"*It's Stored, Hopefully, on an Encrypted Server*": Mitigating Users' Misconceptions About FIDO2 Biometric WebAuthn

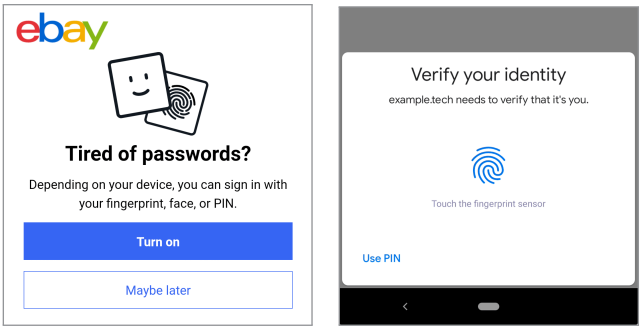
Leona Lassak[◇], Annika Hildebrandt[†], Maximilian Golla[⋆], Blase Ur[†]
◇ *Ruhr University Bochum,* † *University of Chicago,* ⋆ *Max Planck Institute for Security and Privacy*

Abstract

While prior attempts at passwordless authentication on the web have required specialized hardware, FIDO2's WebAuthn protocol lets users sign into websites with their smartphone. Users authenticate locally via the phone's unlock mechanism. Their phone then uses public-key cryptography to authenticate to the website. Using biometrics (e.g., fingerprint, face) for this local authentication can be convenient, yet may engender misconceptions that discourage adoption. Through three complementary studies, we characterized and sought to mitigate misconceptions about biometric WebAuthn. We also compared it to non-biometric WebAuthn and traditional passwords. First, 42 crowdworkers used biometric WebAuthn to sign into a website and then completed surveys. Critically, 67% of participants incorrectly thought their biometrics were sent to the website, creating security concerns. In remote focus groups, 27 crowdworkers then co-designed short notifications to mitigate biometric WebAuthn misconceptions. Through a 345-participant online study, we found that some notifications improved perceptions of biometric WebAuthn and partially addressed misconceptions, yet key misconceptions about where the biometric is stored partially persisted. Nonetheless, participants were willing to adopt biometric WebAuthn over non-biometric WebAuthn or passwords. Our work identifies directions for increasing the adoption of biometric WebAuthn by highlighting its security and usability.

1 Introduction

Despite their drawbacks, passwords remain widely used. A typical user can have hundreds of password-protected online accounts [44, 59]. To be secure, the user must create (and remember) a unique password for each service. If they reuse passwords across services, they are vulnerable to credential stuffing attacks, in which attackers exploit credentials breached from one service to attack accounts on other services where the user has a similar password [23, 43]. In response, online services have introduced two-factor authentication (*2FA*) [11, 51] and risk-based authentication [22, 62].



(a) WebAuthn **notification** used by eBay (June 2021, edited).

(b) WebAuthn **instructions** on a Google Pixel 3a (Android 11).

Figure 1: Examples of a site-specific notification used by eBay and OS-specific instructions for authenticating.

A user can also adopt a password manager to facilitate unique passwords [45]. Sadly, adoption rates for these mechanisms remain low [37, 45] and industry has thus continued to search for an alternative to passwords for signing into websites [26].

One of the most promising approaches for passwordless web authentication is the FIDO2 Project [5] and its web authentication (*WebAuthn*) protocol. The core idea is to use public-key cryptography in place of a password. To register with a website, the user's *authenticator* (hardware token or other device) creates a public-private keypair unique to that website. Subsequent authentication attempts proceed via a challenge-response protocol, overcoming many disadvantages of passwords and also stopping phishing attacks [32]. Hardware tokens (e.g., YubiKeys) are commonly used as authenticators [5]. Recent user studies have demonstrated that using WebAuthn with a hardware security key achieves substantial benefits relative to passwords in both security and usability [18, 32]. The cost and inconvenience of security keys, however, are impediments to widespread adoption [18, 32].

Fortunately, smartphones can also be used as FIDO2 authenticators. Smartphones' ubiquity and familiarity to users makes this support a crucial advance beyond prior attempts at passwordless online authentication. The private key is stored

in a trusted enclave on the smartphone. The user authorizes each use of the private key via their usual mechanism for unlocking their phone. This unlock mechanism ultimately is a PIN, pattern, or password. However, schemes like Apple's Touch ID [9] and its Android equivalent enable users to unlock their phone with a biometric, such as a fingerprint or face. As a result, these biometrics can also be used to sign into a website [5], an interaction we term *biometric WebAuthn*.

This ability to authenticate to websites using only a fingerprint or other biometric holds great promise. Because of WebAuthn's basis in public-key cryptography, it is far more secure than a password. Similar to how support for biometrics made phone unlocking much more convenient [7, 14, 63], biometric WebAuthn promises better usability than passwords. Furthermore, support for WebAuthn is quickly being added by major websites, including eBay, Microsoft, and Yahoo [40].

Unfortunately, when biometric phone unlocking was introduced, misconceptions were initially widespread [7, 14, 63], and we hypothesized the same would hold true for biometric WebAuthn. The superficial appearance that a user is signing into an online service with only their fingerprint or face suggests the potential for even more problematic misconceptions about biometric WebAuthn's security and usability. These misconceptions could discourage the adoption of biometric WebAuthn. Thus, our research focused on users' initial encounters with biometric WebAuthn and their resultant expectations. We anticipate that many users will encounter biometric WebAuthn for the first time via a small notification on a website encouraging them to adopt the technology. Such short notifications cannot possibly capture FIDO2's technical complexities. However, after looking at such a notification for a few seconds, many users will form expectations about the scheme's security and usability, ultimately deciding whether to adopt biometric WebAuthn based on very little information. While future work should examine how to better educate users about how FIDO2 actually works, we focused on understanding and improving these initial impressions and perceptions.

We thus conducted three complementary user studies to understand and mitigate misconceptions about biometric WebAuthn, as well as to compare it to *non-biometric WebAuthn* (e.g., using a PIN) and site-specific *passwords*.[1] For all studies, we did not expect participants to have any prior knowledge of biometric WebAuthn or how it worked. Rather, our intention was to understand their initial expectations relating to security, privacy, usability, and trust. Study 1 and 3 were conducted on participants' personal Android phones, which was the most common, fully supported FIDO2 configuration at the time of the study [20]. While we focused on phones, FIDO2 aims to be widely available on many other platforms.

Our first research goal was to understand how a user who encounters biometric WebAuthn in the aforementioned brief encounter extrapolates about its properties. Thus, in Study 1, 42 crowdworkers used biometric WebAuthn on an Android phone to register and later authenticate at a website we controlled. To understand participants' preconceived notions, we intentionally gave little information about how WebAuthn worked. Through surveys, we investigated how participants thought biometric WebAuthn worked and explored potential misconceptions suggested by either the literature or WebAuthn's design. Critically, 67% of participants incorrectly thought their biometrics were sent to our website or elsewhere outside their phone, leading to other misconceptions.

To help mitigate misconceptions we observed, especially those that might discourage adoption, we then focused on the design of short *notifications* that websites can display. For example, Figure 1a shows eBay's current notification. Designing any notification that conveys complex technical concepts in a short and simple format is a challenge in many areas of security [21]. To this end, Study 2 engaged 27 participants in seven online focus groups. After a moderator taught participants how biometric WebAuthn worked and the group discussed their perceptions of WebAuthn, participants engaged in iterative co-design of new notifications. We distilled participants' ideas into six potential notifications. To align notifications with what users would actually want to learn, our co-design focus groups took an unrealistically large amount of time to help non-technical users better understand how biometric WebAuthn works. They then collaboratively proposed new notifications. This co-design approach, which aims to benefit from end-users' creativity and opinions [39], was motivated by prior research in which notifications for 2FA, TLS, cryptographic APIs, and phishing prevention benefited from similar focus groups [4, 25, 49, 60].

Finally, Study 3 compared these notifications inspired by our co-design focus groups. Each of 345 crowdworkers was assigned to use biometric WebAuthn (with one of those notifications), non-biometric WebAuthn, or a site-specific password. Similar to Study 1, participants created an account on our website and answered survey questions. One notification was a baseline representing how early-adopter companies currently advertise WebAuthn. Relative to this baseline, some notifications created through co-design improved perceptions of biometric WebAuthn's security and partially addressed some key misconceptions. Furthermore, most participants were willing to adopt biometric WebAuthn over non-biometric WebAuthn and passwords for trustworthy websites. Nonetheless, many participants still held key misconceptions, especially about where their biometric is stored, highlighting the need for more expansive education efforts.

Collectively, our findings provide the first comprehensive examination of user misconceptions and perceptions about using biometrics on phones for authentication on the web. We discuss how our findings can influence how websites communicate with users about biometric WebAuthn, helping to spur adoption and move toward a passwordless web.

[1] Our survey instruments and screenshots of the notifications we tested are in our extended version [31]. Our FIDO2 implementation is available at: https://github.com/UChicagoSUPERgroup/fido2biometrics.

2 Background and Related Work

We first detail how FIDO2 and WebAuthn work. We also review the literature on FIDO2, biometric authentication, and security warning design. We finish with related work applying participatory design in the security and privacy domain.

2.1 FIDO2 and the WebAuthn Protocol

The Fast IDentity Online (FIDO) Alliance is an industry association formed to build a passwordless user experience by authenticating users via public-key cryptography. In the past, the FIDO Alliance was best known for its Universal 2nd Factor (U2F) specification enabling strong cryptographic two-factor authentication [26]. U2F's successor, FIDO2, enables passwordless web authentication via two components. First, the Client to Authenticator Protocol 2 (*CTAP2*) standardizes communication between a client and (external) authentication hardware. Second, the Web Authentication (*WebAuthn*) specification defines a JavaScript-based API allowing web services to authenticate users via public-key cryptography.

To register on a web service, the user's *authenticator* (hardware) creates a public-private keypair unique to that website. The authenticator can either be an external hardware key (*roaming authenticator*) connected to a device via USB, NFC, or Bluetooth, or a trusted module on the user's existing computer or smartphone (*platform authenticator*). To sign into a web service, the authenticator signs a cryptographic challenge received from the server. The server then verifies the signature using that user's public key, received during account registration. In contrast to password-based authentication, FIDO2 resists phishing, replay attacks, and breaches of the server.

Users authorize their authenticator's use of the private key either by confirming their presence via a button press or by authenticating locally (*user verification*). When using a smartphone as a platform authenticator, the phone's unlock mechanism is typically used for this verification step. While *non-biometric* unlock mechanisms (e.g., PIN, pattern, password) can be used, so can *biometric* mechanisms (e.g., fingerprint, face, iris). The latter is particularly notable because the use of biometrics for phone unlocking is perceived as very convenient and is widely adopted [7, 14, 63], making biometric WebAuthn a highly promising alternative to passwords for authentication on the web. In the rest of this paper, we use *biometric WebAuthn* as shorthand for using a biometric for user verification within the FIDO2 protocol suite.

