The Kids Are All Right: Investigating the Susceptibility of Teens and Adults to YouTube Giveaway Scams

Elijah Bouma-Sims, Lily Klucinec, Mandy Lanyon, Julie Downs, Lorrie Faith Cranor
Carnegie Mellon University
{eboumasi, laklucin, holbrook, downs, lorrie}@andrew.cmu.edu

***Abstract*—Fraudsters often use the promise of free goods as a lure for victims who are convinced to complete online tasks but ultimately receive nothing. Despite much work characterizing these "giveaway scams," no human subjects research has investigated how users interact with them or what factors impact victimization. We conducted a scenario-based experiment with a sample of American teenagers (n = 85) and adult crowd workers (n = 205) in order to investigate how users reason about and interact with giveaway scams advertised in YouTube videos and to determine whether teens are more susceptible than adults. We found that most participants recognized the fraudulent nature of the videos, with only 9.2% believing the scam videos offered legitimate deals. Teenagers did not fall victim to the scams more frequently than adults but reported more experience searching for terms that could lead to victimization. This study is among the first to compare the interactions of adult and teenage users with internet fraud and sheds light on an understudied area of social engineering.**

I. Introduction

Fraud is a persistent and growing problem online. The FBI Internet Crime Complaint Center (IC3) received over 880,000 reports in 2023, an increase of 90% from 2019 [30]. A common social engineering tactic used by fraudsters is to create websites that purport to offer free goods or services (e.g., iPads, gift cards, mobile game currency, access to movies, etc.). Sometimes referred to as "giveaway scams," fraudsters advertise their websites through social media advertisements [20], YouTube videos [11], [19], hijacked websites [15], and other means [6]. On such websites, victims are directed to complete surveys or other tasks, such as downloading mobile apps, but ultimately receive no compensation. Although this scam often does not cause direct financial harm, victims waste their time and may end up giving their sensitive personal information to fraudsters. Moreover, the frequent use of mobile game currency as a lure may make teenagers (minors aged 13 to 17) particularly susceptible to these schemes [6], [11].

Much prior research has characterized different variations of giveaway scams [20], [46], which are often distributed on social media platforms like YouTube [6], [11], [19], [82]. These types of scams may entrap a significant number of victims. For example, Badawi et al. estimated that a subset of giveaway scams called the "Game Hack Scam," received 60 million impressions from 2014 to 2018 [6]. Additionally, cryptocurrency giveaway scams perpetuated on YouTube have been found to cause millions of dollars of loss [82].

Despite this evidence of harm, no human subjects research has characterized how or why participants fall for these schemes. Moreover, there is also a lack of research on social engineering involving types of media other than static text [14]. To help address these gaps, we conducted an online, scenario-based experiment to evaluate both adults' $(n = 205)$ and teens' $(n = 85)$ reactions to giveaway scams advertised via YouTube videos. Participants were asked to advise a hypothetical friend who was searching either for free in-game currency for Roblox, called Robux, or access to Spotify Premium. The experiment was designed to simulate the path a victim would go down, with participants first viewing search results, then a YouTube video, and finally, a video of the websites linked by the YouTube video. At each step, participants indicated what their friend should do as well as why they would recommend this action. Participants were assigned to view one scam and one video presenting a legitimate free offer (e.g., a free trial) randomly selected from three scams and three legitimate offers for each topic. Stimuli used in the experiment were derived from actual YouTube search results found during a qualitative investigation of giveaway scams.

Through our experiment, we sought to explore three main research questions.

- **RQ1:** How do users reason about YouTube giveaway scams, and what actions do they recommend when encountering them?
- **RQ2:** Do teens and adults interact differently with YouTube giveaway scams? If so, do these differences in behavior result in teens being more vulnerable?
- **RQ3:** Do other demographic and behavioral factors (e.g., income, experience with scam victimization, etc.) impact users' interactions with giveaway scams?

We found that most participants would not fall for these scams, with less than 10% of participants indicating that they believed that the scam stimuli were legitimate. Users generally

Network and Distributed System Security (NDSS) Symposium 2025
24-28 February 2025, San Diego, CA, USA
ISBN 979-8-9894372-8-3
https://dx.doi.org/10.14722/ndss.2025.240342
www.ndss-symposium.org

immediately exited the less convincing scams, while the more convincing scams led them to want to do more research into the offer (RQ1). Teens were not significantly better or worse at identifying giveaway scams, nor were they more likely to indicate that their friend should follow the instructions from the scam. We did note some significant differences in behavior, with teens more likely to recommend that their friends report the scam videos that they saw. Teens were also significantly more likely to have previously searched for free Roblox Robux, the keyword we used to find scam videos on YouTube (RQ2). None of the behavioral or demographic factors we investigated significantly influenced users' behavior in relation to scams, although women expressed more uncertainty and thus were slightly less likely than men to accurately identify scams (RQ3). Ultimately, our study provides insight into an understudied type of social engineering and presents one of the first studies comparing adults' and teens' interactions with internet fraud.

