"It's time. Time for digital security.": An End User Study on Actionable Security and Privacy Advice

Anna Lena Rotthaler*, Harshini Sri Ramulu*, Lucy Simko†, Sascha Fahl‡, and Yasemin Acar§
*Paderborn University
†Barnard College
‡CISPA Helmholtz Center for Information Security
§Paderborn University & The George Washington University

***Abstract*—Digital security advice is the focus of much research, with unsatisfying results: End users do not follow experts' security advice, and users and experts struggle to prioritize existing advice. Several studies point out that users are overwhelmed by the amount of available security advice, and make recommendations on how to improve existing advice. Nevertheless, we still do not know how to effectively give security advice. Inspired by daily habit apps, we developed a set of 30 pieces of short and actionable advice, and the *Security App*, an Android smartphone app to provide this advice to end users, to reduce mental effort, and to build secure habits. We conducted a 30-day online end-user (N=74) study to evaluate whether the set of advice is actionable and meaningful to users, whether users adopt the advice, and whether the app has an impact on security awareness and behavior. Our results show that the app is an appropriate tool to provide security advice to end users. Participants perceive the majority of tasks as comprehensible, actionable, and useful, and we show that the app in fact introduces secure behaviors. Our results can serve as a basis for future research on security advice and creating secure habits, and the possibility to effectively teach secure behavior.**

1. Introduction

Advice sources on digital security and privacy are plenty, and include the internet, friends and family, or the newspaper [1]. The amount of available advice is overwhelming for users, and sometimes advice can be confusing or contradictory [2]. Thus, many users decide to trust their own assessment, or stop caring about security advice, because they think they are "not important enough to hack" [3], [4]. But even if users know how to behave securely online, they may not be able to avoid insecure behavior when reacting subconsciously [5]. In a world of stimulus overflow, we often have to react spontaneously, or react "by habit". Habit apps can help people to adopt new, "desired" behaviors, and they are very common in health, fitness, or meditation [6]–[8]. However, habits are rarely considered in digital behavior change interventions [9].

The usable security and privacy community has worked on digital security and privacy advice, investigating reasons for adoption and rejection of advice, and "what makes good security advice" [2], [3], [10]–[12]. However, adoption of experts' advice remains low, and experts also fail to prioritize advice [2]. Thus, we argue that a new approach is needed that implements recommendations for "good advice" from prior work, and presents digital security and privacy advice in a user-centric way. We therefore take expert-reviewed security and privacy advice, restate it to be actionable and feasible in 10 minutes a day, and deliver it to users in an app for a month, via the popular "habits app" concept.

Contributions. We developed an Android application with daily tasks that implement actionable security advice. In a 30-day online controlled quasi-experiment, we evaluated the tasks and the impact of app use on security awareness and behavior, answering the following research questions:

1) Is the advice perceived as (a) comprehensible, (b) actionable, and (c) useful?
2) Does the app improve security awareness and behavior?
3) Is the app a promising tool for communicating security advice?

We find that participants use the app, follow the advice, and improve their security behavior. We show that (a) the set of advice we created is perceived as comprehensible, actionable, and useful, and (b) our *Security App* helps communicate security and privacy advice and improve security awareness and behavior. We therefore recommend further exploring the delivery of security and privacy advice in this user-centric format.

2. Related Work

We discuss related work in three key areas: security advice, security behavior, behavior change, and habit apps, and usability, user behavior, and user experience. We also illustrate the novelty of our work.

2.1. Security Advice

Prior work found that security advice is largely stable over time [10], [13], but is often inconsistent or confusing [14], and difficult to prioritize [1], [3], [14]. In our study, we focus on the *design* of security advice.

Criteria for Good Advice. In 2008, MacGeorge et al. defined criteria for good security advice: it should be useful, comprehensible, relevant, and effective [15]. Additionally, it should be likely to be accomplished by the user and not include too many limitations and drawbacks. Later, Reeder et al. focused on similar criteria from public awareness communications: advice should be effective, actionable, consistent, and concise [14]. In 2020, Redmiles et al. developed a measurement approach for the quality of security advice [2], including the following criteria: perceived actionability (including confidence, time consumption, disruption, and difficulty), perceived efficacy, and comprehensibility. We use these criteria to scaffold and evaluate our presentation of advice.

Design Guidelines for (Security) Advice. Although researchers identified criteria for good security advice, users struggle to follow advice [2], [16], [17]. Research suggest improvements for security advice [3], [12]: Advice should be simplified and explained, it should be straightforward, should contain step-by-step instructions, and avoid technical jargon while clearly explaining why following the advice is beneficial for user security [3]. Advice must neither be inconsistent nor confusing [14]. Reeder et al. propose to obtain a balance between generic and specific advice. While generic advice covers a larger scope, it sometimes requires certain skills and judgment that end users lack [14]. Researchers also suggest improving the communication of security advice: features and functionality of tools should be better explained; data protection outside of privacy policies should be highlighted, and advice credibility should be transparent [3], [12].

Zou et al. propose that experts should identify the most effective and actionable advice [12]. Beautement et al. recommend taking the "compliance budget" into account when considering the amount of advice for particular users [18]. Therefore, the set of advice should be minimized, and highly valuable to users [3], [14].

Reeder et al. recommend customizing and tailoring advice to users by identifying particular audiences and customizing advice for each audience rather than communicating the same set of advice to everyone [14]. Redmiles et al. propose personalized security advice based on users' costs, risks, and values [19]. A "one-size-fits-all" approach for all users can even be disadvantageous. It could lead to significant market and individual losses. The right way to limit the amount of advice might not be clearly defined yet, but as Herley et al. stated in 2014: more is not the answer [11].

Importance and Prioritization. Herley et al. state, "the correct course of action is seldom as clear as we might like" [11]. It is, e.g., unclear which advice topics are important and how to prioritize them [20]. Redmiles et al. illustrated that users and experts struggle with advice prioritization [2]. Besides, end users and experts disagree on the advice they recommend [10]. Zou et al. suggest that experts identify the most effective and actionable advice from each area [12]. According to Herley, users value the importance of advice from an economic point of view [20]. Further, Schmüser et al. highlighted that prioritization of security advice becomes paramount during a crisis, and lack of prioritization can cause overwhelm [21]. Advice prioritization is a prevalent problem well documented in research—also from an advice curator's perspective [22].

Sources of Security Advice. Advice communication is critical for users to follow it [14]. Ion et al. suggested communicating advice differently, more effectively [10], and Rader et al. proposed using stories for advice communication [23].

Prior work has shown that users rely on several different sources when looking for security information. Websites are one of the most common sources of advice [1], [2]. Sources of security advice differ between users of different education levels: more educated people are more likely to report coworkers and government websites as sources [24]. Another essential aspect is the credibility of the source: While some research suggests users rely on computer security professionals, recommending that security advice clearly state the author's qualifications in computer security [3], other work indicates that users also rely on advice from strangers in online forums for advice, indicating non-experts' struggles with finding reliable advice sources [25], [26].

2.2. Security Behavior, Behavior Change, and Habit Apps

The human-in-the-loop framework describes the process for communicating computer security information to users as a set of six steps: attention switch, attention maintenance, comprehension, knowledge acquisition, knowledge retention, and knowledge transfer [27]. In this model, knowledge acquisition can be seen as a critical component of security behavior change [2].

Research into users' motivations and decision-making regarding computer security advice, illustrates that self-confidence influences user behavior more than actual knowledge and social considerations were trumped by individual perceptions [19], [28]–[31]. Users' decisions are rational from an economic point of view [11], [19], [20]. They consider risks and costs when deciding which advice to follow or not to follow, but they struggle to identify less obvious risks [19]. Users weigh benefits and risk perceptions and see the benefits of their decision higher [16].

Users' concerns are not always in line with their actual behavior. The privacy paradox describes that user-reported concerns regarding privacy often differ significantly from their actual behavior, which can, again, be explained, at least in part, by overwhelm [32].

Behavior Change. Behavior is an outcome of both conscious and non-conscious processes. To change their behavior, users should be consciously motivated and habits as non-conscious behavior should also be taken into account when designing behavior changes [9]. Habits are defined as learned impulses to perform a particular behavior, triggered outside of conscious awareness by specific cues [5], [33],

[34]. An automated response increases the likelihood that behavior will persist long-term, even when motivation fades and willpower is low [35]. To form new habits, repetition is needed, and prior work has provided repetitions of interventions for different durations ranging from less than four weeks [36], [37] to more than 24 weeks [38]. Lally et al. evaluated that it took 18 days for easy tasks like drinking more water and 254 days for complex tasks like regularly exercising [39]. Creating habits is a process, where starting small is key, and simplicity wins [40].

Digital behavior change interventions can support habit formation at different stages, e.g., by priming, reminding, rewarding, and self-monitoring. Next, we briefly discuss research on habits apps.

Habit Apps. Websites, software applications, and mobile devices can be used to change attitudes and behavior [41], especially for behavior relating to technology. There is a growing area of research for digital behavior change interventions (DBCIs) [7], [9]. Currently available habit formation apps are not grounded in habit literature and do not actually support habit formation [42]. The provision of information is the most common DBCI technique, but simply advising people to, e.g., exercise more, is ineffective [9]. There is no consensus on the best behavior change theory or model that should be used for DBCIs, and few interventions provide an in-depth understanding of the underlying mechanisms of behavior and attitude change they use [6], [9]. Nevertheless, there are beneficial effects of mobile health interventions on behavior change [6].

There are several design recommendations for the development of habit apps: reminders and streaks are cited as the most effective features [35], implementation intentions, a user-tailored design of interventions, e.g., with a pre-intervention survey [43], and a user tailored design of notifications are recommended [44]; the intervention frequency should be daily (as daily exposure drives engagement), and for the intervention content, both diversity and repetition matter [43]. Positive reinforcement supports the creation of habits, where rewards can be extrinsic (financial) or intrinsic (pleasure, satisfaction) [42]. It can also be provided through gamification elements, e.g., in the field of mobile health apps [45] but their application is often not driven by behavior change theory [46].

Reminders can be seen as a good start of a new habit, as the automaticity of the new behavior might develop faster than the decay of effectiveness of the reminder [42], [43]. In the long term, there is evidence that users depend on reminders and streaks to continue the desired behavior [35].

Prior work on habit apps achieved promising results for behavior change in physical activity [6], found that DBCIs can positively influence health behaviors [7], and evaluated mindfulness meditation apps finding that these apps mostly focused on guided meditation and giving design recommendations for future apps [8].

Pinder et al. conclude that a focus on habits is crucial to designing effective DBCIs, because there are multiple opportunities for pervasive computing technology to deliver interventions that can target both conscious and non-conscious processes [9]. Additionally, prior work shows that gamification and behavioral training tailored to ensure security and privacy can positively influence security and privacy awareness [47], [48]. All in all, habit apps seem promising to influence security behavior, our study offers empirical evidence for this intuition.

2.3. Usability, User Behavior & User Experience

Usability plays a vital role in usable security and privacy research and we measure usability of our *Security App* in this study with multiple measures. There are various attempts to measure the usability of a system. The most used questionnaire is the System Usability Scale (SUS) [49], [50]. Kortum and Sorber [51] and Kaya et al. [52] measured mobile app usability using SUS. Redmiles et al. proposed some criteria for suitable security advice that can be evaluated using Likert scale questions [2].

The Self-Report Habit Index (SRHI) measures habit strength and automaticity of behavior without measuring the behavioral frequency [53]. Measuring baseline pre-intervention and post-intervention behavior is essential to see whether there is a behavior change and whether this change is sustainable [9].