Note that the user interface for verification (cf. Figure 1b) differs by OS, browser, and vendor. Furthermore, a web service can require the use of either roaming or platform credentials. Similarly, services can specify whether user verification (as opposed to mere presence) is required.

With the standardization of WebAuthn by the W3C in 2019, popular online service like Dropbox, eBay, Facebook, GitHub, Microsoft, and Twitter have begun to implement FIDO2-based single- or two-factor authentication. A key challenge is to migrate existing users from passwords to WebAuthn. FIDO2 involves complicated technical concepts and terminology. Thus far, services have abstracted away most of these technical details and security properties (cf. Figure 1a), instead focusing on convenience and ease of use to encourage adoption. Unfortunately, users might be left with incorrect mental models of WebAuthn as a result.

2.2 Prior Studies of FIDO2 and WebAuthn

While we focus on biometric verification within FIDO2, prior work has primarily studied hardware security keys. In a lab study with 94 participants, Lyastani et al. [32] studied user perceptions of using FIDO2-compatible hardware security keys as a single factor for authentication. Participants were randomly assigned to register and sign into a website with either a security key or a site-specific password. Participants generally preferred the security key over traditional passwords, but identified limitations. They had concerns about several hardware issues, such as access on computers without USB ports. They also desired the ability to recover and revoke access if the security key was lost. Unsurprisingly, participants' mental models of FIDO2 lacked the natural understanding of traditional passwords.

In a field setting, Farke et al. [18] observed the authentication routines of 10 employees in a small software company. Employees were given the choice between using a FIDO2-compatible security key and a traditional password to log in. Over four weeks, several employees stopped using the key as its security benefits were perceived as unnecessary and it was slower than using their browser's password manager.

Oogami et al. [40] had 10 participants use biometric WebAuthn to register their Android phones with their existing Yahoo! Japan accounts. Some participants were confused by the user interface, mistakenly pressing the fingerprint icon on screen rather than the actual fingerprint sensor. While their results highlighted some usability issues with biometric WebAuthn, their small sample limited generalizability.

Independent of the use of biometrics, FIDO2 and WebAuthn have usability drawbacks [2, 33, 41, 42, 54]. The inability to transfer private keys across devices requires users to register multiple times (e.g., on both a phone and laptop). There is no secure fallback if authenticators are lost or broken, shifting the problem from primary to fallback authentication. There is also no security benefit if *insecure* methods, like a password, remain valid even after FIDO2 is enabled.

2.3 Misconceptions About Biometrics

While we are among the first to study the use of biometrics within FIDO2 and WebAuthn, prior work has investigated users' mental models of biometric authentication in other

contexts. In this section, we highlight (and number) key misconceptions identified in prior work. We investigate these potential misconceptions, among others, in our user studies.

Specifically, several user studies have investigated misconceptions about using biometrics to unlock smartphones [9, 12, 34, 36, 57]. In a survey of smartphone users, De Luca et al. [14] found that usability was one of the major factors for choosing biometrics to unlock mobile devices. Apple's Touch ID was considered as ① **easy and fast** as the normal *slide to unlock*. Interestingly, security and privacy concerns did not play a large role in decisions about adopting biometric unlocking. Bhagavatula et al. [7] also found usability to be crucial to user acceptance. They reported misconceptions about the ② **storage location**, with a few participants thinking that biometrics were sent over the network or to the cloud even when unlocking the phone. Revisiting this misconception is a focus of our study because we expected misconceptions that websites process a user's biometric itself to heavily influence perceptions. Participants considered biometric authentication ③ **more secure** than PINs, not realizing that PINs remained enabled as fallback authentication. Many participants were unaware of security risks like ④ **spoofing**.

To delve further into misconceptions about biometric phone unlocking, Wolf et al. [63] conducted semi-structured interviews. Based on misconceptions about ⑤ **biometrics being processed** (whether the data stored enables reconstruction of a user's face/fingerprint), both experts and non-experts expressed concerns about biometric data being ⑥ **accessed by third parties**. They also highlighted misunderstandings around using ⑦ **multiple devices**, ⑧ **delegating access** to others, and ⑨ **availability** due to wet/oily fingers.

2.4 Notification Design

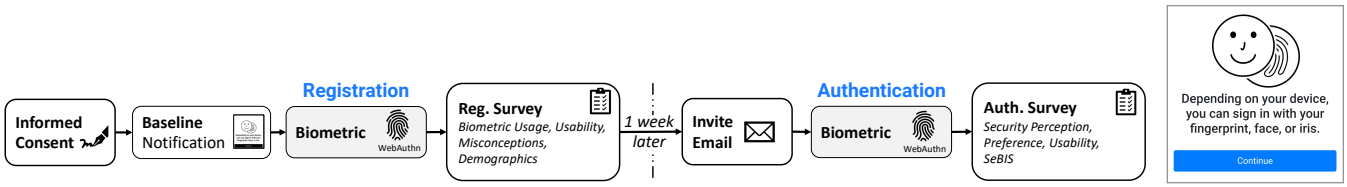
A large body of prior work evaluated and improved warning messages and notifications, including browser warnings in general [1, 6, 8, 28], as well as warnings about phishing [15, 16, 46], malware [3], and PDF downloads [29]. Early work by Egelman et al. [15] studied the effectiveness of phishing warnings. They found that the warnings were ineffective overall, with high click-through rates. They recommended clear action instructions, making them more distinguishable from less severe warnings to prevent habituation, and to make them blocking, full-screen, active warnings. The most extensive set of work studied TLS warnings [6, 21, 52, 53]. To improve the adherence of warnings, the use of opinionated design proved to be effective in a study by Felt et al. [21]. Egelman et al. [16] showed how small design changes can increase the time users spend looking at a notification. More recently, Reeder et al. [50] conducted a survey on browser warnings in situ with Google Chrome and Mozilla Firefox users. They did not find major issues in modern browser warnings, concluding that future improvements only need to be made on smaller, contextual misunderstandings. These real-world improvements show the value and importance of research on designing effective notifications. The best practices for communicating about security identified in this literature informed the design of our notifications. However, our notifications do not aim to warn or stop users, but instead aim to correct misconceptions to spur the adoption of WebAuthn.

2.5 Participatory Design and Focus Groups

Warning designers have commonly applied heuristics or expert views during the development and improvement process [21, 46, 56]. For the design of the notifications in this work, we use focus groups and apply a participatory design (*PD*) approach. PD describes a technique where prospective users are actively involved in the development and design process of new products or interfaces. End users can contribute valuable insights about issues experts are unaware of by challenging implicit assumptions and preconceptions the experts might have [39]. Within security and privacy, Weber et al. [60] applied PD to develop new TLS warning messages. They described the approach as "suitable and versatile" for interface design in the security domain. Their participants stated that existing notifications were too long, complicated, technical, and that they would prefer warnings that are short, focus on recommended actions, and use more concrete and alarming language, which falls in line with prior research on warning design [6]. Research by Redmiles et al. [49] is closest to our approach since they also studied notifications that aim to encourage adoption. Their PD sessions with demographically diverse users revealed that using personalized headlines, bullet points, and the color blue can increase 2FA adoption. Contrary to prior work, their participants chose to avoid graphics since they found those less professional. Gorski et al. [25] used PD to target professional developers to improve security-related console warnings. They found that design recommendations that apply to end users do not necessarily align with experts' wishes. Althobaiti et al. [4] worked with focus groups on improving the usability of phishing reports [4]. McNally et al. [35] applied PD with children to improve and extend the functionalities of mobile child protective apps. Chouhan et al. [10] used PD to design a smartphone app that allows collaborative decision-making for privacy and security. To elicit mental models of HTTPS, Krombholz et al. [30] used a drawing task, a technique we also use.

3 Study 1: Online Study of Misconceptions

The goal of this study was to understand what misconceptions users might have about using biometric WebAuthn. In this two-part study, participants registered and authenticated at a website we created, *ExampleTech*, using their personal mobile device. Our implementation is based on Spomky-Labs' PHP WebAuthn Framework [38]. We modified the



(a) Study 1 procedure. (b) Notification.

Figure 2: An overview of the structure of Study 1 (L), as well as the intentionally vague baseline notification (R) we used.

account registration steps and WebAuthn settings such that only platform authenticators were allowed, user verification was required, and timeout occurred after 60 seconds. Our code is available on GitHub.[1] Note that we used the same WebAuthn implementation for Study 3 (Section 5).

3.1 Method

Figure 2a depicts the overall study flow. Participants were recruited via Prolific for a study with two parts, *registration* and *authentication*, conducted a week apart. We required participants be age 18+, live in the US or UK, and have a 95%+ approval rating. We required that participants have an Android phone (running Android 7+), Google Chrome, and biometric phone unlocking configured and enabled. FIDO2 fully supports this configuration [20] and it reflects the devices and software supporting WebAuthn at the time we conducted our study. All our study protocols were approved by the University of Chicago Institutional Review Board (IRB). We paid participants $5 for each of the two phases of the study.

Registration Phase: Participants first created an account on the *ExampleTech* website. The main goal of Study 1 was to establish a baseline for misconceptions and opinions about WebAuthn. We were interested in participants' opinions about WebAuthn's pros and cons overall, as well as relative to passwords. Thus, we crafted an intentionally vague **baseline notification** that informed participants, "Depending on your device, you can sign in with your fingerprint, face, or iris." This was based on real-world notifications (cf. Figure 2b). After showing this notification, we simulated a sign-up page by asking participants to provide their age and gender. When participants pressed "register," the WebAuthn protocol began, launching an OS-specific WebAuthn instruction screen. At this point, the user *locally* authenticated on their phone (e. g., with their fingerprint or fallback mechanism, such as a PIN). The wording and graphics on the WebAuthn instruction screen varies across vendors. Figure 1b shows an example for a Google Pixel 3a device running Android 11 with a PIN as a fallback. Our instructions requested participants authenticate using a biometric, not a PIN, pattern, or password.

If the participant successfully authenticated, their account was created and they were redirected to complete the registration survey. If they failed to create an account, they completed an alternate survey aiming to understand why the failure occurred. The registration survey began with questions about the participant's use of biometrics both in registering on *ExampleTech* and in unlocking their phone. Participants also responded to the System Usability Scale (*SUS*) about using WebAuthn to register for *ExampleTech*. To gauge participants' mental models about WebAuthn and the technology it replaces (passwords), we then asked participants to describe how they believed WebAuthn and passwords worked behind the scenes. We also asked specific questions about where their authentication data (e. g., biometric or password) or data derived from it is stored. We hypothesized misconceptions about where biometrics are stored, and with whom they are shared, might heavily influence opinions about WebAuthn.

The next section solicited a series of multiple-choice responses using Likert scales and free-text justifications related to additional potential misconceptions surrounding security and usability when registering an account for a website using WebAuthn. We developed this series of questions about misconceptions through iterative piloting to investigate the relevant misconceptions observed in prior user studies about biometric authentication outside the WebAuthn context (cf. Section 2.3). The survey ended with demographic questions.