The remainder of the paper proceeds as follows: In section II we review prior work on fraud, security behavior across age groups, and teen safety online. In section III, we describe a qualitative investigation of these scams on YouTube that we used to find stimuli for our user study as well as our analysis to quantify the number of giveaway scams similar to those presented in our survey. We discuss the methods for our human subjects experiment in section IV and the results in section V. We then discuss the implications of our results and possible guidelines for better protecting users from giveaway scams in section VI. We conclude in section VII.

II. Background and Related Work

In this section we review related work and other background information that underpins our work, focusing particularly on internet fraud and factors that affect users' vulnerability to victimization, the relationship between age and different security behaviors, research related to teen online safety, and an overview of giveaway scams on YouTube and other platforms.

A. *Internet fraud and factors affecting victimization*

Fraud is a broad term referring to any criminal behavior where trickery is used to obtain financial gain [16]. Different forms of fraud or "scams" are among the most common types of cybercrime reported to government agencies. For example, 33.9% of the crimes reported to the FBI Internet Crime Complaint Center in 2023 were phishing [30], a form of fraud where an attacker impersonates a trustworthy entity (e.g., an online service, bank, government institution, etc.) in order to obtain sensitive private information [50]. Other common forms of internet fraud include online romance scams, imposter scams, and fraudulent investments [30].

There is a large body of research investigating what human factors affect users' susceptibility to different types of internet fraud. Due to its prevalence, phishing is one of the most studied forms of internet fraud [26], [59], [92], but researchers have also explored factors affecting susceptibility to other common types of fraud, such as IRS imposter scams [69], online shopping scams [67], and romance scams [88]. Knowledge of scams has been shown to mitigate susceptibility, with training reducing rates of victimization [47], [48], [69], [75], [76]. Differences in individual psychological characteristics, such as personality, may also affect rates of fraud victimization [59], [92]. For example, higher impulsivity and lower self-control have been linked to greater susceptibility to fraud [18], [64], [67], [88].

The content of a fraudulent message is another important factor in understanding users susceptibility. Fraudulent schemes are most effective when aligned with the interests of potential victims. For example, phishing emails are more successful when they align with users' interests [92]. In the context of our investigation, the promise of mobile game currency is unlikely to entice those who do not play the game. Regardless of the specific topic of a scam, fraudsters often employ persuasive techniques in their communications to exploit cognitive biases or try to evoke a particular emotional state. For example, a phishing email may purport to come from an employee's boss in order to take advantage of the tendency to obey authority. The presence of different persuasion tactics has been found to increase susceptibility to phishing emails [83], [89].

B. *Age and Online Security Behavior*

Age is a frequently investigated predictor of behavior in security research [85] due to differences in both developmental stage and experiences by virtue of generational milieu [28]. Older users may be more likely than younger users to report engaging in good security practices [12], [63], [85], [93]. In contrast, younger adults may be more likely to report adoption of privacy-protecting strategies [45], [85], [93], including use of private browsing [37] and use of tracker blockers [55]. This could be explained by the different sources of security information used by younger and older users [66], older users' greater awareness of security [56], [93], or older users' greater willingness to adopt positive security behaviors in response to social pressure [25].

Despite the abundance of past research comparing younger and older users, the relationship between age and internet fraud susceptibility remains unclear. Those 60 or older are the most likely demographic to report falling victim to internet fraud [30], [81], but this could be affected by reporting bias. Research indicates that consumers who are more highly educated or those who experience a greater loss are more motivated to report fraud victimization [22], [43], [44], [65].

The results of past experimental studies on phishing are mixed [74], [85], [92], with some older studies finding victims are more likely to be younger [71], [75], [84], other more recent studies finding them more likely to be older [7], [34], [35], [52], and still others finding no relationship to age [62], [70]. These studies uniformly focus on adult populations, with no prior work comparing minors' and adults' interactions with fraud. Demographic analyses of other types of internet fraud are limited but suggest that different types of frauds may target different age groups, such as middle-aged people being more susceptible to online romance scams [88].

C. Teens' Online Safety

Teenagers are an important demographic to study due to their distinctive online behavior patterns and susceptibility to digital risks. Unlike prior generations, which encountered computers later in life, modern American teenagers have grown up with digital devices and the internet. Most American teenagers own or have access to a smartphone (95%) or a desktop/laptop computer (90%) [4]. YouTube, TikTok, Snapchat, and Instagram are the most popular social media websites among teens [4], [68]. Adults also frequently use YouTube and Instagram but seem to be more likely to use Facebook and Pinterest and less likely to use TikTok or Snapchat [40]. In addition to browsing social media, other common activities for teens include watching online videos and playing video games, both activities which teens seem to do at a higher rate than adults [29], [40], [68].

