using a YubiKey.
(2) I would not always be able to find my YubiKey when I needed to log in.
(3) I could easily keep YubiKey around me to use whenever I log into all my accounts.
(4) Using the YubiKey got easier to use by the second or third time I set it up.
(5) It would be worth the extra work to use the YubiKey to protect my personal accounts.



Account Security User Study

We are conducting research on the user-friendliness of a security tool. We are looking for people who own a laptop running Windows 10. Participants should also use Gmail, and Facebook.

- Read more & Sign up at isrl.youcanbook.me
- The study will last for 4 weeks and include two lab visits
- Compensation will be $75
- You will use the tool to protect your own accounts during this study.

Internet Security Research Lab
2236 TMCB
Provo, UT 84602-6576
(801) 422-7893

For more info, contact
Kent Seamons
Phone: (801) 422-3722
Email: seamons@cs.byu.edu

Sign up at isrl.youcanbook.me

Sign up at isrl.youcanbook.me

Sign up at isrl.youcanbook.me

Sign up at isrl.youcanbook.me

Sign up at isrl.youcanbook.me

Sign up at isrl.youcanbook.me

Sign up at isrl.youcanbook.me

Sign up at isrl.youcanbook.me

Sign up at isrl.youcanbook.me

Sign up at isrl.youcanbook.me

Sign up at isrl.youcanbook.me

Sign up at isrl.youcanbook.me

Sign up at isrl.youcanbook.me

Fig. 7. Recruitment poster used to find participants for the longitudinal study

(6) I would rather get a code texted to my phone than use a YubiKey.

Have you ever used a 2-step login (a.k.a. 2-factor authentication) system before?
Yes, but I stopped using it; Yes, and I still do; No; I do not know

What did you like about the experience of setting up the YubiKey?
Free response

What would you improve about the experience of setting up the YubiKey?
Free response

How does a YubiKey make an account safer?
Free response

The last 4 questions are demographic information to help us learn about our participants' backgrounds.

Choose the range that includes your age.
18–25; 26–35; 36–45; 46–55; 56+

How would you describe your computer skills?
Beginner; Intermediate; Advanced

What is the highest level of education you have completed?
Some high school; High school diploma or equivalent; Some college, no degree; Associate's Degree; Bachelor's Degree; Post-Graduate Degree

What is your gender?
Male; Female

D. *Laboratory Study Coordinator Instructions*

(1) Invite the participant in. Leave the door to the room open. Give the participant the consent form and audio release to review and sign.
(2) Give participant the YubiKey 4 in its original packaging.
(3) Read the following paragraph to the participant:
During the course of this study, you will be using a security tool called a YubiKey. It is in this shipping envelope. Please, use the next five minutes to learn more about the YubiKey using the Internet.
(4) After five minutes, stop the participant and read the following 2 paragraphs:
If you have questions during the study, please ask them; however we may not be able to answer them in order to maintain the integrity of the study. Please comment on things you find easy to use and things you find hard to use. Let me know when you finish each of the three tasks. I may ask you to skip to the next section to make sure there is time to complete all the tasks.
During this study, we ask you to use the accounts provided on the worksheet, not your own personal accounts. Also use our phone numbers. Ask us if you have questions.
(5) Give them their instruction sheet and indicate which order to do the tasks as you write it down in your study notes.
(6) After 15 minutes on each task, cut them off. Thank them and ask them to please go on to the next task. Do not convey that they are incapable, but rather that our limited time requires that we move on. If you are short on time, divide the time evenly so they have 5-10 minutes at the end of the hour to fill out the survey and get the compensation form.
(7) Note how long each task took, mistakes made, and whether you cut them off.
(8) After all three tasks have ended, send them to the link to the Qualtrics survey.

E. *Laboratory Study Instructions to Participants*

Please complete the three parts of the study **in the order written below by your study coordinator.** You will have 15 minutes to complete each task. You may find the task easy or hard. Please keep the good and the bad experiences you have in your mind so you can tell us about them in the final questionnaire. The study coordinator is not allowed to help you during these tasks.