Piloting: We developed our questions based on previously documented misconceptions about biometrics, WebAuthn, and 2FA. We focused on misconceptions that were relevant in the biometric WebAuthn context (cf. Section 2). Since our main goal was to document users' initial expectations and potential misconceptions when interacting with biometric WebAuthn, participants were not expected to have any prior knowledge of WebAuthn, nor expected to have any technical knowledge. Therefore, to minimize biased, confusing, or technical wording, we conducted formal think-aloud cognitive pilot interviews with three domain experts and four non-technical pilot participants. Based on responses from the pilots, we iteratively refined the survey wording and flow.

Authentication Phase: The authentication phase followed a similar procedure. Participants were asked to sign into the *ExampleTech* website with the account they had created previously. Participants who were not able to create an account in the previous week were not invited to complete this follow-up session. The one-week waiting period was intended to minimize the effect of account creation on the login process. Typically, users log in more frequently than they create accounts, so it was important to explore both separately.

Once participants had signed in, they were asked a set of questions regarding possible misconceptions with the login. If they failed to log into their account, they were given two more attempts. After this, they were redirected to an alternate survey aiming to understand why the failure occurred.

Survey questions were developed similarly to the registration phase. The authentication-phase survey asked participants more direct questions about their opinions regarding the use of biometric WebAuthn. These questions were asked at the end of the authentication phase to avoid priming participants when they were responding to questions that aimed to understand their initial expectations.

We also provided a help page with a note that they needed to use the same device as they used for account registration. If participants tried to use a device or browser that was not Android or Google Chrome, they were automatically redirected to this help page, where we also provided an option to exit the study early with partial compensation.

3.2 Participants

We recruited 50 participants, 42 of whom registered successfully. 41 registered with their fingerprint, and one with their face. Of the eight participants who were unable to register, four did not meet the study requirements (three had no locking mechanism configured, while the other ran Android 6), one failed to authenticate within 60 seconds, and three encountered an unspecified error indicating an issue with their phone or settings. The entire registration process, including account creation and the survey, took a median of 18 minutes. The entire authentication process, including the survey, took a median of 9.5 minutes.

Of these 42 participants, 39 returned for the authentication phase, and 33 of them were able to authenticate successfully. Of the 42 participants who were successfully registered, 23 were men and 19 were women. Our participants tended to be young, with 8 who were 18-24, 21 who were 25-34, 9 who were 35-44, and 4 who were 45+. Participants' highest level of education attainment was as follows: 9 had a post-graduate degree, 20 had a college degree, 5 had completed some college without a degree, 7 had a high school diploma, and 1 had not completed high school. Finally, 26 had no background in technology/IT, 14 did, and 2 did not answer.

3.3 Key Security Misconceptions

The most severe misconception we identified was the belief that biometric data is stored in the online service's database.

Storage Location: The key misconception held by participants was where biometrics were stored when using biometric WebAuthn. Only 14 participants (33%) correctly identified that biometrics were stored on their phone. The majority, 20 participants (48%), believed biometrics were stored on the server or in a remote database operated by the website ("*Within the ExampleTech servers which would hopefully be secure,*" P07). Eight participants expressed uncertainty in where their biometrics might be stored. When we asked if an employee of *ExampleTech* would have access to their biometric data, the majority (83%) disagreed, yet only 12 participants (28%) justified their answer based on their biometric data being stored locally. Among those who thought the biometric data was stored somewhere other than their phone, reasoning ranged from believing it was stored in an encrypted format to believing sites had a moral obligation to protect private data (it would be a "*breach of trust*"). Two participants argued that an employee would not have physical access to their phone, so they would not be able to access the biometric information.

Processing of Biometric Data: Only 24 out of the 42 participants correctly thought their biometric would be safe from an attacker who stole data from the website's database. Seven thought the attacker would have their biometric, indicating that they likely believed it would be stored in the website's database. Another 11 were unsure whether an attacker would have access to biometrics stored on their phone, indicating uncertainty about how the biometric data is processed and whether the resulting data would allow an attacker to reconstruct a participant's face or fingerprint.

Third-Party Access: Prior research [63] found that some users are concerned about their biometric data being transmitted to third parties. Our participants did not hold this concern. However, many did not realize their data never leaves their device. 14 participants thought their biometric or data derived from it are sent to the *ExampleTech* server. Only four were positive that their biometric data is not sent outside their phone. Eight participants vaguely described their understanding as a local verification of their biometric on their device.

Lost Phones: We asked participants if someone who found their phone could access their account. 39 participants said no, stating that this person would not have access to their physical biometric ("*No one can steal your fingerprint from you,*" P36). Only three participants said the person probably could access their account. No participant mentioned the possibility of logging in using the PIN, pattern, or password instead.

3.4 Key Usability Misconceptions

The most problematic usability misconception was that participants believed they could sign into their *ExampleTech* account using a different device.

Availability: A key misconception that participants had was how the fallback mechanism used to unlock the phone (e.g., PIN, pattern, or password) could be used to log in if the biometric failed. Only 12 participants believed they would still be able to sign into their account if their biometric failed, while 25 incorrectly believed they would be unable to do so. Five participants were aware that they could use their phone's PIN or password in place of their biometric. Participants commonly stated that, if their biometric failed, they would

not be able to sign in because they had not yet set up a fallback method ("*I won't because, that's the only sign in method that I used during registration,*" P45). Five participants stated they would need to create a separate password or contact *ExampleTech*'s support hotline. Other participants did not even identify the possibility of a backup system, believing the biometric was the only option to authenticate.

Multiple Devices: Misconceptions around device sharing were common, with 11 participants indicating that they would be able to log into their account on a device other than the one where they registered ("*My fingerprint wouldn't have changed so I should be able to log in,*" P29). This finding again indicates a misunderstanding about the underlying functionality of WebAuthn. The current WebAuthn specification [5] does not permit transferring the private key across authenticators [54], requiring a roaming authenticator or an alternative scheme to register a new device. The biometric or its fallback scheme (PIN, pattern, or password) are only used to decrypt and unlock the private key on the device. Even if participants were aware that their biometric data is not stored with the website and that they cannot log in from another device, the explanations given for not being able to sign in were incorrect. 18 participants thought they could not sign in because their biometric data is not registered in their friend's phone. Only six participants correctly explained that the login and fingerprint is tied to the device that they used for registration ("*. . . because it's linked to the device I created it on,*" P24). Broadly, this misconception is reasonable as signing in from multiple devices is possible with traditional passwords.

Delegating Access: When asked if a trusted person could be given access to the account without the participant present, 39 participants thought there was no way since the friend would not have their biometric ("*They wouldn't be able to except if they cut my hand or there's another form like a password,*" P14). Only one participant mentioned a potential fallback option, and only three pointed at the possibility of registering a friend's biometric on their phone to grant access.

3.5 Versus Other Authentication Methods

We investigated whether the misconceptions we observed were specifically related to WebAuthn or also applied to biometric authentication in other contexts. In contrast to biometric WebAuthn, participants thought that their biometric data is only stored locally when it comes to phone unlocking. When it comes to passwords, participants had a better understanding of the processing and storage.

Comparison to Non-biometric Methods: Participants showed a clearer understanding of where their biometrics are stored when they unlock their phone. 30 out of the 42 participants said they believed their biometrics are only stored on their personal device. The remainder either thought they were stored on the cloud or with their phone manufacturer. However, only 8 participants reasoned that biometrics are only stored locally; 4 more argued that an employee of the phone manufacturer would not have physical access to their device. 7 participants stated that an attacker having access depends on how biometrics are stored, such as in an encrypted format. 12 participants had similar reasoning when considering an employee at the phone manufacturer not having access. This indicates that participants still lack a full understanding of how biometric data is used to unlock their phone, which is a more familiar process than biometric WebAuthn.

Comparison to Passwords: We also asked participants if they thought an employee of a website on which they use a traditional password would have access to their plaintext password. The majority, 26 participants, correctly understood that an employee would not have access because the password is "encrypted" or more generally that access to it is restricted by law. Nevertheless, we also identified misconceptions surrounding password security. Six participants argued that the password is stored with the website so the employees must have access, while three more said that only employees like IT administrators would have access. When considering hackers, most participants showed a correct understanding of the relevance of the storage format. Some argued that "encryption" (hashing) will prevent an attacker from actually having their password. Others noted it is relatively easy to circumvent the security precautions taken with passwords. Four participants mentioned personal experience with password breaches ("*happened in the past and has been in the news*," P14).

We also asked participants whether they considered passwords or biometric login to be more secure. In line with previous research [7], most participants argued that the biometric login is more secure. They mentioned well-known attacks on passwords, like shoulder surfing, or they stated a belief that a biometric cannot be copied or guessed ("*Unlike passwords, one's fingerprint can never be guessed,*" P15).

Participants strongly preferred biometric WebAuthn over passwords. Most argued from a convenience point of view, with 20 mentioning the process was easy and seven stating it was fast. Ten argued that using biometrics for authentication is more secure. Nine pointed out that, unlike a password, the biometric data cannot be forgotten and that there is no need to remember it in the first place. Two participants noted that no one can impersonate them as biometric data is unique.

4 Study 2: Co-design Focus Groups

In our second study, we followed a co-design (participatory design) approach to create more effective ways to communicate the security and usability advantages of biometric WebAuthn. In particular, we hoped to counteract the misconceptions identified in Study 1. Participants were asked to come up with single-screen notifications that addressed misconceptions and communicate the advantages of biometric WebAuthn.

4.1 Method

As detailed in Section 2.3, co-design focus groups have been used in past security research to help elicit user perceptions that may not surface in individual interviews. Participants in groups challenge the researchers', and each others', views and preconceptions. This facilitates identifying a middle ground. Inexperienced end-users can be more creative, open minded, and less biased than the researchers, which enriches the notification design process. We conducted 7 online focus groups with 2 to 7 participants per group. Focus groups lasted 75 minutes. Participants were compensated $25. Each group had at most one participant with technical background knowledge. Individuals were recruited via Prolific and were asked to participate in a small group meeting via a video conferencing platform. To protect their privacy during a session, we encouraged participants to select a pseudonym. After asking for consent, we audio recorded each session. All sessions started with a series of warm-up questions where participants described their feelings towards passwords and experiences with biometrics. Similar to Lyastani et al. [32], we created a video[2] to present the mechanics of account creation and sign-in because biometric WebAuthn would be unfamiliar to many participants. The video intentionally did not try to explain the underlying public-key-cryptography-based authentication process. To allow participants to form their own opinions, we did not mention any potential advantages or disadvantages of biometric WebAuthn. Participants were asked to share their initial impressions afterwards.