Adolescence is an important period of development, during which emotional and social networks in the brain mature faster than the prefrontal cognitive-control network. This makes teenagers more susceptible to heightened impulsivity, sensation seeking, and emotional reactivity. These developmental dynamics contribute to a general trend of increased risk-taking behaviors among teenagers that declines with age [51, pgs. 528 – 530]. Because risk-taking is associated with fraud susceptibility (see section II-A), teens may be more likely to fall victim to fraud than adults, although thus far, no research has tested this hypothesis.

Research on teen online safety has focused on privacy concerns (including social media [41], [53], [54]), password management [78], and teen-specific behaviors such as online sexual exploitation [2], [87] and cyberbullying [91]. Researchers have also explored approaches to protecting teens from online dangers via parental controls and other interventions [24], [33], [57].

A small number of studies have investigated how teenagers are victimized by phishing. For example, Lastdrager et al. [49] conducted a study with Dutch children between 8 and 13 years of age ($n = 353$) to evaluate the effectiveness of anti-phishing training in the form of a 40-minute presentation on cybersecurity designed with children in mind (i.e., using story-telling and examples focused on children). As expected, children who received training performed better than those who did not receive training. Controlling for training, older children performed significantly better than younger children at correctly labeling emails. Similarly, another experiment [3] evaluated the effectiveness of embedded phishing awareness training with children from 7 to 13 years of age located in Saudi Arabia ($n = 30$), and concluded that phishing awareness education improved children's performance, although these findings are not statistically significant. Finally, Nicholson et al. [58] conducted an evaluation of the ability of English children aged 12 to 17 ($n = 83$) to recognize phishing emails in an experimental setting. Participants generally performed poorly, with an overall success rate of 59% correctly discriminating between legitimate and fraudulent emails, but the researchers observed no difference with age.

Beyond studies on teens and phishing, there is not much research on how minors interact with non-phishing forms of internet fraud. Our study helps fill this gap by exploring how teens and adults reason about and interact with YouTube giveaway scams.

D. Giveaway Scams

Previous research has contributed to the understanding of giveaway scams, which purport to give valuable rewards for free in exchange for completing a task, but generally involve a never-ending string of surveys or other tasks with no actual reward [6], [11], [19], [20], [46]. For example, Clark et al. [20] analyzed spam posts on Facebook in 2013 and found that over 70% of them contained links with surveys designed to collect and sell respondents' data without compensation. Kharraz et al. developed [46] an automated system for identifying this type of survey scam, finding that websites hosting fake surveys might also contain malware or potentially unwanted programs (PUPs), creating additional security concerns around these types of scams.

A specific type of giveaway scam that is relevant to teenagers is the Game-Hack Scam (GHS), which uses similar methods as the broader giveaway scam described above to obtain user information or install potentially dangerous content on respondents' devices with the promise of free access to rare or paid in-game materials [6]. Bouma-Sims et al. [11] found a high prevalence of GHS scams using organic searching on YouTube.

III. Exploring YouTube Giveaway Scams

In this study we examine teens' and adults' reactions to giveaway scams and test whether they have differential ability to distinguish scam from legitimate websites using an ecologically valid set of stimuli. To this end, we first investigate and describe the landscape of YouTube giveaway scams, and then use these findings to identify stimuli for our human subjects experiment. The basis for this investigation is two-fold: 1) conducting organic searches on YouTube and 2) identifying giveaway scams in a sample of 1% of YouTube collected for a previous unrelated research study in 2021 [1].

A. Organic Searching on YouTube

We conducted three rounds of organic searches on YouTube to identify giveaway scams that appeared to target teenagers. Organic searches were used to best simulate how teenagers interact with YouTube. Prior work has also used organic searching to find giveaway scams [6], [11]. Our multi-stage method of "progressive focusing" was inspired by Wei et al.'s [86] study of surveillance described in TikTok posts. The first round of searching was designed to determine if search results differed based on demographic features, and the second round of searching narrowed down the relevant search terms.

Finally, the third round of searching used the narrowed list of terms to find videos to include in our survey.

To perform the searches in the first round, we created six YouTube accounts to cover all combinations of age (13 or 17) and gender options (male, female, or prefer not to say). These ages were selected to compare results associated with each end of the age range. Each account was created and used within a separate emulated Android environment, using the Genymotion Android emulator. We used the following search terms and analyzed the first 10 results: "make money online for teenagers," "free robux," "free walmart gift cards," and "free ipad." These specific search terms were selected based on the expectation that they would reveal giveaway scams [11] and their relevance to teenagers [68]. We did not find enough qualitative differences in the search results between the accounts and decided to only use one account for the sake of time moving forward.