To minimize recall bias, we use telemetry combined with experience sampling (short surveys after task performance) for insights into participants' behavior and their reasoning for it [54]–[57].

We build on recommendations from prior work on how to change security advice: minimizing advice and empirically evaluating whether advice is implemented [2], [14], and changing its presentation to be more user-centered [10]. We implement these recommendations by designing short and actionable advice and presenting it in a new, modern, and effective way: as a mobile app. Thus, we are the first to combine research on actionable security advice with habit apps.

3. The Security App

In this section, we present our *Security App*, a smartphone application for end users that provides computer security and privacy advice in a 30-day challenge. The daily tasks are chunks of advice that can be implemented in 10 minutes or less. We developed the app for Android as it is the most commonly used smartphone operating system.[1] We developed the app using a Google Pixel 6 with Android 13.0, and tested it on multiple other devices. In the following sections, we describe the app features and the tasks presented in the *Security App* in detail.

3.1. App Features

When the user first opens the app, they are asked to sign in or register. After registration, the user gets a short

1. https://www.statista.com/statistics/272698/global-market-share-held-by-mobile-operating-systems-since-2009/

introduction to the app, describing its purpose and features. The home screen of the app shows the user whether they have a current task or have already completed a task, an overview of the study which consists of the study day, the number of completed tasks, and a detailed calendar which links to the respective completed tasks, as shown in Figure 1.

One task a day. The app is meant to be used for 30 days. During these 30 days, the user can complete one task per day. If the user misses a day, the task originally scheduled for the missed day will appear on the next day until the user engages with it.

Settings. The app provides a daily notification function that is preset to 6 p.m. and can be changed or turned off in the app settings. It says: "*It's time. Time for digital security.*" Besides, the user can access a list of all completed tasks and the introductory text about the app in the settings menu.

Completing tasks. To complete a task, the user has to choose one of the following options on the last page of the current task: "*I applied this security advice today.*" (success), "*I already applied this security advice before today.*" (always), "*I did not manage to apply this security advice.*" (fail), and "*I decided against applying this security advice.*" (reject). Each of the four options leads to a different survey which we describe in Section 4. After filling out the survey, the user is redirected to the home screen which then shows a green (success/always) or yellow (fail/reject) check mark. We track timestamps for these interactions.

Considerations. We developed the app using best practices for mobile design, as well as established nudges and gamification features. We also implemented pre-test feedback for UX decisions, e.g., determining the amount of notifications.

3.2. Tasks

In this section, we describe the genesis of the daily tasks, including our considerations and task choice, task format and order, and our thoughts on the recommended tools.

Considerations. Redmiles et al. identified 374 unique pieces of security advice [2]. These are very general, and partially contradictory pieces of advice.

Our goal was to work out 30 pieces of actionable advice that are feasible to do in about 10 minutes, where the user does not have to think about or look up how to implement the advice, or choose a certain tool. Kuo and Horn found that daily tasks are likely to give people feelings of satisfaction and drive engagement [43]. Taking into account the *compliance budget* which means that users have only limited time and resources for security behavior, we consider 10 minutes per day reasonable [18].

Kuo and Horn suggest to tailor interventions to the individual participant, e.g., by using a pre-intervention survey, which seems promising, but exceeds the scope of our work [43]. We designed the tasks for end users that primarily use an Android smartphone and a Windows computer, do not have computer security and privacy education, and feel frequently overwhelmed with security and privacy.

Task choice. We included six out of twelve topics in security advice identified by Redmiles et al. [2] as appropriate for short and actionable advice for our target group: *Browsers*, *Data Storage*, *Device Security*, *Passwords*, *Privacy*, and *Software*.

Three pieces of advice were considered most important by a majority of experts in a study of Ion et al.: "*installing system updates,*" "*using a password manager,*" and "*using two-factor authentication*" [10], and are also included in the complete list of advice worked out by Redmiles et al.[2] From the 374 pieces of advice identified by Redmiles et al. [2], we selected advice based on our professional expertise as researchers. We chose advice describing concrete action, e.g., "*use a password manager*" but not "*watch out for spam emails*" or "*be suspicious of links*". This "not-actionable" advice category is labeled "*Mindfulness*" by Ion et al. and excluded for our study [10]. We obtained 12 distinct pieces of advice ("*Use privacy settings,*" "*Install latest OS updates,*" "*Limit the amount of personal info being collected about you online,*" "*Lock your computer when you're away from it,*" "*Lock your smartphone with passcode or touch ID,*" "*Regularly back up your data,*" "*Update applications & Turn on automatic updates for applications,*" "*Use 2+ factor authentication,*" "*Use a password manager,*" "*Use ad blocker extensions,*" "*Use end-to-end encryption for communication & Verify who you are talking to,*" and "*Use unique/strong passwords*") that we split into 30 total 10-minute tasks, which necessitated actionability and specificity. We define a task as actionable if it is concrete, concise, effective, feasible to do in 10 minutes, and matches our technical requirements.

Task format. We present a concrete task in one sentence, following up with specific instructions for Android/Windows (which we required for all participants), splitting larger tasks into chunks that are achievable in 10 minutes. For interested users, we explain the reasoning and benefits behind the tasks on a single screen (drawing on [2]). We iterated over wording significantly with security experts and through a nine-participant pilot study to address actionability and specificity from others' perspectives. Figure 1 shows an example task presented in the *Security App*.

A complete list of the final set of advice can be found in Appendix A; details can be found in the replication package.[3]

Task order. Some tasks had to be done in order (e.g., first install a password manager, then store passwords in it). Twenty tasks, relating to the password manager, backup, and first and last tasks were fixed; the remaining ten tasks were randomized using balanced Latin squares.[4] Fixed tasks are marked in color in Appendix A.

Tools recommended. We gave instructions for Android and Windows devices, and for the most commonly used browsers. We recommended built-in tools like the Google

2. https://securityadvice.cs.umd.edu
3. https://securityapp-2024.netlify.app/
4. https://damienmasson.com/tools/latin_square/

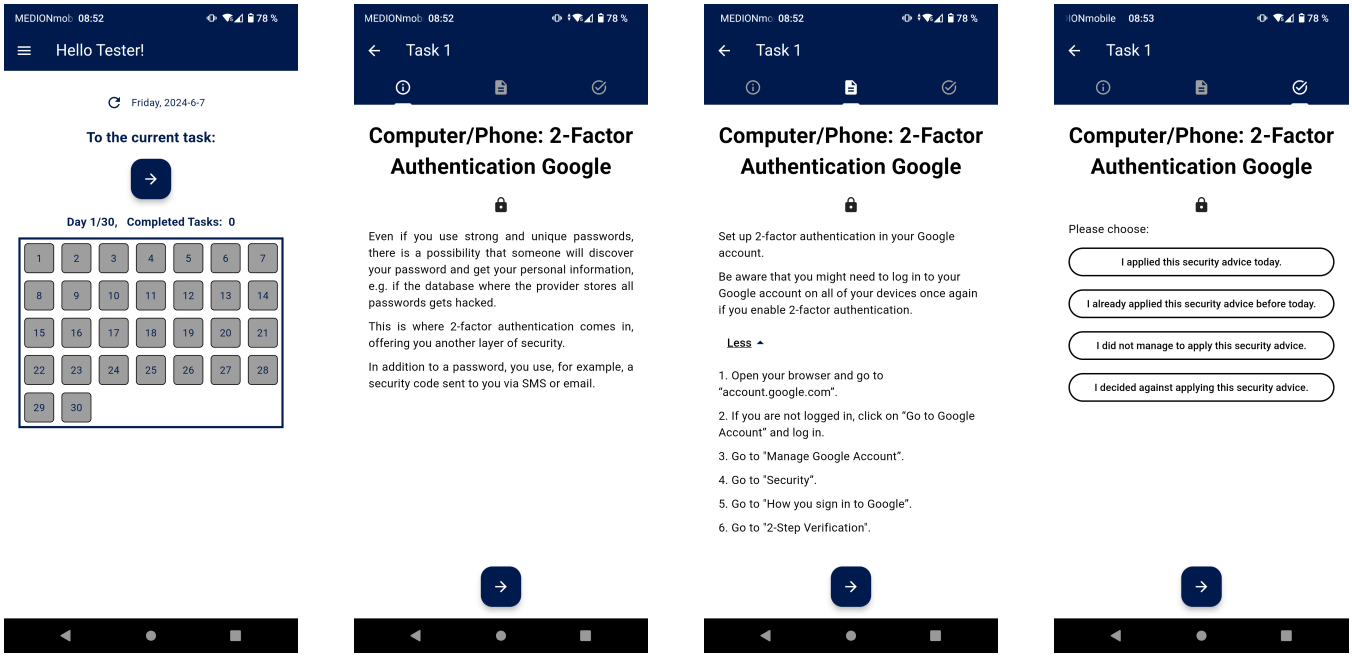

Figure 1. *Security App*: (1) home screen, (2) task information, (3) task instructions, (4) task completion.

password manager[5] and free tools like the uBlock Origin add-on[6]. Besides tool recommendations, we always made clear that it is fine to keep using whatever tool they already used.

Task content. The 30 tasks handle the topics: *Browsers* (2), *Data Storage* (4), *Device Security* (2), *Passwords* (13), *Privacy* (3), *Software* (5). The 30th and final task asks participants to plan their next steps and write them down with the help of guiding questions.

4. Longitudinal Study

We conducted a user study to answer the research questions and to evaluate the quality of the *Security App*.

4.1. Recruitment

Participants were recruited via Prolific and qualified if they were (1) over 18; (2) comfortable participating in English; (3) primarily using an Android smartphone and a Windows computer (desktop, laptop, or tablet), (4) no IT experts (no computer security and privacy education), and (5) willing to install new software on their devices, as well as to change passwords or account settings. We scoped recruitment to the USA and Germany, because our team felt comfortable in their knowledge of widely used technology and widely known computer security and privacy threats in both countries. Participants filled an eligibility survey about their operating systems and commonly used devices, their computer security and privacy education and experience, their mindset on computer security and privacy, and certain security behaviors during the last 30 days. Additionally, we asked some demographic questions including age, gender, place of residence, education, employment, and income to balance our sample.

After the pre-screen survey, participants were assigned to either the "app group" (three surveys over 60 days, 30-day app use), or "control group" (three surveys over 60 days, no app use). We can therefore assume that effects are caused by the treatment (app use) and not by repeatedly filling out surveys or external security-related notable events. We tried to balance participants in the two groups by country and gender.

Participants were paid £1 for the pre-screen survey, £1.50 for the first assessment survey that started the study, £5 for the second (after app use) survey, and £5 for the third survey (10 minutes each). Additionally, participants who used the app received £1.50 per completed task as bonus payment. Each participant could therefore receive a compensation of £57.50 (app group), and £12.50 (control group).

4.2. Study Design

The study consists of four parts: an *assessment survey* to start, 30 days of app usage including daily tasks and a one-minute *in-app survey* each day, a *completion survey* after the 30 days of app use, and a *final survey* 30 days after the *completion survey*. The study procedure is displayed in Figure 2. All surveys can be found in Appendix B.