We then provided participants with 1 out of 4 resources from trusted sources [19, 24, 27, 64] that explained WebAuthn. In selecting articles, we required they have imagery, mention biometric login, and include an explanation of WebAuthn. They should take no longer than 5 minutes to read and contain no technical details (e.g., code snippets). We found appropriate articles on the first 5 pages returned when searching for terms like "What is WebAuthn" and "Passwordless Authentication." Moreover, we provided another document specifically addressing the misconceptions identified in Study 1. This document can be found in our online appendix [31].

After participants had read the articles, we asked them to explain what they understood about the WebAuthn login process, making sure to address any confusions or inaccuracies. Participants also elaborated on the most surprising aspect of the process, where they thought their biometrics was stored, and whether they would use it. We asked participants to identify what was unclear, left out, or satisfactory in the resources. At the end, participants were asked to each come up with a phone-screen-sized notification briefly explaining biometric WebAuthn to someone without prior knowledge. We also asked participants to draw a sketch that would support their explanation. Finally, everyone presented their explanations and drawings, and the group as a whole decided on the most crucial points that should be part of a "perfect" explanation. Those central elements were shared with future focus groups.

[2] Video demonstrating signing into a website using WebAuthn: https://youtu.be/wPzfEGTlcfA, as of June 2, 2021.

4.2 Participants and Overall Perceptions

Overall, 29 people participated in 7 focus groups. We excluded the data from two participants because they did not participate fully in activities or discussions due to technical issues. Of the 27, 19 were women and 8 were men. 69% of the participants were between 18 and 34 years old, 27% were between 35 and 44, and 4% were 45+. 86% of participants had at least some college education, the majority with a bachelor's degree. 18 participants were iPhone users, while the rest were equally distributed among Samsung, Sony, and Huawei. We asked their opinion on different authentication mechanisms. They were surprisingly positive when speaking about passwords. Only a handful explicitly mentioned disliking passwords. P1 was the most emphatic, saying, *"I hate passwords with a passion."* The most common complaint was the number of passwords that need to be remembered. Most participants had experience with using biometrics. A third of participants expressed liking biometrics due to their convenience. Five mentioned trust issues with biometrics.

Confusion About WebAuthn: The provided resources [19, 24, 27, 64] helped to identify further misconceptions. During discussion, two participants expressed the misunderstanding that WebAuthn was a platform where you create an account, which then handles your logins for you. Two other participants confused passwordless WebAuthn with two-factor authentication (the biometric functions as a second factor). Three participants showed misunderstandings surrounding hardware security tokens. One of them interpreted the token as a device to store the biometric. Confirming a finding from previous work [13], two participants thought the token was an external fingerprint scanner. Three participants misinterpreted the challenge that is signed with the private key during authentication as a strong password. In general, fallback authentication was a major concern.

4.3 Desirable Features of Notifications

Text Content: We observed two central features in most of the notifications participants created. Participants tended to stress either the ① **convenience** or the ② **security** of biometric WebAuthn. Overall, 7 participants' notifications described WebAuthn as **fast**, 9 as **easy**, and 21 as **safe** and/or **secure**.

The ③ **storage location** of the biometric data was a key component of the notifications. 16 participants mentioned where the biometric was stored, and 13 included who had access to their biometric data. Biometric data being "*only stored on your device*" was the most common wording, used by 13 participants. Three participants chose the wording "*it never leaves your device*," and one used "*it is only stored lo-*

cally." Three participants mentioned that "*no one except you*" has access to your biometrics. That "*no third parties*" have access to biometric data was mentioned by three participants. Four explained that the "*the website*" has no access either. The fact that ④ **hackers cannot get a hold of biometric data** was included by four participants. ⑤ **Comparison to passwords** was a common approach participants used. From a convenience point of view, WebAuthn eliminates the necessity to remember many passwords, which was mentioned by 11 participants. Seven said passwords are easy to hack, and three said that the biometrics in WebAuthn cannot be hacked. A controversial point followed by an enthusiastic discussion was whether it would be beneficial to include that WebAuthn is supported and co-developed by ⑥ **popular brands** like Microsoft, Google, or Apple. Participants preferred not to include technical details. The complete list of notification elements can be found in our online appendix [31].

Supporting Visuals: The most common style of the supporting images participants drew was a *protocol flow* with arrows representing the inner workings or steps a user would have to take to log in with WebAuthn. A third of participants preferred to draw a representation of a *login interface*.

The most common elements of images, drawn by 16 participants, were personal devices like ① **phones or computers**. 15 participants drew ② **biometric features**, such as fingerprints, eyes, or faces. To better explain the communication between the device and the website, 7 participants drafted a representation of a ③ **website or a server**. Popular visual metaphors were ④ **physical keys and locks**. Adding to the discussion about trustworthiness of certain ⑤ **brands**, 2 participants added logos or mentioned well-known brands like Google and Apple. To convey that WebAuthn is not tied to specific websites, 4 participants included representations of services like Facebook or Amazon. Interestingly, even though the ⑥ **storage location** of the biometric data played a central role in the written explanations, only 5 participants represented this in their images. The complete list of drawing elements can be found in our online appendix [31].

Consensus Notifications: At the end of each focus group, we asked the group to reach consensus on the central points of a notification. From this, we identified four key aspects:

1. **Security:**
   (a) WebAuthn is safe, secure, and private.
   (b) My biometric data is stored locally on my phone. Nobody has access to it. It cannot be hacked.
   (c) WebAuthn was developed by trusted companies.
2. **Convenience:** WebAuthn is fast, easy, and convenient.
3. **Comparison to Passwords:** WebAuthn is better than passwords, which have security/convenience drawbacks.
4. **Availability:** WebAuthn can be used on different websites, but you cannot access your accounts from multiple devices (if you have not registered them first).

We used these items as the starting point for the notifications we designed for Study 3 to address these misconceptions.

5 Study 3: Comparison Study

Study 3 had two goals. First, we aimed to compare the biometric WebAuthn notifications co-designed with participants in Study 2. These notifications themselves aimed to address misconceptions identified in Study 1. Second, we wanted to compare *biometric WebAuthn* to (i) *non-biometric WebAuthn* using a smartphone as a platform authenticator and (ii) traditional site-specific *passwords*. To this end, we conducted a between-subject protocol similar to Study 1 (Section 3).

5.1 Method

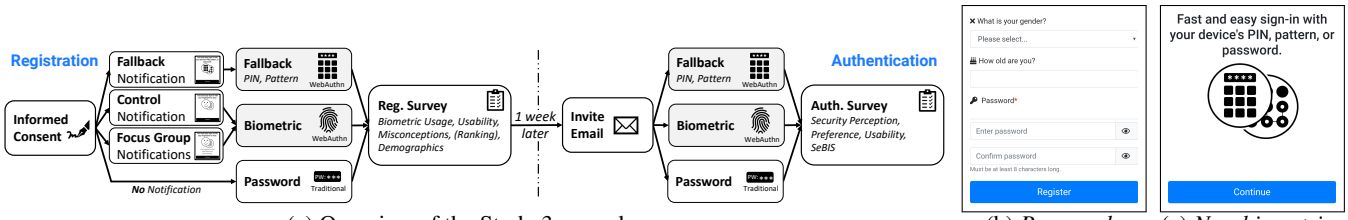
Figure 3a summarizes the protocol. Participants were randomly assigned to one of three groups specifying that they use *biometric WebAuthn*, *non-biometric WebAuthn* (e.g., unlock PIN, pattern, or password), or a *site-specific password*. Additionally, biometric WebAuthn participants were assigned one of six different notifications addressing misconceptions. The notification was shown directly before account creation. Participants were again recruited via Prolific, and those from Studies 1 and 2 were excluded. Compensation was $5 each for the registration and authentication parts.

Design of Biometric WebAuthn Notifications: We developed the notifications based on the consensus participants came to in Study 2. "Security" and "convenience" were the two broad categories those participants wanted to emphasize. We developed baseline language for those concepts through an 80-participant pre-study following the same protocol as Study 3. We compared "Fast and easy…," "Safe and secure…," and "Safe, secure, fast, and easy…," each followed by "…sign-in with your fingerprint or face" (displayed in our online appendix [31]).

The majority of the participants in the pre-study rated "fast and easy" as their favorite, so we used this language for all notifications in Study 3. Our baseline notification, ***Biometric-Control***, contained only this language. The five other notifications appended other concepts participants in the focus groups wished to emphasize, using the terminology that emerged from the focus groups:

- ***Biometric-Brands***: "Backed by Microsoft, Google, and Apple."
- ***Biometric-Hacked***: "Unlike passwords it can't be hacked."
- ***Biometric-Leaves***: "Your fingerprint or face never leaves your personal device."
- ***Biometric-Stored***: "Your fingerprint or face is only stored on your personal device."
- ***Biometric-Shared***: "Your fingerprint or face is never shared with *ExampleTech* or third parties."

Most notifications address where the biometric is stored, a key concern from the previous studies. *Biometric-Hacked* also compares WebAuthn to passwords. Even though the trust aspect *Biometric-Brands* represents only appeared in two focus groups, we tested it since it spurred substantial discus-



(a) Overview of the Study 3 procedure. (b) *Password* (c) *Non-biometric*

Figure 3: An overview of the Study 3 protocol and the visuals for the *Password* and *Non-biometric* conditions from Study 3.

sion in those groups. The ***Password*** condition saw a typical password-creation screen (Figure 3b). The ***Non-biometric*** condition saw the parallel "Fast and easy sign-in with your device's PIN, pattern, or password" (Figure 3c). Figure 4 shows the notifications for the six biometric conditions.

Survey Design: The surveys for both the registration and authentication phases were largely the same as in Study 1. However, in the registration survey, participants who successfully registered were also asked questions aimed at understanding their impressions of WebAuthn after being presented with all of the different notifications. The order in which the notifications appeared was randomized in order to avoid any ordering bias. Additionally, participants were only shown all of the notifications after they answered all questions relating to misconceptions in order to avoid priming them. As with Study 1, participants were not expected to have any technical expertise or prior knowledge of WebAuthn. Survey questions sought to understand participants' initial expectations.

Analysis Methods: Whereas Study 1 was primarily qualitative, Study 3's between-subjects design enabled quantitative comparisons across conditions. When comparing either numerical variables (e.g., timing) or ordinal responses on Likert scales, we first performed an omnibus Kruskal-Wallis H test (*KW*). In cases where the omnibus test was not significant, we report the distribution of responses across all conditions. If the omnibus test was significant, we performed (and report) pairwise, post-hoc Wilcoxon rank-sum tests. For categorical data, we used Fisher's Exact Test (*FET*). We also asked a few questions that compared all notifications within-subjects. Because each participant answered all questions in a repeated-measures design, we use the Friedman test, performing pairwise, post-hoc tests using Eisinga et al.'s method [17]. We set $\alpha = .05$. To control for multiple testing, we corrected p-values using the Benjamini-Hochberg method within each family of tests, as well as within each set of pairwise contrasts.