For the second round of searching, we created a new account with gender listed as prefer not to say and age as 15 years old. These demographics were selected to fall in the middle of our age range and be neutral in terms of gender. We collected the first 10 results from each of the following search terms: "easy online jobs," "free gift cards," "free netflix," "free roblox robux," "free spotify premium," and "make money online for teenagers." Changes to search terms were informed by the results of the first round of searching (e.g., finding no scams from the terms "make money online for teenagers" or "free ipad"). We analyzed the search results for both legitimate opportunities for financial gain, often through programs such as Microsoft Rewards, or for giveaway scams [6], [11], [20], [46]. Of the six terms, only "free spotify premium" and "free roblox robux" produced the variety of results desired, and the other terms were dropped.

For our final round of searching, we searched for videos matching our desired criteria to include as stimuli in our study. We defined three major categories of videos related to monetary gain without payment on YouTube a priori: legitimate, clickbait, and scams. We defined legitimate videos as real and largely safe opportunities for people to receive a monetary benefit, such as earning points through Microsoft Rewards or signing up for a free trial of a streaming platform. We defined clickbait videos as videos having titles unrelated to their actual content, mainly to trick people into viewing them without delivering what they claimed, or including non-harmful but ineffectual "methods" of receiving a benefit, such as editing the HTML of a website. Finally, we define scam videos as those that encourage viewers to pursue unsafe methods of obtaining rewards that will never be received, such as "human verification" to unlock a reward, which leads to an endless loop of tasks to complete, such as filling out surveys with personal information or downloading apps. This criterion was adapted from Bouma-Sims et al. [11], which broadly defined a scam video as one "that attempts to attract viewers through misrepresentation, including fraudulently offering tangible, intangible, or financial awards." Two researchers independently classified each video, discussing to resolve disagreements.

We conducted searches over a two-week period on the same account used for the second round of searching until we found three results in each aforementioned category for both the "free spotify premium" and "free roblox robux" search terms. All of the legitimate Spotify videos involved free trials through Spotify directly or brand partnerships between Spotify and other trusted companies like PayPal. Similarly, all of the legitimate Roblox results involved earning points through Microsoft Rewards to trade in for Robux. Most of the scams across both search terms involved websites posed as coupon hosting services, with the exception being a website with many "Download" buttons that lead to an "Unlock Link" page once clicked.

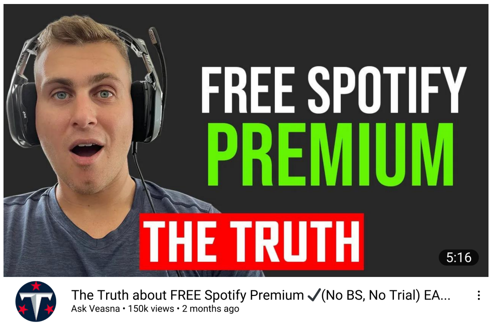
B. Identifying Giveaway Scams on 1% of YouTube

To quantify the number of giveaways scams that are similar to those presented in our survey, we analyzed data from a previous, unrelated study, collected using "random prefix sampling" representing approximately one percent of all videos on YouTube between August 2020 and March 2021, for a total dataset of "about 10 million videos that have at least 800 views" at the time of sampling [1]. We analyzed the data using regex commands on JSON files in the dataset to pull out all videos with a variation of the words "robux" or "spotify" in the title, as these were the two search terms with the most relevant results in our organic searches. We found many results for the "spotify" search, but none were the giveaway scam format we desired, which is likely due to the high usage of "spotify" in videos posted by music artists. We found 451 relevant results for the "robux" search, and proceeded to analyze each result.

We next attempted to view each result on YouTube and assigned one of five codes based on each video's availability on YouTube as of April 2024: "available," "not available," "private," "unlisted," and "not English." Only the videos that were coded as "unlisted" or "available" were investigated further, leaving a total of 161 Robux-related videos to be analyzed out of the initial 451 (approximately 36% of the relevant results). We analyzed each of the remaining result videos, and those matching the giveaway scam format were flagged as scams. Of the 161 available videos, we determined 17 of them to be scams, which is approximately 11% of the available video results from the "robux" search. These results indicate that there would be approximately 1,199 Robux giveaway scams on YouTube between August 2020 and March 2021 based on this random sample of approximately 1.4% of videos on the platform [1]. We anticipate that the actual number is larger based on our organic searching, as many of the effective scams were likely taken down over the past four years.

IV. Methods

We describe our online behavioral experiment designed to evaluate adult and teenage participants' ability to recognize YouTube giveaway scams and assess how they reason about and interact with these scams. We first discuss our experimental design, before discussing recruitment, data analysis, ethical considerations, and limitations of our work. All materials



The Truth about FREE Spotify Premium ✓(No BS, No Trial) EA...
Ask Veasna • 150k views • 2 months ago

Fig. 1: Example of a search result displayed to participants in the Spotify condition (Spotify Scam 2).

necessary to replicate our experiment (e.g., survey instruments, recruitment materials, stimuli, etc.) are available in the associated artifact (see Appendix A).