Study Duration. Prior work has demonstrated that four weeks of repetition is a sufficient time period for automatic-

5. https://passwords.google.com/
6. https://ublockorigin.com/

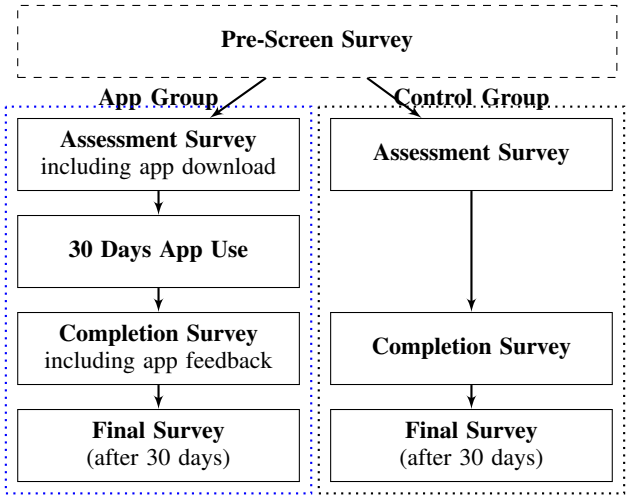

Figure 2. Study flow for "app group" participants (with app use) and "control group" participants (without app use).

ity of new tasks and habit formation [36], [42], and that new habits can be picked up through constant repetitions, sometimes around two weeks [58], and even 18 days for simpler tasks [39]; task repetitions depend on individuals and task complexity [39]. Regular reminders and actionability also matter [42]. We chose a 30-day time period for our study, in line with other behavioral studies [36], [37], [42].

Study Procedure. Participants who were eligible after filling out the *pre-screen survey* were divided into "app group" and "control group" and were then invited to the *assessment survey*. Once they filled out the *assessment survey*, participants in the "app group" received instructions on where they could download the *Security App* and how to install it. We provided the app through our research group's nextcloud.

After they created an account in the app, participants received an introduction to the study and the functions of the app. The day they registered in the app was the first of the 30 days they would use the app.

During these 30 days they were able to complete one task per day. To complete a task, participants had to fill out a one-minute questionnaire (*in-app survey*) with answer options *success, always, fail, reject* (described in Section 3.1). For an immediate impression, we measure the usability criteria for good security advice directly after participants engage with the task.

After the 30 days of app use, the participants were invited to the *completion survey*.

Participants were invited to the *final survey*, 30 days after they filled out the *completion survey*.

All surveys were conducted via Qualtrics. Most questions were required but we included a "do not want to disclose" option for demographic questions such as age and gender as proposed by Redmiles et al. [59]. Following expert recommendations for minimizing respondent fatigue, we used questionnaires that would take at most 10 minutes [1].

Before-and-after comparison. We chose existing measures to evaluate the impact of the app use on security awareness and behavior. Reusing existing, validated questions can be beneficial especially when dealing with topics that arise frequently in many security and privacy studies [59].

We measured baseline pre-intervention and post-intervention behavior to see whether there is a behavior change and whether this change is sustainable [9]. We used SA-13 [60], RSeBIS [30], and SRHI [53] for measurement of security awareness and behavior before and after the app use. We did not add "I don't know" or "I don't want to disclose" for these questionnaires because we could not have computed the scores and thus could not have compared them properly.

Assessment survey. The *assessment survey* includes the SA-13, RSeBIS, and SRHI. For the app group, it also includes the download instructions and link for the *Security App*.

Completion survey. In the *completion survey* the SA-13, RSeBIS, and SRHI were repeated to see whether security awareness and behavior changed with app use. Additionally, we asked questions concerning the tasks, the app, and the survey in general. Regarding the tasks, participants were asked which tasks they found helpful and not helpful and why. Regarding the app, the system usability scale (SUS) [61] was used. Besides, participants were asked to tell what they liked and disliked about the app, whether they would recommend the app to their friends and why, and whether they would use the Security App in their everyday life and why. Asking for recommendation to a friend allows us to compute the Net Promoter Score (NPS)[7], a customer experience metric that measures the participants' perception of the app. Regarding the study, participants were asked if they think participating in the study improved their computer security and privacy behavior and awareness, and whether they concurrently received any other security and privacy advice apart from this study. In addition, we repeated two questions from the pre-screen survey: whether they are overwhelmed with computer security and privacy (*security overwhelm*), and whether they wish they could improve their personal online behavior, but do not know how (*wish to improve*).

Final survey. In the *final survey*, the SA-13, RSeBIS, and SRHI were repeated again, to see if the daily app usage impacts security awareness and behavior in the long term. In addition, we again repeated the *security overwhelm* and *wish to improve* questions.

In-app survey. We wanted the daily *in-app surveys* to be as short as possible to minimize participants' burden, and take at most two minutes [54], [55].

In the *in-app surveys*, we wanted to measure the usability of the tasks, especially the following criteria for good security advice proposed by Redmiles et al. [2]: perceived comprehensibility, actionability, and usefulness.

No standard questionnaire worked to measure the usability of the tasks. Therefore, we created our own five questions about the comprehensibility and the difficulty of the task, the amount of information that was presented in the task,

7. https://www.qualtrics.com/uk/experience-management/customer/net-promoter-score/

whether they would recommend the activity to a friend, and whether the activity increases their security and privacy.

We created four different *in-app surveys*, according to the four options (success, always, fail, reject). If the participant applied the security advice (success or always), they were only shown the usability questions. If the participant failed to apply the security advice, they were additionally asked about their problem and for how long they had tried. If the participant decided against applying the security advice, they were additionally asked to describe the task and why they decided against doing it.

Pilot Study. After a lot of pre-testing, we conducted a pilot study with nine participants. All participants were students, used the app for 30 days, and completed on average 28 tasks. The evaluation revealed that overall, the tasks had a good comprehensibility, actionability, and a relatively high usefulness. The computed SUS scores ranged from 80 to 100, with a median of 90, and a mean of 89.2. Before and after comparisons revealed promising positive trends in the measures of security awareness and behavior.
Based on the pilot study, we adjusted the tasks presented in the app, and iterated for usability, readability, and comprehensibility. Additionally, we changed and replaced tasks after expert reviews or because they became outdated.

4.3. Data Analysis

Prior to data collection, we conducted a power analysis with G*Power[8] using the following parameters: ANOVA repeated measures, within-between interaction, effect size=0.2, alpha error=0.05, power=0.8, groups=2, measurements=3, correlation=0.3, which yielded a total sample size of 58 required participants; 74 participants total participated for the full study.

For the comparison of the SA-13, RSeBIS, and SRHI scores between the app and control group over time, we used a mixed ANOVA [62] with post-hoc t-tests using Benjamini-Hochberg correction for multiple testing [63].

For the in-app questions for task comprehensibility, actionability, and usefulness, we used descriptive statistics, giving mean and median values, as well as ranges. For free text responses, we group them by meaning through single-coder qualitative content analysis and report quotes when appropriate.

4.4. Limitations

This study has several limitations: For pre-existing scales SUS, SA-13, RSeBIS, and SRHI, we kept the original response options and did not offer the option to respond "I don't know" or "I do not want to disclose". This is required to calculate comparable scores, but participants may have chosen responses to the questions they did not understand. We only recruited participants from the US and Germany, and conducted the study only in English; we also included only Android and Windows users. We think expanding the app to be more inclusive would be crucial future work, however, we had to limit the scope in our study for comparability and feasibility.

Sometimes, people fast on technology, and our study did not consider this. Participants did not encounter any drawbacks from missing a day here and there and e.g., used to the "every day" app use of apps for language learning or meditation. Another limitation is that developing strong habits takes time, and participants only had 30 days, which is the minimum number of days needed to create habits. Although the *final survey* results 30 days after app use indicate that automaticity for dealing with computer security is higher than before the study, we cannot determine the sustainability of this habit, and this would require a longer study duration. Besides, we used money as the only *extrinsic* reward, so we are not certain about the influence of extrinsic rewards on the adoption of our tasks. However, we tried to minimize any effects by offering compensation even when participants rejected tasks. Determining what incentives would motivate app usage was out of scope; we implore future work to look into extrinsic motivators that would drive habitual security behavior and learning.

We also rely mostly on self-reported user data and cannot verify that participants completed the tasks as reported—which would have been an extreme privacy intrusion. However, their qualitative answers and the fact that they report failure, too, are consistent with honest behavior. Finally, we only considered relatively safe users and mostly gave advice for general security and privacy; future work may explore how to advise vulnerable populations or those facing crisis [64], [65].

4.5. Ethical Considerations

The study was approved by our institution's ethical review board. We adhered to ethical principles from the Menlo Report, as well as data protection guidelines. Participants were given a detailed consent form that informed them of the purpose of the study, study procedures, data handling, and their rights: Participation was voluntary, and participants were able to stop participation at any time. Demographic questions such as age and gender were optional, we otherwise did not collect personal information. The app required no additional authorization. The study data was de-identified, and only members of this research project had access to raw study data. Some of the tasks may contain risks such as access loss due to loss of the master password for the password manager, or loss of privacy due to sharing personal data with Google, but participants were able to decide against doing tasks without any harm or loss of compensation. We paid participants above local minimum wages, but not excessively, to strike the balance between fair compensation and recruiting those genuinely interested in the study. Beyond ethical treatment of participants, we think that gamified experiences may become addictive to users and may lead users to spend more time on them than they rationally would; we therefore limit our app to one task

8. https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower

a day. Clearly, a similar app could also be used to train users on misinformation; however, we think that our research does not facilitate this beyond the current state of habit-building apps.

5. Results

We describe the results of our study, describing participant demographics, that and how they used our app, how they found the tasks and app useful, and how app use improved security awareness and behavior.

5.1. Demographics

Of 699 participants who filled out the pre-screen survey, 138 were eligible, 97 filled out the assessment survey (52 app group, 45 control group), 47 registered in the app, 80 completed the completion survey (39 app group, 41 control group), and 74 completed the final survey; we report results for 74 full participants (36 app group, 38 control group).

Participants were aged between 19 and 73; 24 identified as men, 35 as women, two as non-binary, 12 preferred to self-describe (many of whom did not report a gender), and one preferred not to disclose. The distribution of the participants to the app group and control group can be seen in Table 2 in the Appendix.

5.2. App Usage

Task completion. The 36 participants of the app group used the *Security App* for 30 days and completed a total of 926 tasks; 26 completed tasks on average (range: 15-30). We define a task as *completed* if a participant filled out an in-app survey for it.

Of 926 times participants completed a task, they reportedly applied the advice successfully 476 times (51.4 %), already applied the advice before 346 times (37.37 %), did not manage to apply the advice 36 times (3.89 %), and decided against applying the advice 68 times (7.34 %).

On average, 31 participants completed each task (range nine to 36). Apart from task 29 (third, "update to more Secure Passwords"), which was fixed on day 29 and thus only completed by nine participants, all tasks were completed by at least 26 participants.

All but nine participants missed one or more tasks. We observe no pattern to missing tasks, neither within our study nor related to week days.

After completing a task, 27 participants looked up previously completed tasks at least once. One participant mentioned that they liked this feature: "*The ability to go back and check each task after you've finished it was very useful.*" (P14)

Advice Adoption. We define a task as *adopted* if the participant applied the security advice (option *success*) or had already applied the security advice (option *always*). Out of 926 tasks, 792 tasks were adopted (88.77%).

The task "back up photos from your phone" was the only task that all 36 participants completed and also adopted. Tasks that were adopted by all participants who completed the respective task are: "set up a phone unlock code" (29), "perform a privacy check in your Google account" (28), "update Windows" (27), "update your browser" (31), and "back up documents from your computer" (30).