5.2 Participants

A total of 345 participants completed the registration phase, while an additional 29 failed to register for an account (similarly to Study 1, due to incompatible hardware, an incompatible web browser, or the phone failing to recognize a fingerprint). Of the 345 participants who successfully registered, 322 returned for the authentication phase, and 303 authenticated successfully. The registration phase (including the associated survey) took a median of 21 minutes, while the authentication phase took a median of 15 minutes. Between 40 and 49 participants were randomly assigned to each condition.

Of the 345 participants who successfully registered, 197 were men, 143 were women, 4 were non-binary, and 1 preferred not to answer. Participants were again relatively young, with 19% age 18–24, 39% age 25–34, 25% age 35-44, 11% age 45–54, and the remaining 5% age 55+. Among participants, 24% had a post-graduate degree, 40% had a college degree, 23% had completed some college without a degree, and 13% finished high school. Finally, 65% of participants had no background in technology/IT, 33% did, and 2% preferred not to answer. Asked if they had "heard of the terms WebAuthn or FIDO2," 18% reported they had. Most of them (79%) had first encountered it within the last year.

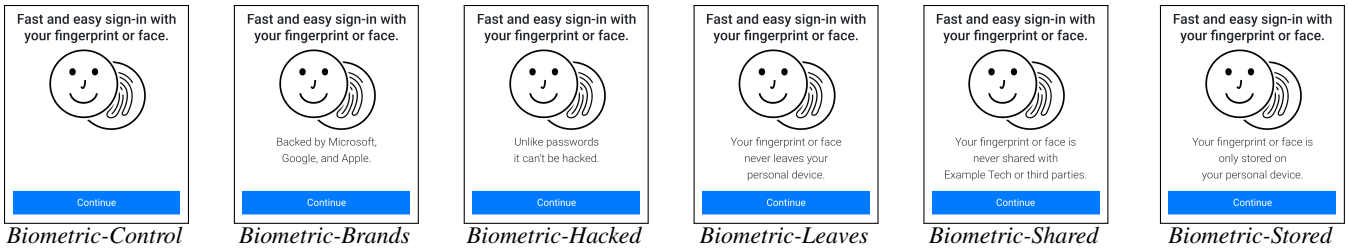
Among participants, 44% had a Samsung phone, 15% a Huawei phone, and 12% a Google phone. Across all 345 participants, 95% had enabled fingerprint unlock, 24% face unlock, and 5% iris unlock. Participants used either a four-digit PIN (48%), pattern (26%), PIN of another length (22%), or password (8%) as their non-biometric fallback mechanism.

5.3 Registration and Authentication

The 44 participants in the *Password* condition created an *ExampleTech*-specific password, of which 6 appeared (based on heuristics) to have been auto-generated by Chrome. Following recommendations from the literature [55], we required passwords be 8+ characters long and have a *zxcvbn* [61] strength score of 3+ (resisting $\geq 10^8$ guesses). The median PGS [58] *min_auto* guess number was 10^{13}, and the mean zxcvbn strength score was 3.4. The 40 participants in the *Non-biometric* WebAuthn condition used the method they typically use for unlocking their phone: a four-digit PIN (21 participants), pattern (14), PIN of another length (4), or password (1).

Most participants assigned to a biometric condition used their fingerprint. Of those 261 participants, 256 authenticated with a fingerprint, three with their iris, and two with their face. This preference toward fingerprints was also evident in the methods participants had enabled for phone unlocking.

The time it took participants to register an account on *ExampleTech* varied across conditions (KW $\chi^2(7) = 104.9$,



Biometric-Control *Biometric-Brands* *Biometric-Hacked* *Biometric-Leaves* *Biometric-Shared* *Biometric-Stored*

Figure 4: The notifications shown to participants in the six biometric WebAuthn conditions.

$p < .001$). The median time for the six biometric conditions ranged from 4.6 to 5.1 seconds, compared to 9.7 seconds for *Non-biometric* and 22.6 seconds for *Password*. These differences were statistically significant between all six biometric conditions and both *Non-biometric* (all $p < .001$) and *Password* (all $p < .001$). Registration required a median of a single attempt in all conditions. Participants in all conditions found the registration process highly usable, with a median score of 90.0 on the System Usability Scale (*SUS*).

The time it took to authenticate a week later also varied across conditions (KW $\chi^2(7) = 32.1$, $p < .001$). The median time to authenticate in the six biometric conditions ranged from 3.9 – 4.9 seconds, compared to 5.9 seconds for *Password* and 7.6 seconds for *Non-biometric*. The difference between *Non-biometric* and all seven other conditions was statistically significant (all $p \leq .001$), though no other pairwise differences (including compared to *Password*) were significant. As with registration, participants in all conditions found the authentication process highly usable, with a median SUS score of 95.0, which did not vary significantly across conditions.

Participants assigned to make an *ExampleTech*-specific password were less successful at authenticating than participants who used biometric WebAuthn. In our omnibus test, the proportion of participants who successfully authenticated varied across conditions (FET, $p < .001$). Whereas only 76% of *Password* participants successfully authenticated, 93% of *Non-biometric* participants and 95%–100% of participants in the six biometric conditions did so. In pairwise, post-hoc contrasts, we found that the difference between all six biometric conditions and *Password* was either significant or marginally significant (FET, $.030 \leq p \leq .086$ for all six comparisons). The 10 *Password* participants who were unable to log in reported forgetting their password.

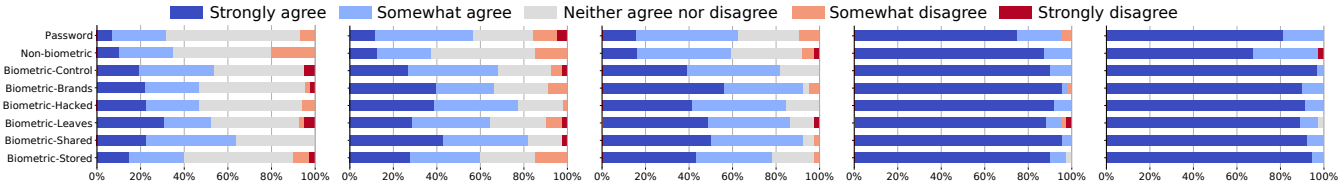
5.4 Overall Perceptions of Security/Usability

Participants responded on a Likert scale to broad statements about the security, privacy, and easiness of the process of creating an account. We found that the authentication mechanism, and to a lesser extent the biometric WebAuthn notification shown, impacted perceptions of security and privacy.

As shown in Figure 5b, the distribution of responses to the statement "I think account creation at *ExampleTech* is secure" varied across conditions (KW $\chi^2(7) = 29.4$, $p < .001$). In each of the six biometric conditions, at least 60% of participants strongly or somewhat agreed with this statement. At the high end, 82% of *Biometric-Shared* and 78% of *Biometric-Hacked* participants strongly or somewhat agreed. In contrast, only 38% of *Non-biometric* participants strongly or somewhat agreed. For *Password*, this number was 57%. Agreement that account creation is secure was significantly higher for *Biometric-Shared* than for *Non-biometric* ($p < .001$) and *Password* ($p = .003$), while the difference with *Biometric-Stored* was marginally significant ($p = .084$). Similarly, agreement was significantly higher for *Biometric-Hacked* than *Non-biometric* ($p < .001$) and *Password* ($p = .009$). It was also significantly higher for *Biometric-Brands* than for *Non-biometric* ($p = .020$), and marginal compared to *Password* ($p = .080$). Finally, agreement for *Biometric-Control* was significantly higher than for *Non-biometric* ($p = .043$), as well as marginal for *Biometric-Leaves* compared to *Non-biometric* ($p = .080$). Across biometric groups, the most common justification for perceiving registration as secure was the general fact that they used biometrics (20% of participants). For example, P56 wrote, *"Biometrics are usually pretty reliable."*

Responses to "I think account creation at *ExampleTech* protects the privacy of my *fingerprint/PIN/...*" also varied by condition (KW $\chi^2(7) = 17.1$, $p = 0.025$), as shown in Figure 5a. Agreement was significantly higher for *Biometric-Shared* than for *Non-biometric* ($p = 0.025$) and *Password* ($p = 0.025$). Whereas 64% of *Biometric-Shared* participants felt their data was kept private, only 35% and 32% of *Non-biometric* and *Password* participants, respectively, thought so. No other contrasts were significant. In free-text justifications, 33 participants wrote that their assigned mechanism protects their privacy because the biometric stays on their phone. Of those participants, 24% saw *Biometric-Shared*, 24% saw *Biometric-Stored*, and 30% saw *Biometric-Leaves*.

In all conditions, participants found account creation easy. Across conditions, 89% of participants "strongly" agreed, and 9% "somewhat" agreed with the statement "I think account creation at *ExampleTech* is easy." Figure 5d shows these responses, which did not vary significantly by condition. The primary justification for perceiving WebAuthn as easy was that it was fast. Several participants in the biometric groups also noted that there was no need to remember passwords. For example, P163 wrote, "*I don't have to remember any pass-*



(a) Registration protects **privacy** of *{fingerp., face, PIN, pass., ...}*. (b) Registration is **secure**. (c) Sign-in is **secure**. (d) Registration is **easy**. (e) Sign-in is **easy**.

Figure 5: In Study 3, participants responded to Likert-scale questions about the security, privacy, and ease of use of registering for an account using their assigned mechanism: biometric WebAuthn, non-biometric WebAuthn, or a site-specific password.

words and I don't have to worry about losing the password." These results were echoed in the authentication phase.

In the second part of the study, we asked similar questions about whether sign-in was secure and easy. As shown in Figure 5c, perceptions of security again varied by condition (KW $\chi^2(7) = 34.0$, $p < .001$). Agreement was significantly higher in all six biometric conditions than in *Non-biometric* (all $p \leq .031$). It was also significantly higher in all six biometric conditions than in *Password* (all $p \leq .040$). Whereas 60% of *Non-biometric* participants and 63% of *Password* participants "somewhat" or "strongly" agreed, the same was true for 78%–93% of participants in the biometric conditions. As shown in Figure 5e, perceptions that sign-in was easy also varied by condition (KW $\chi^2(7) = 21.8$, $p = 0.008$). *Non-biometric* participants rated sign-in as significantly less easy than *Biometric-Control* ($p = .044$), *Biometric-Hacked* ($p = .045$), *Biometric-Shared* ($p = .045$), or *Biometric-Stored* ($p = .044$). 40% of biometric participants, but only 7% of those in *Password*, said sign-in was easy because it was fast.

5.5 Security Misconceptions

Storage Location: To quantify participants' mental models about where data about their biometric or non-biometric fallback (e.g., PIN) is stored when using WebAuthn, we asked a multiple-choice question: "From the list below, where do you think your *chosen biometric/fallback secret* (or data derived from it) is stored when you created an account at *ExampleTech*? Select all that apply." Based on the misconceptions observed in Study 1, we provided the following options: "on my phone"; "on *ExampleTech*'s computers"; "on your phone manufacturer's computers"; "on the computers of a third-party that handles the login process"; "on the computers of another third-party"; and a fill-in-the-blank "other."