A. *Experimental Design*

To evaluate our hypotheses, we employed a scenario-based design meant to mimic the process of YouTube scam victimization. In the experiment, participants were asked to imagine they were advising a friend searching on YouTube for ways to earn free Roblox Robux or gain free access to Spotify Premium. We decided to have participants advise a friend so stimuli could be presented randomly, without suggesting that the participant had done something they did not want to do. Previous studies of security and privacy behavior have also asked participants to imagine they are advising a friend [13], [42]. Participants were randomly assigned to one of these two topics and told the normal cost of Spotify Premium or Roblox Robux accordingly. Participants then completed several tasks.

1) Search result selection task: Participants first viewed nine YouTube search results in a random order (see Figure 1). Among these search results were three legitimate videos, three clickbait videos, and three scam videos (see section III). Participants were asked to select the videos they would advise their friend to click on or indicate if they would not select any of the videos. Participants were asked to justify their response by selecting which of the six features of the search result they considered (i.e., number of views, date of publication, the title of the video, thumbnail image, name of the channel, profile picture of the channel, or the length of the video). They were also asked to explain how they selected search results in an open-ended question.

2) Video review task: Participants were next shown two YouTube videos, one presenting a scam and the other presenting a legitimate free offer. These videos were selected at random from the three scam videos and three legitimate videos for each topic, regardless of which search results the participant indicated that they would select in the previous task. They were asked to watch each video and advise their friend on what to do next. Participants could opt to visit the website, exit the video, report it to YouTube, search online to learn more, look at the comments, leave a comment, or indicate that they would do something else. They were also asked to justify their answer and identify which features of the video influenced their decision. Immediately after answering questions about each YouTube video, participants were told that their friend had gone to the website linked in the video. They were then shown another video of someone scrolling through the website. As with the YouTube video, they were asked to advise their friend on what to do next and justify their answer. The legitimate stimuli and scam stimuli were shown in a random order to control for potential order effects. Participants needed to watch at least half of each video before they could advance in the survey.

3) Scam rating task: After collecting survey responses from 30 teen participants and 10 adult participants, we added a question to the end of the survey directly asking whether participants thought each of the videos they saw was a scam or not. Participants were shown the search result for each of the two videos they saw and were asked to rate on a five-point Likert scale, ranging from "definitely legitimate" to "definitely fraudulent." 55 (64.7%) of the teen participants and 195 (95.1%) of the adult participants answered these questions.

4) Demographic and behavior questions: Before and after the experimental tasks, participants were asked questions to contextualize their responses and measure potential predictors of fraud victimization. At the beginning of the survey, participants were asked questions about their internet usage habits developed based on the Common Sense Media census [68] (e.g., what social media services they use, how much time they spend doing digital entertainment activities, etc.). After the experimental tasks, participants were asked about their prior usage of Spotify/Roblox and whether they had ever looked for free Roblox Robux/free Spotify Premium before. Participants were then asked how often they displayed behaviors that we hypothesized may be associated with scam victimization (e.g., prior victimization by similar online scams, cryptocurrency usage, etc.). To encourage truthful answers, we included distractor questions about benign behaviors (e.g., using coupons, receiving cash back from a credit card, etc.). In the adult version of the experiment, the survey concluded with a series of demographic questions. These questions were also asked of the parents of teenage participants during the consent process (see section IV-B).

B. *Recruitment*

We recruited two samples of participants: a sample of minors aged 13 to 17 (n = 85) and an age-balanced sample of adult crowdworkers (n = 205).

1) Recruiting Teen Participants: Teen participants were recruited through a digital flier distributed to their parents/guardians. The flier informed parents about the study and included a link to an informed consent form. We used nonspecific language to describe the study tasks to avoid priming participants about the study (i.e., "[we are] investigating how teens interact with social media posts on YouTube"). The flier was distributed in various ways over time, as explained below.

More details on the recruitment procedure can be found in our companion publication [10].

Once a parent/guardian gave consent for their child to participate in the study, they were redirected to a survey asking questions about their child's demographics (e.g., state of residence, age, household income, race, etc.). Parents were then given a link to an assent form for their child to complete. The child assent form explained the study procedure and the potential risks of participation at an 8th-grade reading level (based on Flesch reading ease). If teen participants affirmed that they were eligible for the study (i.e., they were fluent in English, located in the US, used YouTube, and were between 13 and 17 years old) and that they assented to participate in the study, they were then redirected to the main survey instrument. After completing the survey, they were redirected to a final survey where they could enter their email address to be compensated with a $5 Amazon gift card.

The recruitment flier was initially distributed via snowball sampling by sending the flier to parents known to the authors who have children in the appropriate age range for the study. These parents were also asked to share the flier with other families that have eligible children who may be willing to participate. We recruited 30 teenage participants using this approach from November 11, 2023, to December 7, 2023.