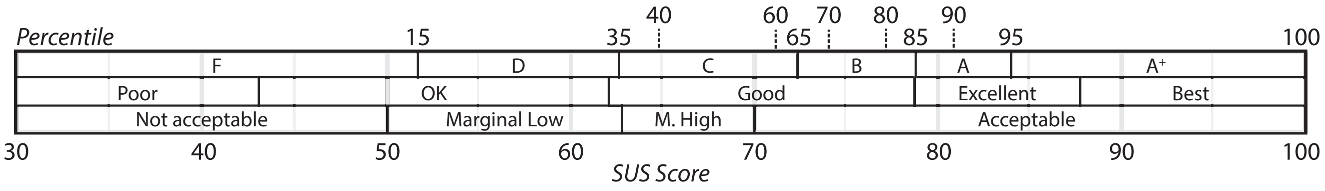


Fig. 8. Adjective-based ratings to help interpret SUS scores

First, do part [B; ordering changed per participant]
Second, do part [A; ordering changed per participant]
Last, do part [C; ordering changed per participant]

(A) Set up the YubiKey to be part of the login process for the following account at facebook.com. You can use the Internet to help you complete this task. When you are finished, let the study coordinator know you are ready to continue.
 (a) Username: [Facebook username]
 (b) Password: [Facebook password]

(B) Set up the YubiKey to be part of the login process for the following account at gmail.com. You can use the Internet to help you complete this task. When you are finished, let the study coordinator know you are ready to continue.
 (a) Username: [Google username]
 (b) Password: [Google password]

(C) Set up the YubiKey to be part of the login process on this computer for Windows 10. You can use the Internet to help you complete this task. When you are finished, let the study coordinator know you are ready to continue.
 (a) Username: [Windows username]
 (b) Password: [Windows password]

F. Longitudinal Study Survey

In the following survey, the word system refers to the YubiKey you used. All questions must be answered. If you feel you cannot answer one of the items, mark the center of the scale. Please record your initial reaction after carefully reading each question.
Strongly Disagree; Disagree; Neither Agree nor Disagree; Agree; Strongly Agree

(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 be able 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.

The last 4 questions are demographic information to help us learn about our participants' backgrounds.

Choose the range that includes your age.
18-25; 26-35; 36-45; 46-55; 56+

How would you describe your computer skills?
Beginner; Intermediate; Advanced

What is the highest level of education you have completed?
Some high school; High school diploma or equivalent; Some college, no degree; Associate's Degree; Bachelor's Degree; Post-Graduate Degree

What is your gender?
Male; Female

G. Longitudinal Study Coordinator Instructions

(1) Welcome participant. Give them the consent form to sign.
(2) Verbally explain what will happen.
(3) Have participant choose a time for their final interview.
(4) Setup 2FA with Google (let them choose an authenticator app or text). Demo login in an incognito window and on a mobile device. Print backup codes.
(5) Setup 2FA with Facebook (let them choose an authenticator app or text). Demo login in an incognito window and on a mobile device. Print backup codes.
(6) Setup YubiKey with Windows Hello (do not learn their PIN). Demo login process.
(7) Warn them that changing settings could lock them out. They do so at their own risk.
(8) Explain the expectation and purpose of the Authentication Journal.
(9) Give them the packet of instructions explaining how remove the YubiKey from their systems and account, the backup codes, and Authentication Journal.

H. Longitudinal Study Final Interview Guide

(1) What was the best part of using the YubiKey with a Google account?
(2) What would you improve about the user-friendliness of using the YubiKey with a Google account?

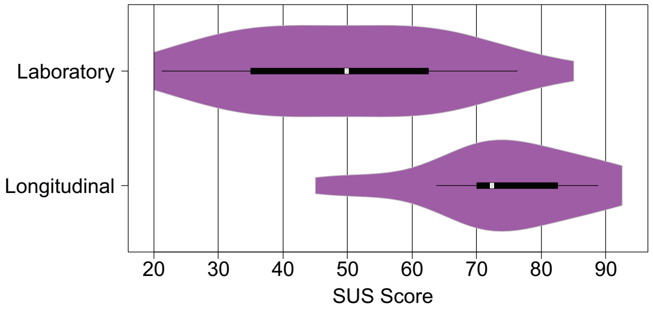


Fig. 9. Participant SUS scores for the laboratory and longitudinal studies

(3) What was the best part of using the YubiKey with a Facebook account?
(4) What would you improve about the user-friendliness of using the YubiKey with a Facebook account?
(5) What was the best part of using the YubiKey with a Windows 10 account?
(6) What would you improve about the user-friendliness of using the YubiKey with a Windows 10 account?
(7) What do you think are the benefits of using the YubiKey?
(8) What do you think are the problems caused by the YubiKey?
(9) How easy was it to have your YubiKey with you whenever you needed it?
(10) Did you ever get around having to use your YubiKey? If so, what were the circumstances?
(11) Please tell me why you would either like to use or not like to use a YubiKey as part of your everyday life.

I. SUS Scores

A violin plot of the SUS scores in the laboratory and longitudinal study is given in Figure 9.