Tasks with the lowest adoption rate are "install an ad blocker on your phone" (61 %), and "in your most used messenger, verify the security number of a contact" (69 %).

Participants adopted 128/130 *Data Storage* tasks (98 %), 141/145 *Software* tasks (97 %), 57/65 *Device Security* tasks (88 %), 75/85 *Privacy* tasks (88 %), 347/413 *Password* tasks (84 %), and 44/58 *Browser* tasks (75 %). We provide an overview of all the means of task adoption, completion, and usability in Table 3, in the Appendix.

Notifications. The default notification time was six p.m. (in any time zone). Most participants set notifications but changed the notification time. Nevertheless, participants completed tasks at various times of the day and around the clock.

Did they get any (other) advice. We asked participants of both groups if they received any (additional) security advice during the study period to discover any influences on scores and other data that were not caused by this study. Only seven participants in total mentioned other advice, namely: "*Google Password Help Section*" (P9), "*my companies' security platform*" (P16), "*AVAST*" (P38), "*a company I work with*" (P41), "*Search engines about VPN*" (P57), "*you tube*" (P63), and "*https://www.infosecurity-magazine.com/*" (P72).

5.3. App Evaluation

In this section, we present the results for general app usability, comprehensibility, actionability, and usefulness of the tasks, problems while applying advice, reasons for not applying advice, tasks participants found (un)helpful, and additional sources for help if needed. Overall, users perceive the app as usable and the tasks as comprehensible, actionable, and useful.

5.3.1. App Usability. We evaluate the app's usability through participants' responses to the System Usability Scale (SUS) as well as their responses to open-ended questions. Both describe the app as highly usable, as detailed below.

System Usability Scale (SUS): "very usable" on average. Scores on the SUS range from 42.5 to 100, with a median of 92.5, and a mean of 88.26. According to prior work [51], [52], the *Security App* is considered "very usable".

Participants liked that the Security App was easy to use. Participants gave free text responses that said that they liked the app itself because it was fast (1), without problems (1), easy to use (13), informative (1), useful (1), because of the reminders (3), the daily tasks (1), the concept of gamifying security practices (1), the good UI (1), and the simplicity of the app (7): "*I loved how easy and user friendly the app was to use. Everything was simple and worked exactly how*

I expected it to. The ability to go back and check each task after you've finished it was very useful." (P14)

Participants liked the tasks because they were easy to understand (2), not overwhelming (1), diverse (1), concise (2), provided straightforward, clear, step by step instructions (3): "*I liked that it gave you one task each day and that it divided the tasks, so it wouldn't be overwhelming. Like change 5 passwords today and change the other 5 passwords tomorrow.*" (P8)

Additionally, participants liked the effect of using the *Security App*. It helped them to learn about security and gain confidence (2), and made them feel safer (2): "*It made me feel a lot safer, it explained things in a very easy-to-understand way, it taught me a lot.*" (P6)

Participants encountered technical issues or constraints that impacted their experience. For some participants, notifications did not work properly (6). Other participants did not like that the app restarts after closing (1), that they can do only one task a day (2), and that they could not opt out of tasks, e.g., the password manager tasks (2): "*I didn't like that I couldn't just opt out of all the password manager stuff, that felt like a waste, as I wasn't going to use one and I'd have rather been trying/learning something else.*" (P21)

Participants mentioned that the tasks had too many steps (1), that they would have liked more information/articles in addition to the tasks (1), that they were not convinced the advice is good (1), and that there was no personalization (1): "*Personalization options would be cool, e.g., selecting the preferred browser and then only getting relevant tasks.*" (P9)

Two participants did not like the interface, and one participant mentioned that the app "*could be more motivational*" (P17).

Overall, three participants mentioned that they liked "nothing" about the app, while 19 participants found "nothing" that they did *not* like about the app.

Most participants are likely to recommend the Security App to their friends and colleagues. We asked participants how likely they were to recommend the *Security App* to a friend or colleague on a scale from 0 to 10. Responses ranged from 0 to 10 with a mean value of 7.583 and a median of 8. To compute the Net Promoter Score (NPS), responses are categorized into three categories: 14 participants are promoters (39 %, responses 9 and 10), 14 are passives (39 %, responses 7 and 8), and 8 are detractors (22 %, responses 0 to 6). This results in a NPS of 39 - 22 = 17. An NPS above 0 shows that more participants are satisfied with the app than not, and can be seen as a "good" score.[9] As a reason for their recommendation, several participants mentioned that the app is easy to use (6), informative and understandable (3), useful (5), and that it helped them a lot (4): "*I'd recommend it because it teaches one how to be safe online in a very easy and understandable way*" (P26) Other participants gave no specific reason (7). Reasons for not recommending the app were that it is boring (1), unnecessary (1), and that it only serves as a reminder for which they do not need an app (1).

9. https://www.qualtrics.com/uk/experience-management/customer/good-net-promoter-score/

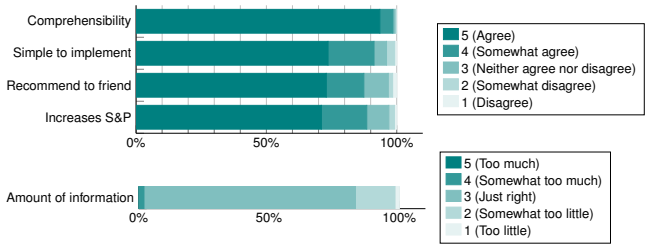

Figure 3. Participants' overall ratings of *comprehensibility*, *simple to implement*, *recommend to a friend*, *increases security and privacy (S&P)*, and *amount of information presented in the task*.

A majority of participants would like to use the Security App in their everyday lives. Out of 36 participants, 23 would like to use the app in their everyday lives because it provides very good security tips (3), it is easy to use (4), it helps to protect (11), it raises awareness (2), they think there is more to learn (2), or because it reminds them of security and privacy (5): "*An app that makes you aware (or reminds you) of what you should do to protect your data just makes sense.*" (P35)

Out of 36 participants, 13 would not like to use the app in their everyday life because it is too much effort (1), doing more would stress them out (1), they do not need it (2), they do not need reminders (1), it was not applicable (1). Other participants would prefer monthly reminders (1), would rather do it all in one go than over a month (1), do not want to go through all these different tasks (1). One participant mentioned that the app does not seem like it was designed for long-term use, and another participant states: "*It's kind of a waste of time now that I know what to do.*" (P13)

5.3.2. Comprehensibility of Tasks. We report our findings for task comprehensibility overall, and by task and topic. Participants largely found the tasks comprehensible, with software tasks most comprehensible and privacy tasks slightly less comprehensible.

The majority of completed tasks are perceived as comprehensible. Overall, participants agreed that they understood the tasks presented in the *Security App* (median: 5, mean: 4.912). Of 926 completed tasks, 883 completed tasks received a rating of four or five (98.55 %), seven completed tasks received a rating of three (0.78 %), and only six completed tasks received a rating of one or two (0.67 %). The overall ratings of *comprehensibility* are displayed in Figure 3.

Software tasks were most comprehensible. Four out of five software tasks received the best possible comprehensibility rating (mean: 5): "set up a phone unlock code", "update Android", "update Windows", "automatically update apps", and "update your browser". Three password related tasks also received a mean value of five ("save accounts in your password manager", "enable 2FA", "update to more secure passwords"), and one data storage task ("back up documents from your computer").

Privacy tasks were rated least comprehensible. Out of the six (of 926) times that participants rated slightly worse for comprehensibility, three were privacy related tasks: "perform a privacy check for your Google account" (mean: 4.82), "in your most used messenger, verify the security number of a contact" (mean: 4.5), and "check messenger privacy settings" (mean: 4.81). The others are "install a password manager on your computer" (mean: 4.72), and "back up other data from your computer" (mean: 4.86). Only two of these six times participants failed to do the task, the other four times they successfully adopted the task.

Overall, the difference between the topics' mean values of comprehensibility was rather small, with the lowest value of 4.72 for privacy and the highest value of 4.99 for software.

5.3.3. Actionability of Tasks. We report how *actionable* participants found the tasks, defining *actionable* as how simple the task is to implement and the amount of information presented in the task. Participants mostly found tasks actionable, with software and device security tasks rated as most actionable.

The majority of completed tasks are perceived as actionable. Overall, participants agreed that the tasks presented in the *Security App* are simple to implement (median: 5, mean: 4.60). Of 926 completed tasks, 818 completed tasks received a rating of four or five (91.3 %), 43 completed tasks received a rating of three (4.8 %), and 35 completed tasks received a rating of one or two (3.9 %).

Overall, participants agreed that the amount of information presented in the tasks is just right (median: 3, mean: 2.83). Of 926 completed tasks, participants rated the amount of information presented in the task as just right for 727 completed tasks (81.14 %, rating of 3), somewhat too much for 18 completed tasks (2.01 %, ratings of 4), and (somewhat) too little for 151 completed tasks (16.85 %, ratings of 1 and 2). The overall ratings of *simple to implement* and the *amount of information* are displayed in Figure 3.

Software and device security tasks were most actionable, privacy tasks were least actionable. "Automatically update apps" received the best rating for simple to implement (mean: 5) and for the amount of information (just right, mean: 3). "Update to more secure passwords" also received the best rating for simple to implement (mean: 5). "Update your browser" and "install an ad blocker on your computer" received the best rating for the amount of information (just right, mean: 3).

Tasks that were rated as (at least somewhat) simple to implement by all participants (ratings of 4 and 5) are "set up a phone unlock code" (mean: 4.97), "update android" (mean: 4.93), "update windows" (mean: 4.93), and "update your browser" (mean: 4.97).

The task that was least simple to implement was "in your most used messenger, verify the security number of a contact" (mean: 3.65), and its amount of information was considered somewhat too little (mean: 2.77), similar to "choose a password manager" (mean: 2.53).

"In your most used messenger, verify the security number of a contact" (mean: 3.65), "install a password manager" (mean: 4.11), "install an ad blocker on your phone" (mean: 4.11), and "update to more secure passwords" (mean: 4.27) received the most (low) ratings of one and two.

5.3.4. Usefulness of the Activity. In this section, we summarize the findings for task usefulness overall, and by task and topic, including if participants think the activity increases their security and privacy, and whether they would recommend the activity to a friend. Participants largely found tasks useful for increasing their security and privacy, especially updating passwords.

The majority of completed tasks are perceived as useful. Overall, participants agreed that the tasks presented in the *Security App* increase their security and privacy (median: 5, mean: 4.56). Of 926 completed tasks, 793 completed tasks received a usefulness rating of four or five (88.5 %), 76 completed tasks received a rating of three (8.48 %), and 27 completed tasks received a rating of one or two (3.02 %).

Overall, participants agreed that they would recommend the tasks presented in the *Security App* to their friends (median: 5, mean: 4.556). Of 926 completed tasks, 783 completed tasks received a friend recommendation rating of four or five (87.39 %), 84 completed tasks received a rating of three (9.38 %), and 29 completed tasks received a rating of one or two (3.23 %). The overall ratings of *increases security and privacy* and *recommend to a friend* are displayed in Figure 3.

Participants think that updating to more secure passwords increases their security and privacy, and they would recommend it to a friend. The task "update to more secure passwords" was presented three times during the 30 days, and received overall only ratings of four and five for increase of security and privacy (mean: 4.93, 4.96, and 5), and ratings of three to five for recommendation to a friend (mean: 4.7, 4.92, and 5).

"Set up a phone unlock code" (mean: 4.97) and "backup documents from your computer" (mean: 4.97) also received high ratings.