On December 13, 2023, we began distributing fliers via Peachjar,[1] a K-12 school communication platform that allows third parties to pay to distribute digital fliers to families with children enrolled in particular schools. Our goal in using Peachjar was to obtain a larger and more diverse sample, so we submitted fliers to school districts around the US. We initially selected districts based on the median income where the school district was located, as reported in the 2022 American Community Survey (ACS) [79]. We selected the largest high school in school districts with at least one high school evenly from three income groups: less than $42,606, $42,606 to $74,580, and greater than $74,580. $42,606 was the threshold for a three-person household to receive reduced-price lunches starting July 2022 [80]. $74,580 is the national median household income reported in the 2022 ACS [79]. We used this method to select 25 schools to distribute our flier to from December 13, 2023 to January 8, 2024.

On February 26, 2024, we submitted our flier to 50 more schools. We selected school districts by sampling those with the most high schools. We expected these may have the largest student populations and would result in more participants. We also translated our materials for parents (i.e., flier, consent form, and parental survey) into Spanish to send fliers to schools that required Spanish and English fliers. Teenage participants were still required to be fluent in English to take the survey. Five additional fliers were submitted to schools on April 30, bringing the total number of schools to 90.

We initially did not implement measures to prevent spam, as we expected only parents in our selected school districts to see the fliers. We received many responses that appeared to be spam (i.e., duplicate responses, meaningless answers to open-ended questions, failed Captcha, etc.) between January 2-4, 2024. Once we noted this issue, we paused data collection and retracted the 17 Peachjar fliers posted up to this point. Out of caution, we removed all survey responses (122) received during this period, including a small number that wrote plausible responses to open-ended questions.

We relaunched our survey on January 8, 2024 with new anti-spam checks. We generated posters with a unique URL for each school district and required parents to select which school district their child attended from a dropdown list to progress through the survey. Parents who selected a school district or state different from the one corresponding to their flier could not proceed with the study. After implementing this procedure, most attempted spammers were screened out. Moreover, using a custom URL for each flier allowed us to quickly block spam from a particular flier without shutting down the entire survey. We had to remove a flier on March 8, 2024 due to one or more spammers discovering the correct answers to the security questions. 125 responses from this flier on March 7-8, 2024 were removed.

2) Recruiting Adult Participants: Adult participants were recruited using the online crowdworking platform Prolific[2]. We recruited adult participants in 10 gender-balanced age buckets (18-22, 23-27, 28-32, 33-37, 38-42, 43-47, 48-52, 53-57, 58-62, and 63 or older). We recruited 19 to 24 participants from each group, comprising 205 participants. Pre-experiment power analysis was performed using G*Power 3.1.9.7[3] to determine how many adult and teen participants we needed. This calculation used a repeated measures ANOVA model, assuming an effect size of 0.25, power = 0.95, $\alpha = 0.05$, 2 measures, 12 groups (i.e., the total number of legitimate and scam stimuli), and a 0.5 correlation between measures. The result of this calculation was a total sample size of 312.

As with the flier for parents of teens, the Prolific recruitment advertisement described the study broadly so as not to prime participants to expect scam videos. Adult participants were recruited over an extended period. An initial 10 participants were recruited on November 20, 2023. The remaining adult participants were recruited incrementally, one age bucket at a time, as responses were received from teen participants. The final adult participants were recruited on April 10, 2023.

C. Quantitative Data Analysis

We conducted statistical analysis to determine whether participant responses varied significantly ($p < \alpha = 0.05$) during the search result selection task, video selection task, or scam rating task. Participants' responses during the search result selection task (Q9) and video review task (Q12 and Q15) were multiple select questions. Each choice was thus treated as a separate binary outcome variable (e.g., whether the user recommended exiting the YouTube video or not, whether the user recommended reporting the video or not, etc.). The

[1] https://peachjar.com/

[2] https://www.prolific.com/

[3] https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower

responses to the scam rating task were interpreted as a binary variable corresponding to whether the participant correctly identified a video as a scam or legitimate.

The independent variables we investigated are listed in table I. Fisher's exact test [31] was used to determine whether the distribution of dependent variables varied significantly with respect to categorical and binary independent variables. For ordinal independent variables, we used the Cochran-Armitage test, which determines whether there is a significant linear relationship between an ordinal variable and a binary response variable [5]. To quantify the size of statistically significant differences, we computed Cramer's V [23] for categorical or binary dependent variables and Freeman's θ [32] for ordinal dependent variables. Both values range from 0 to 1, with 0 indicating no relationship between two variables and 1 indicating a perfect relationship. All planned comparisons involving the same dependent variable were corrected using the Benjamini and Hochberg procedure [8], [39]

For the search result task, we conducted additional statistical testing to determine whether the selection rate varied significantly between search results. For this purpose, we used Cochran's Q-test, which is suitable for determining whether a binary response variable differs between conditions for within-subjects data [21]. Serlin et al.'s maximum-corrected measure of effect size, η_Q^2, was used to quantify the size of differences for these tests [73]. η_Q^2 ranges from 0 to 1, with 1 indicating a perfect relationship and 0 indicating no relationship.