In WebAuthn, the biometric is stored only on the phone, a design decision that is critical for users' privacy. Unfortunately, only 40% of participants across the 7 WebAuthn conditions correctly chose only "on my phone" for where their biometric or non-biometric fallback secret is stored. While 55% of *Biometric-Stored* and 50% of *Biometric-Leaves* participants chose only "on my phone," compared to between 33% and 39% in the five other conditions, these differences were not statistically significant (KW $\chi^2(6) = 7.7$, $p = 0.265$). 22% incorrectly chose that the data is stored only on *ExampleTech*'s computers, 10% chose that it is stored in both of those places, and 7% incorrectly chose that the data is stored "on the computers of a third-party that handles the login process."

We asked a parallel question about where data is stored for unlocking a phone. In stark contrast, 71% of participants across the seven WebAuthn conditions correctly chose only "on my phone" for where their data is stored for unlocking their phone. Even though the actual storage location is identical, far more participants had misconceptions about where their data is stored in WebAuthn than in phone unlocking.

Other misconceptions followed from misunderstandings about biometric storage. We asked, "Do you expect a member of our research team who maintains the *ExampleTech* website to have access to your biometric *data*?" 15% of participants incorrectly answered "probably" or "definitely yes," mostly because they thought data is stored in *ExampleTech*'s database or that employee access to the data is needed for maintenance reasons. Interestingly, participants in all biometric conditions except *Biometric-Leaves* and *Biometric-Stored* wrote that biometric data does not have any value to employees or that employees have no reason to access it.

Processing of Biometric Data: We asked participants to respond to: "If an attacker stole data from the *ExampleTech* database, do you think the attacker would have your *fingerprint/PIN/...*?" Responses varied across conditions (KW $\chi^2(7) = 41.1$, $p < .001$). 43% of *Non-biometric* participants incorrectly believed attackers could "probably" or "definitely" get their non-biometric mechanism for unlocking their phone. In contrast, no more than 20% of participants in any of the six biometric conditions incorrectly believed the attacker could "probably" or "definitely" get their biometric. These differences were statistically significant between *Non-biometric* and all six biometric conditions (all $p \leq .020$), as well as between *Password* and all six biometric conditions (all $p \leq .005$). Assumptions that biometric data is stored securely or encrypted was the primary reason participants gave as to why an attacker would not have access. In contrast, other participants thought hackers could gain access because they are highly skilled. However, all of these participants missed the key point that the website does not store biometrics at all.

Third-Party Access: Between 23% (*Biometric-Stored*) and 39% (*Biometric-Leaves*) of participants incorrectly thought their biometrics or data derived from them were sent to a third party for processing. That a confirmation or vaguely defined "login token" is sent by the third-party service upon successful authentication was mentioned by between 12% and 20% of participants per biometric condition. In contrast, 14% of *Biometric-Leaves* participants (far more than in any other group) correctly noted that no biometric data is sent.

Lost Phones: Next, we asked if someone who found a lost phone could access the participant's account. Responses varied by condition (KW $\chi^2(7) = 54.6$, $p < .001$). Whereas at most 10% of participants in any biometric condition answered "probably" or "definitely" yes, 41% of *Password* and 15% of *Non-biometric* participants gave those answers. Agreement was significantly higher for *Password* and *Non-biometric* compared to all six biometric conditions (all $p < .001$ and all $p \leq .036$, respectively). Biometric participants' free-text justifications emphasized that someone who found their phone would not have their biometric. Other participants mentioned biometrics' inherent security "features," such as their uniqueness or difficulty to fake. Only 7 participants across all biometric groups correctly noted that attackers could, in fact, just use the device's fallback PIN, pattern, or password.

Cross-Site Usage: We also investigated whether participants thought an employee of a trustworthy/untrustworthy website where the participant is also registered could access their account at *ExampleTech*. 48% of *Password* participants expected an untrustworthy site could leverage their password to sign into other sites, which is the case if the user reuses passwords across sites. In contrast, 40% of *Non-biometric* participants and up to 39% of participants in each biometric condition incorrectly also thought they were putting themselves at risk. We asked a parallel question, replacing "untrustworthy website" with "companies, such as eBay, Google, and Microsoft." Only 15% of participants answered "probably yes" or "definitely yes," highlighting the need for further education that WebAuthn can be used safely even on potentially untrustworthy websites.

5.6 Usability Misconceptions

Availability: Only 38% of participants across the six biometric conditions correctly realized they could sign into their account even if the scanner failed to read their biometric; 46% incorrectly thought they would be unable to do so. Most commonly, incorrect free-text justifications suggested that participants were unaware that biometric WebAuthn always has a non-biometric fallback. Other participants assumed they could use classic reset mechanisms like email recovery or calling a website's service hotline.

Multiple Devices: Among participants, 34% thought they "probably" or "definitely" could log into *ExampleTech* via WebAuthn while using a different device (e.g., a friend's phone), while another 17% were unsure. Most participants who thought they could log in from another device incorrectly assumed their biometric was stored by *ExampleTech*. A few, however, chose this answer, (correctly) realizing that they could register a separate account on another device.

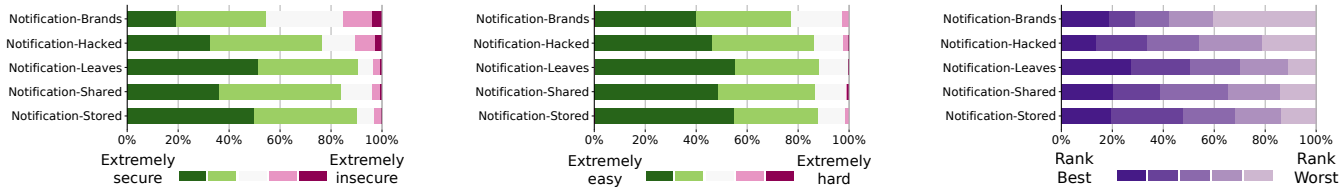
Delegating Access: Asked if they think a friend or family member could sign into the participant's account with their permission, 10% or fewer of the participants in any biometric condition answered "probably" or "definitely yes." This result emphasizes that the vast majority of participants were unaware that they could share their phone and its non-biometric fallback method (e.g., the phone's unlock PIN) to delegate access. While we found no significant differences across biometric conditions, perceptions did vary across conditions (KW $\chi^2(7) = 107.4$, $p < .001$). Specifically, most participants in *Password* (52%) and *Non-biometric* (60%) realized they could delegate access, which is significantly higher than in any biometric group (all $p < .001$).

5.7 Comparison of Notifications

To further understand what the notifications we designed based on Study 2 communicated, we asked a series of questions to all biometric participants about the five notifications other than *Biometric-Control*. Here, condition names refer only to the notification; every biometric participant saw every notification in a randomized order in a within-subjects design. An overview of the results is given in Figure 6.

The notifications created varied impressions. As shown in Figure 6a, responses to "how secure would you feel using your fingerprint or face to create an account at *ExampleTech*?" varied across notifications (Friedman $\chi^2(4) = 196.6$, $p < .001$). Among participants, 51% reported they would feel "extremely secure" after viewing *Biometric-Leaves*; 50% reported the same after viewing *Biometric-Stored*. In contrast, fewer participants felt the same for *Biometric-Shared* (36%), *Biometric-Hacked* (33%), and *Biometric-Brands* (19%). *Biometric-Leaves* was rated as significantly more secure than *Biometric-Shared* ($p = .004$), *Biometric-Hacked* ($p < .001$), and *Biometric-Brands* ($p < .001$). Similarly, *Biometric-Stored* was rated as significantly more secure than *Biometric-Shared* ($p = .006$), *Biometric-Hacked* ($p < .001$), and *Biometric-Brands* ($p < .001$). Participants also felt significantly more secure for both *Biometric-Shared* ($p < .001$) and *Biometric-Hacked* ($p < .001$) compared to *Biometric-Brands*.

As shown in Figure 6b, responses to "how easy do you expect it to be to create an account at *ExampleTech*?" also varied across notifications (Friedman's $\chi^2(4) = 54.5$, $p < .001$). Participants felt that *Biometric-Brands* suggested account creation was more difficult than for the other four notifications (all $p \leq .047$). Whereas 40% of participants felt creating an account would be "extremely easy" after viewing *Biometric-Brands*, between 46% and 55% felt the same after viewing the other notifications.



(a) Notification suggests that creating an account with a biometric is **secure**.

(b) Notification suggests that creating an account with a biometric is **easy**.

(c) Participants' relative **ranking** of notifications.

Figure 6: Participants' relative perceptions and rankings of biometric WebAuthn notifications (within-subjects).

When asked to rank these 5 notifications (cf. Figure 6c), a plurality of participants (28%) ranked *Biometric-Leaves* first. Furthermore, 51% of participants ranked *Biometric-Leaves* first or second, and the smallest fraction (11%) ranked it last. Next best, *Biometric-Stored* was ranked first by 20% of participants, and either first or second by 48%. No other notification was ranked either first or second by more than 39% of participants. In contrast, 40% of participants ranked *Biometric-Brands* last. Differences in rankings were significant (Friedman $\chi^2(4) = 51.1$, $p < .001$). *Biometric-Leaves*, *Biometric-Stored*, and *Biometric-Shared* were ranked higher than *Biometric-Brands* (all $p < .001$) and *Biometric-Hacked* (all $p \leq .046$). *Biometric-Hacked* was itself ranked higher than *Biometric-Brands* ($p = .046$).

5.8 Choosing Biometric WebAuthn

Biometric WebAuthn Preferred Over Passwords: We asked participants to select types of websites on which they would "choose to use a biometric login. . . over a password." Only 5% of participants did not select any of the listed websites. In contrast, 87% indicated they would do so on banking websites, 62% for email, and between 48% and 56% for work, social media, shopping, and education websites. In contrast, when we asked a parallel question about non-biometric WebAuthn, 34% of participants did not select any of the listed websites. For each website, between 27% and 40% of participants would use non-biometric WebAuthn.

If given a choice between biometric sign-in and a password for *ExampleTech*, 66% of participants were "extremely likely" to choose biometric sign-in, while an additional 22% were "somewhat likely" to do so. Only 10% were "somewhat" or "extremely unlikely" to do so. This likelihood varied across groups (KW $\chi^2(7) = 20.3$, $p = .013$). Whereas 89% to 95% of participants in the biometric conditions were likely to do so, only 70% of *Non-biometric* and 69% of *Password* participants responded the same. These results suggest a single experience with biometric WebAuthn makes adoption more likely.