Password managers get worst ratings. The tasks with worst ratings for increase of security and privacy include "create a master password" (mean: 4.14), "choose a password manager" (mean: 4.25), "install a password manager on your computer" (mean: 4.28) and "install a password manager on your phone" (mean: 4.14). Besides password manager tasks, ratings of one and two were given to backup tasks (4), "automatically update apps" (1), ad blocker tasks (2), and privacy tasks (4).

Password manager tasks were also less likely to be recommended to a friend: "create a master password" (mean: 4.14), "choose a password manager" (mean: 4.25), and "install a password manager" (mean: 4.22). Ratings of one and two were also given to "set up 2FA for Google" (mean: 4.25), and "in your most used messenger, verify the security number of a contact" was least likely to be recommended to a friend (mean: 3.96).

Privacy tasks are considered least useful. *Privacy* tasks were rated worst for increasing security and privacy (mean: 4.36), and for recommendation to a friend (mean: 4.26).

Device security tasks were, on average, rated best for increasing security and privacy (mean: 4.88), and *data storage* tasks for recommendation to a friend (mean: 4.83). The only topic that received no ratings of one or two was *software* (mean: 4.62).

5.3.5. Problems while applying advice. Participants were unable to apply advice (option *fail*) 36 times in total (meaning that they were unable to complete the task), describing different problems with 16 different tasks. Participants chose this option 18 times because the activity described in the task did not apply to them, e.g., because "*I do not use either of those [messenger] apps*" (P13). Seven mentioned reasons to reject the task, e.g., "*I do not use a password manager*" (P27) or "*Going through the steps of setting it up seemed inconvenient at the moment.*" (P33). Other problems mentioned were a lack of time (4), a lack of money to buy an external hard drive (1), or they could not find the feature (4).

5.3.6. Reasons for not applying advice. Participants rejected to use a secure computer password because they do not feel the need (4), cannot remember a secure password (1), or need quick access to their computer (1): "*Nobody has access to my computer besides me.*" (P23)

Several password manager related tasks were rejected by 10 participants in total. They did not want to choose and set up a password manager, use a master password, and rejected follow up password manager tasks because they do not trust (free) password managers (3), they use an offline password manager (1), they do not feel the need (2), or they simply do not want to use a password manager (4): "*I don't like the idea of using a password manager.*" (P27)

Reasons for not changing passwords were effort (1), no time (1), they will maybe do it later (1), that it is not usable (1), or that they do not feel the need (2): "*I live alone so security is not an issue for me.*" (P19)

Automatically updating apps was rejected because participants want to remain in control (2): "*I want to select when and what apps are updated.*" (P33)

Participants did not want to set up 2FA because it is not usable (4), it is too time consuming (1), or they simply do not want it (3): "*I don't like this type of security. I've had it before and it was a hassle when my number changed.*" (P28)

Two participants did not want to install an ad blocker because they want to receive ads (2).

Messenger verification was seen as not usable (3), or unnecessary (1): "*I will know if I'm chatting with a stranger.*" (P17)

5.3.7. Which tasks were (not) rated helpful and why. We asked participants which tasks they did (not) find helpful and why and found mixed opinions for various tasks.

Just over half of the participants (18) found the password manager tasks helpful, including creating a master password because it makes managing passwords easier (8), helping change (duplicate) passwords (5), and generally helping to keep them safe (3): "*Password Manager, I wanted to implement this specific security feature for a long time, but was always overwhelmed with all the options. So this was great, because I was led through every step very clearly and understandably*". (P16) Others did not find the password manager tasks helpful because they do not trust password managers (3). One participant mentioned that the changing duplicate passwords tasks were not helpful "*because it made me realize I had a lot of duplicate passwords.*" (P24)

2FA was mentioned as a helpful task (2) because "*Google is my main account*" (P25), and as a not helpful task because it is not usable (4): "*it is extremely inconvenient and I have always hated it.*" (P23)

The phone ad blocker task was seen as helpful because it makes browsing convenient and enjoyable (3), and as not helpful because phone ad blockers are not available for Chrome (2).

Other tasks that were mentioned as helpful: backup (1) and automatic updates for software (3) because they otherwise are not doing it regularly, and checking the Google privacy settings (1). "*I didn't know that entire Google privacy panel existed and I was kind of shocked at how many apps were keeping lots of tabs on me, especially when many of them I don't even use (and have since deleted.)*" (P21)

Other tasks that were not seen as helpful: Whatsapp privacy settings (1), and secure computer lock (1) because the participants were already aware of it. One participant wrote down "*Most, because I already had a lot of them in place.*" (P22)

5.3.8. Where did they ask for help. We asked our app participants where they received help if they needed more help with the tasks than the app provided. Out of the 36 participants, only 5 participants mentioned that they received additional help from Google (3), or from spouses (2): "*I would first google and do some research to verify it is the right way to go.*" (P20)

5.4. Security Awareness and Behavior

We present the results of the statistical tests for the three questionnaires SRHI, RSeBIS, and SA-13, as well as the results of the self-assessment questions regarding security awareness and behavior. We compare differences between "app" and "control" group ("Group" variable) and over time ("time" variable), as well as interaction effects. We consider two group levels ("app" and "control") and three timepoints: assessment survey (t1), completion survey (t2), and final survey (t3).

5.4.1. SRHI. Mixed ANOVA shows a statistically significant interaction effect of "Group" and "time" on SRHI score: $F(2,144) = 6.368$, $p < 0.001$. Post-hoc analyses with Benjamini-Hochberg correction revealed statistically significant differences in SRHI score for the app group between t1 and t2 ($p_{adj} = 0.009$) and t1 and t3 ($p_{adj} = 0.009$). The app group was on average 27 % better after the study (increasing their mean SRHI score from 2.586 to 3.276).

5.4.2. RSeBIS. Mixed ANOVA shows a statistically significant interaction effect of "Group" and "time" on RSeBIS score: F(1.72,124.01) = 14.007, p $<$ 0.001, pes = 0.163. Post-hoc analyses with Benjamini-Hochberg correction revealed statistically significant differences in RSeBIS score for the app group between t1 and t2 ($p_{adj} < 0.001$) and t1 and t3 ($p_{adj} < 0.001$). The app group was on average 18 % better after the study (increasing their mean RSeBIS score from 3.222 to 3.804).

5.4.3. SA-13. Mixed ANOVA does not show statistically significant differences for SA-13 score. Differences for the interaction of "Group" and "time" closely failed significance (p = 0.062). Mixed ANOVA for the subscale SA-6 shows a statistically significant interaction effect of "Group" and "time": F(2,144) = 6.446, p = 0.002. pes = 0.082. Post-hoc analyses with Benjamini-Hochberg correction revealed a statistically significant difference within the app group (p $<$ 0.001). The app group was on average 11 % better after the study (increasing their mean SA-6 score from 3.245 to 3.606).

5.4.4. Self-assessment. To measure participants' self-assessment of their security and privacy awareness and behavior, we used the same questions before and after the study. We compare the results in this section.

Participants are less overwhelmed with computer security and privacy after participating in the study. On a five-point Likert scale asking if they are overwhelmed with computer security and privacy (1=Disagree, 5=Agree), participants of the app group reported a mean score of 3.833 in the pre-screen survey (3.784 for control group), and 2.194 in the final survey (2.459 for control group).

Additionally, we asked them to respond on a five-point Likert scale to the statement: "I wish I could improve my computer security and privacy, but I don't know how." (1=Disagree, 5=Agree). Participants in the app group reported a mean score of 4.222 in the pre-screen survey and a score of 2.111 in the final survey, whereas participants of the control group reported a mean score of 4.184 in the pre-screen survey and a score of 3.026 in the final survey.

The app group thinks participating in the study has improved their computer security and privacy awareness and behavior. Participants were asked if they think participating in the study has improved their computer security and privacy behavior, and they agreed with a mean score 4.314 (range: 1-5, median: 5). Only one participant disagreed, and one participant somewhat disagreed.

Besides, participants were asked if they think participating in the study has increased their awareness of computer security and privacy, and they agreed with a mean score 4.444 (range: 1-5, median: 5). Here as well, only one participant disagreed, and one participant somewhat disagreed.

App participants report more secure behavior after completing the study. We asked participants if they displayed certain behaviors during the last 30 days in the pre-screen survey and in the final survey. Their responses are summarized in Table 1. Additionally, we asked the app group for reasons to (not) continue these behaviors after the study.

None of the participants used a password manager when starting this study. During the study, 29 app participants and 12 control group participant started using a password manager.

App and control group performed similarly for using the same password for different accounts, updating software on their phone, updating software on their computer, and using 2FA for their Google account.

In the pre-screen survey, a similar number of participants reported having backed up documents from their computer as well as pictures from their phone during the last 30 days. After the study, the number of participants who backed up remained similar for the control group but increased for the app group.

Participants continued secure behavior because they want to feel/be more secure. Most participants mentioned general reasons for continuing behaviors, e.g. they want to feel/be more secure (14), out of habit (4), they already did that before the study (4), or it is convenient (5): "*I continued doing them because they felt much easier and quicker to do for me once I knew more about it.*" (P26)

One participant specifically mentioned that they checked for updates and updated some passwords, and another participant mentioned multi-factor authentication: "*I felt I needed to reduce the chances of being hacked and therefore need to include a more secure authentication approach with MFA.*" (P30)

Participants discontinued secure behavior because it is inconvenient. Participants gave general reasons for not continuing behaviors, e.g. time (3), too much effort (2), or inconvenience (5): "*I do what's convenient and I'm not that concerned with my security.*" (P23)

During the last 30 days I...	App Group (N=36)				Control Group (N=38)			
	y	n	idk	nd	y	n	idk	nd
...used a password manager								
before	0	35	1	0	0	35	3	0
after	29	7	0	0	12	23	3	0
...used the same password for different accounts								
before	31	4	0	1	31	4	1	2
after	21	15	0	0	23	14	1	0
...updated software on my phone								
before	29	4	3	0	33	2	3	0
after	32	4	0	0	34	4	0	0
...updated software on my computer								
before	23	11	2	0	26	9	3	0
after	30	5	1	0	32	6	0	0
backed up documents from my computer								
before	13	20	3	0	15	19	4	0
after	27	9	0	0	16	18	4	0
...backed up pictures from my phone								
before	19	16	1	0	20	18	0	0
after	27	9	0	0	20	16	2	0
...used 2FA for my Google account								
before	17	16	3	0	15	15	8	0
after	26	8	2	0	24	14	0	0

TABLE 1. PARTICIPANT BEHAVIORS BEFORE AND AFTER THE STUDY.

6. Discussion

From our development of the *Security App* and controlled quasi-experiment with 74 participants, we learn both about the efficacy of the *Security App* itself, and about new strategies for security and privacy education and calls-to-action for users. Our work is especially poignant because of how it addresses one key issue in user action/education: prioritization of advice and a feeling of overwhelm amongst lay-users [2], [3].

The *Security App* is a promising approach for improving users' security and privacy. Our results quantitatively show that, against a control group, the *Security App* statistically significantly increased self-reported behaviors that align with expert recommendations and greater resilience to certain common threats (e.g., exploitation due to unpatched systems) and issues (e.g., password reuse), and that users showed some maintenance of these behaviors 30 days after the study's completion. Interestingly, our control group, with no intervention except filling in surveys about their security behavior, also started adopting some security measures throughout the study, to a lesser degree than the treatment group. Prior work has identified an overwhelming cacophony of security advice that is difficult for laypeople to prioritize and implement [2], [3], in part responsible for the gap between the privacy users *say* they want and how they protect their assets [32]. Our promising approach guides laypeople *making security and privacy less overwhelming* by (a) curating bite-sized pieces of security-positive actions, and (b) building security as a habit. While there is certainly room for improvement, participants completed most of the tasks (Section 5.2), and reported that most of the tasks were highly comprehensible, actionable, useful, and said that they would like to use the *Security App* daily and would recommend it to friends and colleagues (Section 5.3).