After completing the above analysis, we conducted some additional unplanned comparisons to explore potential confounds. Most notably, we performed statistical testing to compare the responses to behavioral and demographic questions between men and women. We also tested whether certain relationships remained significant when controlling for other variables.

Due to space constraints, we are unable to report all results. We focus on reporting significant results related to scam stimuli. Anonymized data from our experiment, the code used to run statistical tests, and the complete testing results are available in the associated artifact (see Appendix A).

D. Qualitative Data Analysis

Open-ended questions were analyzed via inductive coding performed by two authors. The lead coder reviewed 100 randomly selected responses to develop an initial codebook for each set of questions. The lead coder and another author then independently coded samples of 10% to 20% of the responses, meeting regularly to discuss differences in coding and decide on a consensus code for each set of responses. When new codes were added to the codebook, the lead coder re-reviewed previously coded responses, adding the new codes where necessary.

The coders started by analyzing the responses to the two questions where participants were asked to explain their response to the YouTube video. The lead coder realized during codebook development that the responses to these questions were often interdependent, so these two questions were analyzed together. Similarly, the responses to the two questions in which participants were asked to explain their response to the video of the website linked by the YouTube video were also analyzed together. The codebook developed while analyzing the video explanations served as the basis for the website explanations, as similar reasoning was displayed across these questions. A unique codebook was developed to analyze user explanations for YouTube search result selections.

TABLE I: Potential independent variables.

Variable	Type	Note
Adult	Binary	True if age $\geq$ 18
Condition	Categorical	Video shown
Time on digital entertainment (daily)	Ordinal	Response to Q1
Time watching online videos (weekly)	Ordinal	Response to Q4
Time playing mobile games (weekly)	Ordinal	Response to Q5
Time playing computer games (weekly)	Ordinal	Response to Q6
Time using social media (weekly)	Ordinal	Response to Q7
Time using other websites (weekly)	Ordinal	Response to Q8
Previous use of Roblox or Spotify	Ordinal	Response to Q18
Experience purchasing Robux or Spotify Premium	Binary	Response to Q19
Experience receiving Robux or Spotify Premium for free	Binary	Response to Q20
Experience with shopping scams	Ordinal	Response to Q27
Experience doing online tasks for money	Ordinal	Response to Q29
Experience purchasing cryptocurrency or NFTs	Ordinal	Response to Q31
Experience doing online tasks for money, but not receiving compensation	Ordinal	Response to Q33
Household income	Ordinal	Response to Q34
Gender	Categorical	Response to Q35
Type of community	Categorical	Response to Q39

The codebooks for each set of questions with definitions and example quotes can be found in the associated artifact (see Appendix A). Each codebook is divided into two categories: reasoning codes, which describe participants' decision-making process, and feature codes, which describe the specific elements of the stimuli mentioned by a participant. The majority of participant responses were assigned more than a single code.

E. Ethical considerations

Our study protocol was approved by our Institutional Review Board (IRB). To prevent participants from visiting scam websites, links, and other methods of accessing websites (e.g., images of Google search results) were redacted from all videos, including legitimate videos. We also included a debrief at the end of the survey that identified which videos the user saw were scams and gave advice on avoiding victimization.

Our primary ethical concern in recruiting teen participants was ensuring we received consent from their parents/guardians. If we recruited teenagers directly, it was likely that many would pretend to be their parents when asked to obtain parental consent. In consultation with our IRB, we opted to recruit teens through their parents, as this would ensure that consent was received before the teen even saw the study.

F. Limitations

Participants encountering YouTube videos promoting give-away scams in their normal lives may behave differently than they did in our study. Participants also may have had different perceptions of scams and legitimate websites if they had been able to interact with them directly rather than viewing a video of the website. While redaction of URLs was necessary to protect participants, this reduced participants' ability to detect scams accurately. Much of our quantitative analysis is based on self-reported experiences and behavior, which may be influenced by social desirability bias [36] or other forms of self-report bias. Our qualitative analysis is necessarily subjective, and a different research team may have identified different themes in the same data. Due to the difficulty in recruiting teen participants, our experiment may be slightly underpowered (see section IV-B2 for more details on the power analysis).

While we tried to recruit a diverse sample, our sample is not representative, which may limit the generalizability of our results. In both the teen and adult samples, there is a notable lack of participants identifying as Hispanic or Latino (see section V-A). The adult sample recruited from Prolific may have a higher level of technical expertise than the general population. Tang et al. [77] found that samples from Prolific are comparable to a probabilistic sample of the US population when asking about security and privacy perceptions and experiences, but not security and privacy knowledge. Additionally, teens from low-income households seem to be slightly underrepresented in our sample (see section V-A). While we did not find a relationship between income and scam susceptibility, teens from low-income households may experience different financial pressures, which might lead them to behave differently. Ultimately, future work should seek to explicitly study the behavior of marginalized groups when encountering internet fraud.