Biometrics Considered More Secure: While 75% of participants felt passwords were "slightly" or "much less secure" than biometric sign-in, only 10% felt passwords were "slightly" or "much more secure." Comparing fingerprint and face biometrics, 33% of participants felt they were equally secure, while 58% felt face was less secure than fingerprint. Comparing their non-biometric unlock mechanism to a site-specific password, 51% of participants felt their unlock mechanism was less secure than a site-specific password. This result was heavily influenced by the misconception that guessing the PIN or pattern was sufficient for gaining access. Note that with WebAuthn, physical access to the phone would also be required.

Website Trustworthiness: Due to misconceptions about biometric storage, we found a large gap between participants' willingness to register with biometric WebAuthn on trustworthy and untrustworthy websites. Whereas 86% of participants were "extremely" or "somewhat likely" to use biometric WebAuthn on trusted websites, only 24% answered the same for untrusted websites. Because only a site-specific public key is transferred, registering on a potentially untrustworthy website with WebAuthn does not actually put the user at risk. In fact, it is far safer to register at untrustworthy websites with WebAuthn rather than a (potentially reused) password.

6 Discussion

Key Misconceptions About Biometric WebAuthn: Our participants perceived biometric WebAuthn as more secure than passwords. However, we observed that participants tried to infer how this new authentication system worked based on their existing knowledge about passwords and phone unlocking. While some misconceptions participants expressed (e.g., availability concerns) are well-established in the literature, we identified new issues specific to biometric WebAuthn. The most urgent and salient misconception we identified in the context of WebAuthn is where users believe their biometrics are stored. The fraction of participants we observed reporting that their biometric data is sent to *ExampleTech* is alarming. Thus, when deploying WebAuthn with support for local biometric authentication, such as on mobile devices and some security keys, we urge services to address this misconception.

Our results show a clear usability advantage of biometric over non-biometric WebAuthn (PIN, pattern, or password). Our participants were surprised by how easy and fast the account creation and login process was. Due to the value participants placed on ease and speed of use, we also suggest services emphasize this when communicating with end users.

The notifications we tested to address misconceptions were impactful. In particular, *Biometric-Leaves* and *Biometric-Stored* increased the fraction of participants who correctly reported that their biometric is stored on the device from one-third up to one-half. Still, a single, brief notification is not enough to address this fundamental issue. Users will need more education on where their biometric is stored.

In regards to previous work by Lyastani et al. [32], Farke et al. [18], and Takakuwa [54] about using WebAuthn with hardware security keys, our work confirms their findings (recovery, lost authenticators, revocation, and mental models), but has also identified problems (storage, processing, transferability across multiple devices, and delegating access) that are specific to biometric WebAuthn and should be addressed to make the large-scale deployment of WebAuthn a success.

FIDO2/WebAuthn Implementation Issues: As already touched upon in previous work [40], we identified vendors' different implementations of the WebAuthn "Verify your identity" screen (cf. Figure 1b) to be a major usability hurdle. Depending on the OS version, hardware vendor, biometric sensor position (i.e., back of the phone or in-display), UI appearance settings (light or dark), and configured knowledge-based fallback scheme (i.e., PIN, pattern, or password) the interface is very different. Online services can influence the appearance as well by configuring an authenticator attachment selection criteria (i.e., no preference, platform, cross-platform) that allows user to select an external (roaming) authenticator such as a USB, NFC, or Bluetooth hardware security key or not.

Recommendations and Takeaways: The notifications our participants developed were centered around addressing misconceptions. Specific to biometric WebAuthn, we found that users' mental models differ substantially from other use cases on mobile devices, like phone unlocking. Our participants explicitly pointed out that they could unlock their phone with biometrics even without a data connection, so the biometric must be stored locally. In contrast, we observed that participants commonly applied their mental model for password-based sign-in, concluding that their biometric must be transmitted to the website as a password would. Compared to eBay's notification (cf. Figure 1a), which motivates users by focusing on the weaknesses of passwords, we recommend focusing on the convenience of WebAuthn instead.

Services implementing *biometric* WebAuthn should:

- Explicitly say that biometric data is not sent to, nor stored by, the website.
- Emphasize biometric authentication's speed and ease.
- Focus on WebAuthn's convenience, rather than comparing it to passwords.

Researchers should:

- Aim to solve impediments to adoption (e.g., transferability across devices [54], delegating access)
- Move beyond notifications and study richer interactions (e.g., short videos [32]) aiming to counteract lingering misconceptions (e.g., where biometrics are stored).

Limitations: Our studies have a number of limitations. Responses from our participants may suffer from a social desirability and response bias. To mitigate this, we did not explain that this was a study about usability or security, and we reminded our participants that people might have many different opinions. Like many human-subject studies, there is always the potential for a bias in question wording. To avoid this, we adhered to best practices [47] like keeping the questions short and clear, randomization, and piloting. All questionnaires we used can be found online [31]. As in previous work [32], we relied on a controllable artificial account setting, so our notifications have not been tested with real-world services. Most importantly, our studies are limited by our recruitment method using Prolific. Prior work [48] suggests that online studies about security and privacy behavior can approximate behaviors of populations well. Our studies are based on convenience samples, so they are inherently limited in their ecological validity. Participants were rather young, well-educated, and a fraction reported having an IT background. Most notably, due to our recruitment criteria, our study only included people that unlock their phone using biometrics, which might include a less privacy-concerned population and influence the responses towards a more positive perception of biometric authentication.

Future Work: Biometrics can be used for authentication on mobile devices in many contexts other than websites, including unlocking the device or within apps (e.g., online banking apps or password managers). We intentionally scoped our studies to web authentication and only relied on mobile devices as they offered the easiest-to-control study environment [20]. Nevertheless, users' mental models may differ between these contexts, and we consider studying the difference between biometric usage within apps compared to a website context important future work. Moreover, research on the design of warnings has shown that design factors like icons, colors, notification style, or choice (such as providing an alternative way for account creation) can significantly influence the notification's effectiveness. Promising UI designs like personalization or opinionated nudging could also be evaluated. We therefore consider it meaningful to explore further design patterns or richer interactions that more specifically focus on encouraging WebAuthn adoption.

7 Conclusion

In this work, we studied misconceptions of FIDO2 WebAuthn using biometrics. An online study with 42 crowdworkers revealed that 67% incorrectly thought their biometrics were transmitted to the website. In co-design sessions, we developed short-form notifications aiming to mitigate misconceptions surrounding WebAuthn. Participants focused on security, convenience, availability, and a comparison to (and emphasis of the drawbacks of) passwords. Via a 345-participant online study where we compared notifications, we found that

notifications that focus on the storage location of the biometric are most effective in counteracting misconceptions. Nevertheless, key misconceptions partially persisted. However, participants indicated high interest in adopting biometric WebAuthn over non-biometric WebAuthn and passwords.

Acknowledgments

We thank Anika Bansal and Olivia Morkved for study assistance, as well as Sven Bugiel for shepherding the paper. This research was partially funded by the MKW-NRW research training group *SecHuman* and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2092 *CASA* – 390781972.

Availability

An extended version of this work with screenshots of all notification designs, the full questionnaires, and more is available online [31]. Our FIDO2 implementation, based on Spomky-Labs' PHP WebAuthn Framework [38], is at: https://github.com/UChicagoSUPERgroup/fido2biometrics.

References

[1] Devdatta Akhawe and Adrienne Porter Felt. Alice in Warningland: A Large-Scale Field Study of Browser Security Warning Effectiveness. In *Proc. USENIX Security*, 2013.

[2] Aftab Alam, Katharina Krombholz, and Sven Bugiel. Poster: Let History Not Repeat Itself (This Time) – Tackling WebAuthn Developer Issues Early On. In *Proc. CCS*, 2019.

[3] Hazim Almuhimedi, Adrienne Porter Felt, Robert W. Reeder, and Sunny Consolvo. Your Reputation Precedes You: History, Reputation, and the Chrome Malware Warning. In *Proc. SOUPS*, 2014.

[4] Kholoud Althobaiti, Nicole Meng, and Kami Vaniea. I Don't Need an Expert! Making URL Phishing Features Human Comprehensible. In *Proc. CHI*, 2021.

[5] Dirk Balfanz, Alexei Czeskis, Jeff Hodges, J.C. Jones, Michael B. Jones, Akshay Kumar, Angelo Liao, Rolf Lindemann, and Emil Lundberg. Web Authentication: An API for Accessing Public Key Credentials – Level 1, March 2019. https://www.w3.org/TR/2019/REC-webauthn-1-20190304/, as of June 2, 2021.

[6] Lujo Bauer, Cristian Bravo-Lillo, Lorrie Cranor, and Elli Fragkaki. Warning Design Guidelines. Technical Report CMU-CyLab-13-002, Carnegie Mellon University, 2013.

[7] Chandrasekhar Bhagavatula, Blase Ur, Kevin Iacovino, Su Mon Kywey, Lorrie Cranor, and Marios Savvides. Biometric Authentication on iPhone and Android: Usability, Perceptions, and Influences on Adoption. In *Proc. USEC*, 2015.

[8] Christian Bravo-Lillo, Lorrie Cranor, Julie S. Downs, and Saranga Komanduri. Bridging the Gap in Computer Security Warnings: A Mental Model Approach. *Security & Privacy*, 9(2):18–26, 2011.

[9] Ivan Cherapau, Ildar Muslukhov, Nalin Asanka, and Konstantin Beznosov. On the Impact of Touch ID on iPhone Passcodes. In *Proc. SOUPS*, 2015.

[10] Chhaya Chouhan, Christy M. LaPerriere, Zaina Aljallad, Jess Kropczynski, Heather Lipford, and Pamela J. Wisniewski. Co-Designing for Community Oversight: Helping People Make Privacy and Security Decisions Together. In *Proc. CSCW*, 2019.

[11] 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 *Proc. CHI*, 2018.

[12] Heather Crawford and Karen Renaud. Understanding User Perceptions of Transparent Authentication on a Mobile Device. *Trust Management*, 1(1):1–28, 2014.

[13] Sanchari Das, Andrew Dingman, and L. Jean Camp. Why Johnny Doesn't Use Two Factor: A Two-Phase Usability Study of the FIDO U2F Security Key. In *Proc. FC*, 2018.

[14] Alexander De Luca, Alina Hang, Emanuel von Zezschwitz, and Heinrich Hussmann. I Feel Like I'm Taking Selfies All Day!: Towards Understanding Biometric Authentication on Smartphones. In *Proc. CHI*, 2015.

[15] Serge Egelman, Lorrie Cranor, and Jason Hong. You've Been Warned: An Empirical Study of the Effectiveness of Web Browser Phishing Warnings. In *Proc. CHI*, 2008.

[16] Serge Egelman and Stuart Schechter. The Importance of Being Earnest [in Security Warnings]. In *Proc. FC*, 2013.

[17] Rob Eisinga, Tom Heskes, Ben Pelzer, and Manfred Te Grotenhuis. Exact p-Values for Pairwise Comparison of Friedman Rank Sums, with Application to Comparing Classifiers. *BMC Bioinformatics*, 18(1):68:1–18, 2017.