The *Security App* did not succeed equally at helping users adopt all security behaviors. Despite the general success of the *Security App* in encouraging users to adopt security and privacy-positive behaviors, users failed or declined to adopt a few tasks at a higher rate than others, or reported that some tasks were less actionable or comprehensible than others. Many of these can be resolved with updated instructions or design of the app, but these may also reflect gaps in user education about *why* these behaviors are necessary, or gaps between the behaviors promoted by the *Security App* and its users' needs. One of the greatest challenges in designing the *Security App* was choosing the tasks themselves, given the enormous amount of evidence that security and privacy advice is *not* universally applicable [2], [14]. Thus, future iterations of the app will consider updated designs for the tasks identified as problematic, and will consider other user feedback, such as personalization and the inability to skip tasks, all of which we hypothesize will help address current issues. Future work could also more deeply help users self-evaluate (perhaps as a series of tasks) whether security advice advice fits their lives.

The *Security App* both echoes findings from prior work about persistent non-adoption of security behaviors, and shows that our approach leads *some* people to new adoption of security behaviors. A significant body of prior work has explored why people fail or decline to adopt password managers, finding concerns about password managers securely protecting their passwords [17], [66]–[68], being unusable [69], [70] perceiving them as unnecessary [67], [68], or insufficient awareness [68]. Our results echo these reasons, with the "choose a password manager" task being rated as having too little information (unusability), increasing security and privacy the least (unnecessariness and security concerns) and simply not something participants were willing to do. Participants gave similar reasons for rejecting the 2FA task and the computer lock task. However, the number of people who *did* adopt password managers (and other behaviors) in the study period reveals that a subset of non-adopters can be nudged into adoption of security-positive behaviors though our approach. These results also remind us that security advice cannot be universally applicable, and a task success rate of less than 100% is not a failure: we cannot know if not adopting a password manager in the way our app directed was actually the right choice for some of the participants. While password managers will likely increase the security and usability of most users' password management, the design of password managers themselves is not universally applicable, and does not fit all users [65], [71]. Compared to existing research on habit apps in general that either focused on the design and usability of habit apps [72], [73], analysis of existing habit apps [8], [74], [75], or measured only pre- and post-intervention but not app engagement [76], [77], our study provides detailed insights on the app usage and shows that participants actually engaged with the app on an (almost) daily basis.

We can actually give actionable security advice. While the security and privacy field generally struggles to prioritize advice [2], and while we also felt that we could easily have prioritized tasks differently, we ended up with a best-effort on giving advice—and this approach worked reasonably well. Our results show that our set of advice is overall perceived as comprehensible, actionable, and useful (RQ1), that using the app improved security awareness and behavior (RQ2), and that the app is a promising tool for communicating security advice (RQ3). Within 30 days, participants did make substantial changes. We argue that it may very well be possible to apply similar best-effort approaches to other difficult behavior change areas.

Future directions. For future work, we think it can be impactful to continue to practice "translating" security knowledge from research into actionable tasks, similar to how we explored this in our study, and share this strategy with security researchers. We think that making the app customizable for different technology, user groups, and context can be a promising direction. Additionally, we would like to explore whether the use of stronger gamification features will allow the app to help users be more secure without financial incentives. Future work can determine the real

world contexts in which the *Security App* concept can be used. Further work is needed to investigate how extrinsic motivators like gamification, social recognition, rewards, etc., might encourage continued app usage and learning; prior work on health habit apps has shown regular app use without monetary incentive [37]. Finally, the security and privacy community lacks effective methods to deliver security education; the *Security App* is only *one* possible implementation of habit-forming security and privacy training. Our study shows that this type of training *may* work more generally and *did* work with our specific design. We encourage future work in this direction.

7. Conclusion

We built the *Security App*, an Android app that gives actionable, comprehensible security advice that users implement over 30 days in 10 minutes a day or less. Our evaluation showed the success of this approach—participants implemented the advice, enjoyed the app, and improved on several security-relevant scales.

We hope that this work demonstrates that human-centered best effort can be impactful in spreading security behavior, and hope to expand this approach by tailoring it to users, considering higher-stakes environments (such as with vulnerable populations), and making the app widely available.

Acknowledgments

We would like to thank all of our participants for participating in our study. We also thank the anonymous reviewers and shepherd for their valuable feedback. We thank Ronja Rotthaler for her helpful feedback on earlier versions of this paper.

References

[1] E. M. Redmiles, S. Kross, and M. L. Mazurek, "How i learned to be secure: A census-representative survey of security advice sources and behavior," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, 2016, pp. 666–677.

[2] E. M. Redmiles, N. Warford, A. Jayanti, *et al.*, "A comprehensive quality evaluation of security and privacy advice on the web," in *29th USENIX Security Symposium (USENIX Security 20)*, 2020, pp. 89–108.

[3] E. M. Redmiles, A. R. Malone, and M. L. Mazurek, "I think they're trying to tell me something: Advice sources and selection for digital security," in *2016 IEEE Symposium on Security and Privacy (SP)*, IEEE, 2016, pp. 272–288.

[4] A. E. Howe, I. Ray, M. Roberts, M. Urbanska, and Z. Byrne, "The psychology of security for the home computer user," in *2012 IEEE Symposium on Security and Privacy*, IEEE, 2012, pp. 209–223.

[5] B. Verplanken and H. Aarts, "Habit, attitude, and planned behaviour: Is habit an empty construct or an interesting case of goal-directed automaticity?" *European review of social psychology*, vol. 10, no. 1, pp. 101–134, 1999.

[6] A. Domin, D. Spruijt-Metz, D. Theisen, Y. Ouzzahra, C. Vögele, *et al.*, "Smartphone-based interventions for physical activity promotion: Scoping review of the evidence over the last 10 years," *JMIR mHealth and uHealth*, vol. 9, no. 7, e24308, 2021.

[7] K. J. Thomas Craig, L. C. Morgan, C.-H. Chen, *et al.*, "Systematic review of context-aware digital behavior change interventions to improve health," *Translational behavioral medicine*, vol. 11, no. 5, pp. 1037–1048, 2021.

[8] C. Daudén Roquet and C. Sas, "Evaluating mindfulness meditation apps," in *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–6.

[9] C. Pinder, J. Vermeulen, B. R. Cowan, and R. Beale, "Digital behaviour change interventions to break and form habits," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 25, no. 3, pp. 1–66, 2018.

[10] I. Ion, R. Reeder, and S. Consolvo, ""... no one can hack my mind": Comparing expert and non-expert security practices," in *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*, 2015, pp. 327–346.

[11] C. Herley, "More is not the answer," *IEEE Security & Privacy*, vol. 12, no. 1, pp. 14–19, 2013.

[12] Y. Zou, K. Roundy, A. Tamersoy, S. Shintre, J. Roturier, and F. Schaub, "Examining the adoption and abandonment of security, privacy, and identity theft protection practices," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–15.

[13] K. Busse, J. Schäfer, and M. Smith, "Replication: No one can hack my mind revisiting a study on expert and non-expert security practices and advice," in *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*, 2019, pp. 117–136.

[14] R. W. Reeder, I. Ion, and S. Consolvo, "152 simple steps to stay safe online: Security advice for non-tech-savvy users," *IEEE Security & Privacy*, vol. 15, no. 5, pp. 55–64, 2017.

[15] E. L. MacGeorge, B. Feng, and E. R. Thompson, ""good" and "bad" advice," *Studies in applied interpersonal communication*, vol. 145, 2008.

[16] M. Fagan and M. M. H. Khan, "Why do they do what they do?: A study of what motivates users to (not) follow computer security advice," in *Twelfth symposium on usable privacy and security (SOUPS 2016)*, 2016, pp. 59–75.

[17] H. Ray, F. Wolf, R. Kuber, and A. J. Aviv, "Why older adults (don't) use password managers," in *30th USENIX Security Symposium (USENIX Security 21)*, 2021, pp. 73–90.

[18] A. Beautement, M. A. Sasse, and M. Wonham, "The compliance budget: Managing security behaviour in organisations," in *Proceedings of the 2008 New Security Paradigms Workshop*, 2008, pp. 47–58.

[19] E. M. Redmiles, M. L. Mazurek, and J. P. Dickerson, "Dancing pigs or externalities? measuring the rationality of security decisions," in *Proceedings of the 2018 ACM Conference on Economics and Computation*, 2018, pp. 215–232.

[20] C. Herley, "So long, and no thanks for the externalities: The rational rejection of security advice by users," in *Proceedings of the 2009 workshop on New security paradigms workshop*, 2009, pp. 133–144.

[21] J. Schmüser, H. Sri Ramulu, N. Wöhler, *et al.*, "Analyzing security and privacy advice during the 2022 russian invasion of ukraine on twitter," in *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 2024, pp. 1–16.

[22] L. Neil, H. S. Ramulu, Y. Acar, and B. Reaves, "Who comes up with this stuff? interviewing authors to understand how they produce security advice," in *Nineteenth Symposium on Usable Privacy and Security (SOUPS 2023)*, 2023, pp. 283–299.

[23] E. Rader, R. Wash, and B. Brooks, "Stories as informal lessons about security," in *Proceedings of the Eighth Symposium on Usable Privacy and Security*, 2012, pp. 1–17.

[24] E. M. Redmiles, S. Kross, and M. L. Mazurek, "Where is the digital divide? a survey of security, privacy, and socioeconomics," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, pp. 931–936.

[25] A. A. Hasegawa, N. Yamashita, T. Mori, D. Inoue, and M. Akiyama, "Understanding non-experts' security-and privacy-related questions on a q&a site," in *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*, 2022, pp. 39–56.

[26] M. Tran, C. W. Munyendo, H. S. Ramulu, *et al*., "Security, privacy, and data-sharing trade-offs when moving to the united states: Insights from a qualitative study," in *2024 IEEE Symposium on Security and Privacy (SP)*, IEEE Computer Society, 2023, pp. 4–4.

[27] L. F. Cranor, "A framework for reasoning about the human in the loop," 2008.

[28] J. Nicholson, L. Coventry, and P. Briggs, ""if it's important it will be a headline" cybersecurity information seeking in older adults," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–11.

[29] R. Wash and E. Rader, "Too much knowledge? security beliefs and protective behaviors among united states internet users," in *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*, 2015, pp. 309–325.

[30] Y. Sawaya, M. Sharif, N. Christin, A. Kubota, A. Nakarai, and A. Yamada, "Self-confidence trumps knowledge: A cross-cultural study of security behavior," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, pp. 2202–2214.

[31] M. Fagan and M. M. H. Khan, "To follow or not to follow: A study of user motivations around cybersecurity advice," *IEEE Internet Computing*, vol. 22, no. 5, pp. 25–34, 2018.

[32] S. Barth and M. D. De Jong, "The privacy paradox–investigating discrepancies between expressed privacy concerns and actual online behavior–a systematic literature review," *Telematics and informatics*, vol. 34, no. 7, pp. 1038–1058, 2017.

[33] S. Orbell and B. Verplanken, "The automatic component of habit in health behavior: Habit as cue-contingent automaticity.," *Health psychology*, vol. 29, no. 4, p. 374, 2010.

[34] B. Verplanken, B. Verplanken, and Ryan, *Psychology of habit*. Springer, 2018.

[35] I. Renfree, D. Harrison, P. Marshall, K. Stawarz, and A. Cox, "Don't kick the habit: The role of dependency in habit formation apps," in *Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems*, 2016, pp. 2932–2939.

[36] J. Fiedler, C. Seiferth, T. Eckert, A. Woll, and K. Wunsch, "A just-in-time adaptive intervention to enhance physical activity in the smartfamily2. 0 trial.," *Sport, Exercise, and Performance Psychology*, vol. 12, no. 1, p. 43, 2023.

[37] M. Rabbi, A. Pfammatter, M. Zhang, B. Spring, T. Choudhury, *et al*., "Automated personalized feedback for physical activity and dietary behavior change with mobile phones: A randomized controlled trial on adults," *JMIR mHealth and uHealth*, vol. 3, no. 2, e4160, 2015.

[38] C. Höchsmann, D. Infanger, C. Klenk, K. Königstein, S. P. Walz, A. Schmidt-Trucksäss, *et al*., "Effectiveness of a behavior change technique–based smartphone game to improve intrinsic motivation and physical activity adherence in patients with type 2 diabetes: Randomized controlled trial," *JMIR serious games*, vol. 7, no. 1, e11444, 2019.

[39] P. Lally, C. H. Van Jaarsveld, H. W. Potts, and J. Wardle, "How are habits formed: Modelling habit formation in the real world," *European journal of social psychology*, vol. 40, no. 6, pp. 998–1009, 2010.

[40] B. J. Fogg, *Tiny habits: The small changes that change everything*. Eamon Dolan Books, 2019.

[41] B. J. Fogg, "Persuasive technology: Using computers to change what we think and do," *Ubiquity*, vol. 2002, no. December, p. 2, 2002.

[42] K. Stawarz, A. L. Cox, and A. Blandford, "Beyond self-tracking and reminders: Designing smartphone apps that support habit formation," in *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, 2015, pp. 2653–2662.

[43] P.-Y. Kuo and M. S. Horn, "Daily challenges for sustainable lifestyles: Design implications from a mobile intervention study," in *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, 2017, pp. 635–641.

[44] C. Pinder, J. Vermeulen, A. Wicaksono, R. Beale, and R. J. Hendley, "If this, then habit: Exploring context-aware implementation intentions on smartphones," in *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, 2016, pp. 690–697.

[45] M. Schmidt-Kraepelin, P. A. Toussaint, S. Thiebes, J. Hamari, A. Sunyaev, *et al*., "Archetypes of gamification: Analysis of mhealth apps," *JMIR mHealth and uHealth*, vol. 8, no. 10, e19280, 2020.

[46] V. W. S. Cheng, T. Davenport, D. Johnson, K. Vella, and I. B. Hickie, "Gamification in apps and technologies for improving mental health and well-being: Systematic review," *JMIR mental health*, vol. 6, no. 6, e13717, 2019.

[47] A. Ganesh, C. Ndulue, and R. Orji, "Tailoring a persuasive game to promote secure smartphone behaviour," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 2023, pp. 1–18.

[48] N. Gerber, P. Gerber, H. Drews, *et al*., "Foxit: Enhancing mobile users' privacy behavior by increasing knowledge and awareness," in *Proceedings of the 7th Workshop on Socio-Technical Aspects in Security and Trust*, 2018, pp. 53–63.

[49] R. A. Grier, A. Bangor, P. Kortum, and S. C. Peres, "The system usability scale: Beyond standard usability testing," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, SAGE Publications Sage CA: Los Angeles, CA, vol. 57, 2013, pp. 187–191.

[50] J. R. Lewis, "The system usability scale: Past, present, and future," *International Journal of Human–Computer Interaction*, vol. 34, no. 7, pp. 577–590, 2018.

[51] P. Kortum and M. Sorber, "Measuring the usability of mobile applications for phones and tablets," *International Journal of Human-Computer Interaction*, vol. 31, no. 8, pp. 518–529, 2015.

[52] A. Kaya, R. Ozturk, and C. Altin Gumussoy, "Usability measurement of mobile applications with system usability scale (sus)," in *Industrial engineering in the big data era*, Springer, 2019, pp. 389–400.

[53] B. Verplanken and S. Orbell, "Reflections on past behavior: A self-report index of habit strength 1," *Journal of applied social psychology*, vol. 33, no. 6, pp. 1313–1330, 2003.

[54] R. W. Reeder, A. P. Felt, S. Consolvo, N. Malkin, C. Thompson, and S. Egelman, "An experience sampling study of user reactions to browser warnings in the field," in *Proceedings of the 2018 CHI conference on human factors in computing systems*, 2018, pp. 1–13.

[55] I. Myin-Germeys, Z. Kasanova, T. Vaessen, *et al*., "Experience sampling methodology in mental health research: New insights and technical developments," *World Psychiatry*, vol. 17, no. 2, pp. 123–132, 2018.

[56] N. Van Berkel, D. Ferreira, and V. Kostakos, "The experience sampling method on mobile devices," *ACM Computing Surveys (CSUR)*, vol. 50, no. 6, pp. 1–40, 2017.

[57] G. Eisele, H. Vachon, G. Lafit, *et al*., "The effects of sampling frequency and questionnaire length on perceived burden, compliance, and careless responding in experience sampling data in a student population," *Assessment*, vol. 29, no. 2, pp. 136–151, 2022.

[58] B. Gardner, K. Sheals, J. Wardle, and L. McGowan, "Putting habit into practice, and practice into habit: A process evaluation and exploration of the acceptability of a habit-based dietary behaviour change intervention," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 11, pp. 1–13, 2014.

[59] E. M. Redmiles, Y. Acar, S. Fahl, and M. L. Mazurek, "A summary of survey methodology best practices for security and privacy researchers," Tech. Rep., 2017.

[60] C. Faklaris, L. Dabbish, and J. I. Hong, "Do they accept or resist cybersecurity measures? development and validation of the 13-item security attitude inventory (sa-13)," *arXiv preprint arXiv:2204.03114*, 2022.

[61] U. G. S. Administration, *System usability scale (sus)*, https://digital.gov/2014/08/29/system-usability-scale-improving-products-since-1986/.

[62] B. Walther, *Statistik mit R Schnelleinstieg: R einfach lernen in 14 Tagen*. MITP-Verlags GmbH & Co. KG, 2022.

[63] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," *Journal of the Royal statistical society: series B (Methodological)*, vol. 57, no. 1, pp. 289–300, 1995.
[64] R. Bellini, E. Tseng, N. Warford, *et al.*, "Sok: Safer digital-safety research involving at-risk users," *2024 IEEE Symposium on Security and Privacy (SP)*, 2024.
[65] L. Simko, A. Lerner, S. Ibtasam, F. Roesner, and T. Kohno, "Computer security and privacy for refugees in the united states," in *2018 IEEE symposium on security and privacy (SP)*, IEEE, 2018, pp. 409–423.
[66] C. W. Munyendo, P. Mayer, and A. J. Aviv, ""i just stopped using one and started using the other": Motivations, techniques, and challenges when switching password managers," in *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security*, 2023, pp. 3123–3137.
[67] N. Alkaldi and K. Renaud, "Why do people adopt, or reject, smartphone password managers?" In *1st European Workshop on Usable Security-EuroUSEC 2016*, 2016.
[68] S. Pearman, S. A. Zhang, L. Bauer, N. Christin, and L. F. Cranor, "Why people (don't) use password managers effectively," in *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*, 2019, pp. 319–338.
[69] P. Mayer, C. W. Munyendo, M. L. Mazurek, and A. J. Aviv, "Why users (don't) use password managers at a large educational institution," in *31st USENIX Security Symposium (USENIX Security 22)*, Boston, MA: USENIX Association, Aug. 2022, pp. 1849–1866.
[70] S. Chiasson, P. C. van Oorschot, and R. Biddle, "A usability study and critique of two password managers.," in *USENIX Security Symposium*, vol. 15, 2006, pp. 1–16.
[71] N. Huaman, S. Amft, M. Oltrogge, Y. Acar, and S. Fahl, "They would do better if they worked together: The case of interaction problems between password managers and websites," in *2021 IEEE Symposium on Security and Privacy (SP)*, IEEE, 2021, pp. 1367–1381.
[72] R. A. CARDOŞ, R. ŞOFLĂU, A. Gherman, M. SUCALĂ, and A. Chiorean, "A mobile intervention for core needle biopsy related pain and anxiety: A usability study.," *Journal of Evidence-Based Psychotherapies*, vol. 17, no. 1, 2017.
[73] M. Kytö, S. Koivusalo, A. Ruonala, *et al.*, "Behavior change app for self-management of gestational diabetes: Design and evaluation of desirable features," *JMIR human factors*, vol. 9, no. 4, e36987, 2022.
[74] F. Alqahtani, G. Al Khalifah, O. Oyebode, and R. Orji, "Apps for mental health: An evaluation of behavior change strategies and recommendations for future development," *Frontiers in Artificial Intelligence*, vol. 2, p. 30, 2019.
[75] G. Villalobos-Zúñiga and M. Cherubini, "Apps that motivate: A taxonomy of app features based on self-determination theory," *International Journal of Human-Computer Studies*, vol. 140, p. 102 449, 2020.
[76] D. Bakker, N. Kazantzis, D. Rickwood, and N. Rickard, "A randomized controlled trial of three smartphone apps for enhancing public mental health," *Behaviour Research and Therapy*, vol. 109, pp. 75–83, 2018.
[77] N. Ibrahim, S. J. Hui, Y. K. Li, *et al.*, "The effectiveness of a persuasive mobile app to influence habit change," *European Proceedings of Social and Behavioural Sciences*, 2020.

Appendix A. List of Tasks

Topics in brackets, tasks with fixed positions are marked with an asterisk (*).

1) *Computer/Phone: 2-Factor Authentication Google (Passwords)
2) *Computer: Lock (Device Security)
3) Phone: Unlock Code (Device Security)
4) Phone: Android Update (Software)
5) *No device: Create a Master Password (Passwords)
6) Computer/Phone: Google Privacy Check (Privacy)
7) *Computer: Choose a Password Manager (Passwords)
8) *Computer: Install Password Manager (Passwords)
9) *Phone: Install Password Manager (Passwords)
10) Computer: Windows Update (Software)
11) *Computer: Password Manager Save Accounts (Passwords)
12) *Computer: Change Duplicate Passwords (Passwords)
13) Phone: Automatically Update Apps (Software)
14) *No device: Backup Location (Data Storage)
15) *Phone: Back up Data (Data Storage)
16) *Computer: Password Manager Save Accounts (Passwords)
17) *Computer: Change Duplicate Passwords (Passwords)
18) *Computer: Browser Update (Software)
19) Computer: Adblocker (Browsers)
20) *Computer: Update to more Secure Passwords (Passwords)
21) *Computer: Backup Documents (Data Storage)
22) *Computer: Backup Other Data (Data Storage)
23) *Computer: Program Updates (Software)
24) *Computer: Update to more Secure Passwords (Passwords)
25) Computer: 2-Factor Authentication (Passwords)
26) Phone: Adblocker (Browsers)
27) Phone: Messenger Verification (Privacy)
28) Phone: Messenger Privacy (Privacy)
29) *Computer: Update to more Secure Passwords (Passwords)
30) *No device: What's next?