[18] Florian M. Farke, Lennart Lorenz, Theodor Schnitzler, Philipp Markert, and Markus Dürmuth. "You still use

the password after all" – Exploring FIDO2 Security Keys in a Small Company. In *Proc. SOUPS*, 2020.

[19] Fast IDentity Online (FIDO) Alliance. How FIDO Works, August 2020. https://fidoalliance.org/how-fido-works/, as of June 2, 2021.

[20] Fast IDentity Online (FIDO) Alliance. Support for FIDO2: WebAuthn and CTAP – Browsers and Platforms, June 2020. https://fidoalliance.org/fido2/fido2-web-authentication-webauthn/, as of June 2, 2021.

[21] Adrienne Porter Felt, Alex Ainslie, Robert W. Reeder, Sunny Consolvo, Somas Thyagaraja, Alan Bettes, Helen Harris, and Jeff Grimes. Improving SSL Warnings: Comprehension and Adherence. In *Proc. CHI*, 2015.

[22] David Mandell Freeman, Sakshi Jain, Markus Dürmuth, Battista Biggio, and Giorgio Giacinto. Who Are You? A Statistical Approach to Measuring User Authenticity. In *Proc. NDSS*, 2016.

[23] Maximilian Golla, Miranda Wei, Juliette Hainline, Lydia Filipe, Markus Dürmuth, Elissa Redmiles, and Blase Ur. "What was that site doing with my Facebook password?" Designing Password-Reuse Notifications. In *Proc. CCS*, 2018.

[24] Lauren Goodegear. Our Password-Free Future Is Near (But Not Really), April 2018. https://www.wired.com/story/webauthn-in-browsers/, as of June 2, 2021.

[25] Peter Leo Gorski, Yasemin Acar, Luigi Lo Iacono, and Sascha Fahl. Listen to Developers! A Participatory Design Study on Security Warnings for Cryptographic APIs. In *Proc. CHI*, 2020.

[26] Eric Grosse and Mayank Upadhyay. Authentication at Scale. *Security & Privacy*, 11(1):15–22, 2013.

[27] Mark Hachman. Webauthn: What You Need to Know About the Future of the Passwordless Web, March 2019. https://www.pcworld.com/article/3355240/, as of June 2, 2021.

[28] Marian Harbach, Sascha Fahl, Polina Yakovleva, and Matthew Smith. Sorry, I Don't Get It: An Analysis of Warning Message Texts. In *Proc. Financial Cryptography and Data Security*, 2013.

[29] Kat Krol, Matthew Moroz, and M. Angela Sasse. Don't Work. Can't Work? Why It's Time to Rethink Security Warnings. In *Proc. CRiSIS*, 2012.

[30] Katharina Krombholz, Karoline Busse, Katharina Pfeffer, Matthew Smith, and Emanuel von Zezschwitz. "If HTTPS Were Secure, I Wouldn't Need 2FA" – End User and Administrator Mental Models of HTTPS. In *Proc. IEEE S&P*, 2019.

[31] Leona Lassak, Annika Hildebrandt, Maximilian Golla, and Blase Ur. "It's Stored, Hopefully, on an Encrypted Server": Mitigating Users' Misconceptions About FIDO2 Biometric WebAuthn (Extended Version), June 2021. https://www.blaseur.com/papers/fido2biometrics-extended.pdf, as of June 2, 2021.

[32] Sanam Ghorbani Lyastani, Michael Schilling, Michaela Neumayr, Michael Backes, and Sven Bugiel. Is FIDO2 the Kingslayer of User Authentication? A Comparative Usability Study of FIDO2 Passwordless Authentication. In *Proc. IEEE S&P*, 2020.

[33] Robbie MacGregor. Evaluating the Android Security Key Scheme: An Early Usability, Deployability, Security Evaluation with Comparative Analysis. In *Proc. WAY*, 2019.

[34] Liam M. Mayron. Biometric Authentication on Mobile Devices. *Security & Privacy*, 13(3):70–73, 2015.

[35] Brenna McNally, Priya Kumar, Chelsea Hordatt, Matthew Louis Mauriello, Shalmali Naik, Leyla Norooz, Alazandra Shorter, Evan Golub, and Allison Druin. Co-Designing Mobile Online Safety Applications with Children. In *Proc. CHI*, 2018.

[36] Weizhi Meng, Duncan S. Wong, Steven Furnell, and Jianying Zhou. Surveying the Development of Biometric User Authentication on Mobile Phones. *Communications Surveys & Tutorials*, 17(3):1268–1293, 2014.

[37] Grzergor Milka. Anatomy of Account Takeover. In *Proc. USENIX Enigma*, 2018.

[38] Florent Morselli (Spomky-Labs). PHP Webauthn Framework, October 2020. https://github.com/web-auth/webauthn-framework, as of June 2, 2021.

[39] Michael J. Muller and Allison Druin. *Participatory Design: The Third Space in HCI*, chapter 49, pages 1125–1153. Taylor & Francis, Boca Raton, Florida, USA, 3 edition, 2012.

[40] Wataru Oogami, Hidehito Gomi, Shuji Yamaguchi, Shota Yamanaka, and Tatsuru Higurashi. Poster: Observation Study on Usability Challenges for Fingerprint Authentication Using WebAuthn-enabled Android Smartphones. In *Proc. SOUPS*, 2020.

[41] Kentrell Owens, Olabode Anise, Amanda Krauss, and Blase Ur. User Perceptions of the Usability and Security of Smartphones as FIDO2 Roaming Authenticators. In *Proc. SOUPS*, 2021.

[42] Kentrell Owens, Blase Ur, and Olabode Anise. A Framework for Evaluating the Usability and Security of Smartphones as FIDO2 Roaming Authenticators. In *Proc. WAY*, 2020.

[43] Bijeeta Pal, Tal Daniel, Rahul Chatterjee, and Thomas Ristenpart. Beyond Credential Stuffing: Password Similarity Models using Neural Networks. In *Proc. IEEE S&P*, 2019.

[44] Sarah Pearman, Jeremy Thomas, Pardis Emami Naeini, Hana Habib, Lujo Bauer, Nicolas Christin, Lorrie Cranor, Serge Egelman, and Alain Forget. Let's Go in for a Closer Look: Observing Passwords in Their Natural Habitat. In *Proc. CCS*, 2017.

[45] Sarah Pearman, Shikun Aerin Zhang, Lujo Bauer, Nicolas Christin, and Lorrie Cranor. Why People (Don't) Use Password Managers Effectively. In *Proc. SOUPS*, 2019.

[46] Justin Petelka, Yixin Zou, and Florian Schaub. Put Your Warning Where Your Link Is: Improving and Evaluating Email Phishing Warnings. In *Proc. CHI*, 2019.

[47] Elissa M. Redmiles, Yasemin Acar, Sascha Fahl, and Michelle L. Mazurek. A Summary of Survey Methodology Best Practices for Security and Privacy Researchers. Technical Report CS-TR-5055, UM Computer Science Department, 2017.

[48] Elissa M. Redmiles, Sean Kross, and Michelle L. Mazurek. How Well Do My Results Generalize? Comparing Security and Privacy Survey Results from MTurk, Web, and Telephone Samples. In *Proc. IEEE S&P*, 2019.

[49] Elissa M. Redmiles, Everest Liu, and Michelle L. Mazurek. You Want Me To Do What? A Design Study of Two-Factor Authentication Messages. In *Proc. WAY*, 2017.

[50] Robert W. Reeder, Adrienne Porter Felt, Sunny Consolvo, Nathan Malkin, Christopher Thompson, and Serge Egelman. An Experience Sampling Study of User Reactions to Browser Warnings in the Field. In *Proc. CHI*, 2018.

[51] Ken Reese, Trevor Smith, Jonathan Dutson, Jonathan Armknecht, Jacob Cameron, and Kent Seamons. A Usability Study of Five Two-Factor Authentication Methods. In *Proc. SOUPS*, 2019.

[52] Andreas Sotirakopoulos, Kirstie Hawkey, and Konstantin Beznosov. On the Challenges in Usable Security Lab Studies: Lessons Learned from Replicating a Study on SSL Warnings. In *Proc. SOUPS*, 2011.

[53] Joshua Sunshine, Serge Egelman, Hazim Almuhimedi, Neha Atri, and Lorrie Cranor. Crying Wolf: An Empirical Study of SSL Warning Effectiveness. In *Proc. USENIX Security*, 2009.

[54] Alex Takakuwa. *Moving from Passwords to Authenticators*. PhD thesis, University of Washington, 2019.

[55] Joshua Tan, Lujo Bauer, Nicolas Christin, and Lorrie Cranor. Practical Recommendations for Stronger, More Usable Passwords Combining Minimum-Strength, Minimum-Length, and Blocklist Requirements. In *Proc. CCS*, 2020.

[56] Kurt Thomas, Jennifer Pullman, Kevin Yeo, Ananth Raghunathan, Patrick Gage Kelley, Luca Invernizzi, Borbala Benko, Tadek Pietraszek, Sarvar Patel, Dan Boneh, and Elie Bursztein. Protecting Accounts From Credential Stuffing With Password Breach Alerting. In *Proc. USENIX Security*, 2019.

[57] Shari Trewin, Cal Swart, Larry Koved, Jacquelyn Martino, Kapil Singh, and Shay Ben-David. Biometric Authentication on a Mobile Device: A Study of User Effort, Error and Task Disruption. In *Proc. ACSAC*, 2012.

[58] Blase Ur, Sean M. Segreti, Lujo Bauer, Nicolas Christin, Lorrie Cranor, Saranga Komanduri, Darya Kurilova, Michelle L. Mazurek, William Melicher, and Richard Shay. Measuring Real-World Accuracies and Biases in Modeling Password Guessability. In *Proc. USENIX Security*, 2015.

[59] Rick Wash, Emilee Radar, Ruthie Berman, and Zac Wellmer. Understanding Password Choices: How Frequently Entered Passwords are Re-used Across Websites. In *Proc. SOUPS*, 2016.

[60] Susanne Weber, Marian Harbach, and Matthew Smith. Participatory Design for Security-Related User Interfaces. In *Proc. USEC*, 2015.

[61] Daniel Lowe Wheeler. zxcvbn: Low-Budget Password Strength Estimation. In *Proc. USENIX Security*, 2016.

[62] Stephan Wiefling, Luigi Lo Iacono, and Markus Dürmuth. Is This Really You? An Empirical Study on Risk-Based Authentication Applied in the Wild. In *Proc. IFIP SEC*, 2019.

[63] Flynn Wolf, Ravi Kuber, and Adam J. Aviv. "Pretty Close to a Must-Have": Balancing Usability Desire and Security Concern in Biometric Adoption. In *Proc. CHI*, 2019.

[64] Yubico, Inc. What is WebAuthn?, August 2020. https://www.yubico.com/authentication-standards/webauthn/, as of June 2, 2021.