Appendix B. Surveys

Pre-Screen Survey

1) I am primarily using an Android phone.
2) If yes: I would download the study app and use it for 30 days on this Android phone.
3) I would download the study app via download link (external source, not Google Play Store).
4) I am primarily using a Windows computer (desktop, laptop, or tablet).
5) If yes: I would use this Windows computer for the study for 30 days.
6) I use the following browsers.
7) Do you have any of the following: A certificate in computer security. Taken classes on computer security. Have a degree in computer science, cyber security, programming, or a related field. Other professional experience in computer security. None of the above.
8) I am interested in improving my computer security and privacy behavior.
9) I am willing to install new software on my devices, as well as change passwords or account settings.
10) I am willing to participate in a study for 10 minutes every day for 30 days, using my Android phone and Windows computer?
11) I am frequently overwhelmed with computer security and privacy.
12) I wish I could improve my personal online behavior, but I don't know how.
13) If others have security or privacy problems with their computers or mobile devices, they often ask me for help.
14) If I have security or privacy problems with my computer or mobile device, I often ask others for help.
15) During the last 30 days, I...
    - used a password manager.
    - used the same password for different accounts.
    - updated software on my phone.
    - updated software on my computer.
    - backed up documents from my computer.
    - backed up pictures from my phone.
    - used 2-Factor Authentication for my Google account.
16) Please describe one thing about computer security or privacy that you would like to learn more about.
17) Age: What is your age?
18) Gender: What is your gender?
19) Place of residence: In which country do you live?
20) Education: What is the highest level of school you have completed or the highest degree you have received?

21) Income: Information about income is important to understand. Would you give your best guess? Please indicate the answer that includes your entire household income in 2022 before taxes.

Assessment Survey

SA-13

1) I seek out opportunities to learn about security measures that are relevant to me.
2) I am extremely motivated to take all the steps needed to keep my online data and accounts safe.
3) Generally, I diligently follow a routine for security practices.
4) I often am interested in articles about security threats.
5) I always pay attention to experts' advice about the steps I need to take to keep my online data and accounts safe.
6) I am extremely knowledgeable about all the steps needed to keep my online data and accounts safe.
7) I am too busy to put in the effort needed to change my security behaviors.
8) I have much bigger problems than my risk of a security breach.
9) There are good reasons why I do not take the necessary steps to keep my online data and accounts safe.
10) I usually will not use security measures if they are inconvenient.
11) I want to change my security behaviors to improve my protection against threats (e.g. phishing, computer viruses, identity theft, password hacking) that are a danger to my online data and accounts.
12) I want to change my security behaviors in order to keep my online data and accounts safe.
13) I worry that I'm not doing enough to protect myself against threats (e.g. phishing, computer viruses, identity theft, password hacking) that are a danger to my online data and accounts.

RSeBIS

1) I set my computer screen to automatically lock if I don't use it for a prolonged period of time.
2) I use a password/passcode to unlock my laptop or tablet.
3) I manually lock my computer screen when I step away from it.
4) I use a PIN or passcode to unlock my mobile phone.
5) I change my passwords even if it is not needed.
6) I use different passwords for different accounts that I have.
7) When I create a new online account, I try to use a password that goes beyond the site's minimum requirements.
8) I include special characters in my password even if it's not required.
9) When someone sends me a link, I open it only after verifying where it goes.
10) I know what website I'm visiting by looking at the URL bar, rather than by the website's look and feel.
11) I verify that information will be sent securely (e.g., SSL, "https://", a lock icon) before I submit it to websites.
12) When browsing websites, I mouseover links to see where they go, before clicking them.
13) If I discover a security problem, I fix or report it rather than assuming somebody else will.
14) When I'm prompted about a software update, I install it right away.
15) I try to make sure that the programs I use are up-to-date.
16) I verify that my anti-virus software has been regularly updating itself.

SRHI

Dealing with computer security and privacy is something...

1) ... I do frequently.
2) ... I do automatically.
3) ... I do without having to consciously remember.
4) ... that makes me feel weird if I do not do it.
5) ... I do without thinking.
6) ... would require effort not to do it.
7) ... that belongs to my (daily, weekly, monthly) routine.
8) ... I start doing before I realize I'm doing it.
9) ... I would find hard not to do.
10) ... I have no need to think about doing.
11) ... That's typically 'me'.
12) ... I have been doing for a long time.

Completion Survey

1) SA-13
2) RSeBIS
3) SRHI
4) Which task(s) did you find particularly helpful and why?
5) Which task(s) did you not find particularly helpful and why?
6) If you needed help with the tasks that the Security App did not provide, where did you get that help? (e.g. asked a friend, searched on google)
7) SUS
   a) I think that I would like to use this app frequently.
   b) I found the app unnecessarily complex.
   c) I thought the app was easy to use.
   d) I think that I would need the support of a technical person to be able to use this app.
   e) I found the various functions in this app were well integrated.
   f) I thought there was too much inconsistency in this app.
   g) I would imagine that most people would learn to use this app very quickly.
   h) I found the app very cumbersome to use.
   i) I felt very confident using the app.
   j) I needed to learn a lot of things before I could get going with this app.
8) What did you like most about the Security App and why?
9) What did you not like about the Security App and why?
10) On a scale from 0-10, how likely are you to recommend the Security App to a friend or colleague?
11) Why would you (not) recommend the Security App?
12) Would you like to use the Security App in your everyday life and why?
13) I am frequently overwhelmed with computer security and privacy.
14) I wish I could improve my personal online behavior, but I don't know how.
15) Participating in the study has improved my computer security and privacy behavior.
16) Participating in the study has increased my awareness of computer security and privacy.
17) Apart from this study, did you get any other computer security and privacy advice in the last 30 days?
18) Other comments on the study

Final Survey

1) SA-13
2) RSeBIS
3) SRHI

In-App Survey *Success* and *Always*

1) I think I understood this task.
2) I think the task is simple to implement.
3) I think the amount of information that was presented in this task was: (Too much, Somewhat too much, Just right, Somewhat too little, Too little)
4) I would recommend [the activity] to a friend.
5) [The activity] increases my security and privacy.

In-App Survey *Fail*

1) What was the problem? (free text)
2) How long did you try it? (free text)
3) I think I understood this task.
4) I think the task is simple to implement.
5) I think the amount of information that was presented in this task was: (Too much, Somewhat too much, Just right, Somewhat too little, Too little)
6) I would recommend [the activity] to a friend.
7) [The activity] increases my security and privacy.

In-App Survey *Reject*

1) Please describe the task. What was the goal of the task? (free text)
2) Why did you decide not to do it? (free text)
3) I think I understood this task.
4) I think the task is simple to implement.
5) I think the amount of information that was presented in this task was: (Too much, Somewhat too much, Just right, Somewhat too little, Too little)
6) I would recommend [the activity] to a friend.
7) [The activity] increases my security and privacy.

In-App Survey *End* (for task 30)

1) Do you use any other devices in addition to your Android phone and Windows computer?
2) Do you want to set up a password manager on another device? Which one?
3) Do you want to set up a backup strategy for another device? Which device, how and when?
4) Can you list 3 more accounts that you want to save in your password manager?
5) Can you list 3 more accounts for which you want to use 2-Factor-Authentication?
6) Do you still use any insecure and/or duplicate passwords that you want to change? How and when?
7) Do you want to check more security and privacy settings of accounts? Which account will you start with, and when?
8) Schedule the next backup of your data. Which data do you want to back up when?
9) Is there anything else you want to do for your computer security and privacy?

Appendix C. Tables

	App Group (N=36)	**Control Group** (N=38)
Age		
19-30	24	10
31-40	5	16
41-50	6	8
$>$ 50	1	4
Gender		
Man	10	14
Woman	21	14
Non-Binary	1	1
Self-Describe	4	8
Not Disclosed	0	1
Country		
Germany	19	10
US	17	28
Degree		
High school or equivalent	5	7
College	10	13
Bachelor's degree	14	10
Higher than Bachelor's degree	7	8
Employment		
Full-time	11	16
Part-time	7	3
Self-employed	4	4
Out of work	2	1
Stay-at-home-parent	3	2
Student	7	5
Retired	1	1
Unable to work	0	6
Self-describe	1	0
Income		
$<$ 15k	7	9
15k - 29.99k	7	4
30l - 49.99k	7	6
50k - 74.99k	4	7
75k - 99.99k	7	4
100k - 149.99k	1	7
$>$ 150k	2	1
Not Disclosed	1	0

TABLE 2. PARTICIPANT DEMOGRAPHICS.

Task	**Completed**	**Adopted**	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**
1	36	29	4.94	4.67	2.83	4.25	4.69
2	36	28	4.94	4.67	2.75	4.44	4.81
3	29	29	5	4.97	2.86	4.97	4.97
4	30	30	5	4.93	2.97	4.9	4.77
5	36	28	4.83	4.39	2.67	4.14	4.14
6	28	28	4.82	4.71	2.61	4.36	4.39
7	36	31	4.92	4.5	2.53	4.25	4.25
8	36	30	4.72	4.11	2.75	4.22	4.28
9	36	31	4.83	4.56	2.92	4.28	4.14
10	27	27	5	4.93	2.89	4.67	4.7
11	36	31	4.94	4.78	2.92	4.47	4.36
12	36	26	4.92	4.47	2.97	4.5	4.78
13	30	27	5	5	3	4.5	4.53
14	36	35	4.86	4.47	2.92	4.78	4.53
15	36	36	4.97	4.47	2.69	4.86	4.42
16	33	29	5	4.79	2.97	4.54	4.36
17	33	28	4.94	4.67	3.06	4.64	4.7
18	31	31	5	4.97	3	4.87	4.77
19	30	27	4.97	4.9	2.8	4.67	4.4
20	30	28	4.93	4.27	2.9	4.7	4.93
21	30	30	5	4.6	2.83	4.97	4.77
22	28	27	4.86	4.46	2.64	4.71	4.46
23	27	26	4.96	4.78	2.7	4.59	4.56
24	26	22	4.92	4.62	3.04	4.92	4.96
25	30	26	5	4.73	2.9	4.53	4.8
26	28	17	4.89	4.11	2.79	4.57	4.57
27	26	18	4.5	3.65	2.77	3.96	4.31
28	31	29	4.81	4.64	2.68	4.42	4.39
29	9	8	5	5	3	5	5

TABLE 3. OVERVIEW OF TASK RATINGS: COMPREHENSIBLE (Q1), SIMPLE TO IMPLEMENT (Q2), AMOUNT OF INFORMATION (Q3), RECOMMEND TO FRIEND (Q4), INCREASES SECURITY AND PRIVACY (Q5).

Appendix D. Meta-Review

The following meta-review was prepared by the program committee for the 2025 IEEE Symposium on Security and Privacy (S&P) as part of the review process as detailed in the call for papers.

D.1. Summary

The paper discusses a longitudinal study focused on providing security advice using an android app. In the study with 74 participants, the paper discusses in-depth survey results based on 30 day app usage, reminding users with actionable advice. The results show the effectiveness of the approach and positive implementation of security advice.

D.2. Scientific Contributions

- Creates a New Tool to Enable Future Science
- Addresses a Long-Known Issue
- Provides a Valuable Step Forward in an Established Field

D.3. Reasons for Acceptance

1) Well thought-out study design
2) Impactful study to raise security awareness and help people build healthy security habits
3) Reports insightful study results