V. Results

In this section, we review the results of our behavioral experiment. We first discuss participants' demographics and behavior, highlighting the differences between adult and teen participants. We then discuss the results of each of our research questions.

A. Participant Characteristics

Table II contains an overview of participants' demographic characteristics, as well as the distribution of the most important behavioral characteristics. Participants were from 41 US states and Washington, D.C. The most common states where

TABLE II: Overview of selected demographic and behavioral characteristics. Questions about Roblox and Spotify were only asked of people in the appropriate conditions

	Adults		**Teenagers**	
	n	*%*	*n*	*%*
Gender				
Female	98	47.8%	37	43.5%
Male	102	49.8%	46	54.1%
Non-binary	4	2.0%	1	1.2%
Prefer not to say	1	0.5%	1	1.2%
Age				
13			9	10.6%
14			16	18.8%
15			27	31.8%
16			17	20%
17			16	18.8%
18-24	31	15.1%		
25-34	42	20.5%		
35-44	45	22.0%		
45-54	33	16%		
55+	54	26.3%		
Race				
American Indian/Alaskan Native	2	1.0%	2	2.3%
Asian	20	9.8%	7	8.2%
Black/African American	38	18.5%	14	16.5%
Hispanic/Latinx	13	6.3%	5	5.8%
Native Hawaiian/Pacific Islander	0	0%	1	1.2%
White	148	72.2%	62	72.9%
Self describe	3	1.4%	0	0%
Prefer not to say	1	0.5%	5	5.9%
Household Income				
Less than $20,000	20	9.8%	3	3.5%
$20,000 to %39,999	38	18.5%	10	11.8%
$40,000 to $59,999	36	17.6%	16	18.8%
$60,000 to $79,999	33	16.1%	4	4.7%
$80,000 to $99,999	19	9.3%	5	5.9%
$100,000 to $149,999	32	15.6%	14	16.4%
More than $150,000	21	10.2%	15	17.6%
Prefer not to say	6	2.9%	18	21.2%
Community Type				
Rural	38	18.5%	6	7.1%
Suburban	105	51.2%	44	51.2%
Urban	62	30.2%	30	35.2%
No response	0	0%	5	5.8%
Experience w/ Roblox ($n = 140$)				
Never	70	72.9%	10	22.7%
1 or 2 times	15	15.6%	13	29.5%
3 to 5 times	4	4.2%	5	11.3%
More than 5 times	7	7.3%	16	36.4%
Experience w/ Soptify ($n = 150$)				
Never	12	11	0	0%
1 or 2 times	22	20.2%	14	34.1%
3 to 5 times	6	5.5%	8	19.5%
More than 5 times	69	63.3%	19	46.1%
Time watching online videos (weekly)				
None	0	0%	1	1.2%
Less than an hour	10	4.9%	0	0%
1 to 5 hours	70	34.1%	21	24.7%
5 to 10 hours	53	25.9%	24	28.2%
10 to 15 hours	25	12.2%	19	22.4%
15 to 20 hours	19	9.3%	9	10.6%
More than 20 hours	28	13.7%	11	12.9%
Total	205	100%	85	100%

participants resided were California (12.4%), Florida (9.3%), and Texas (6.9%). Teens came from households with slightly higher incomes than adult participants (43.3% of teens vs. 26.6% of adults who reported an income live in households with reported income of $100,000 or more; $p = 0.048$, $\theta = 0.178$). Parents were less likely to report household income on behalf of their teens than Prolific crowd workers (21.2% of parents vs. 2.9% of the adults recruited from Prolific did not provide a household income). Prolific crowd workers may be more comfortable sharing income information due to their greater past experience with online surveys.

Unsurprisingly, the reported behavior of adult and teen participants varied in many ways. Most notably, adult participants reported much more previous experience doing online tasks for money (82.9% of adults vs 7.1% of teens reported doing online tasks five times or more; $p < 0.001$, $\theta = 0.824$). They also reported more experience doing online tasks without being paid (67.8% of adults vs 31.8% of teens reported doing tasks for money without being paid at least once; $p < 0.001$, $\theta = 0.386$). This almost certainly results from our adult sample being recruited from Prolific. Adults also reported more previous victimization by online shopping scams (58.0% of adults vs 37.6% of teens reported purchasing something without receiving the item or a refund at least once; $p < 0.003$, $\theta = 0.222$). Adults reported more experience purchasing crypto assets, such as cryptocurrencies or NFTs (34.8% of adults vs 17.6% of teens reported purchasing crypto assets at least once $p < 0.001; \theta = 0.203$).

Teen participants reported more previous usage of both Spotify and Roblox. Teens had much more experience with Roblox (77.3% of teens vs. 27.1% of adults in the Roblox condition had played Roblox at least once; $p < 0.001$, $\theta = 0.549$), but were only slightly more likely to have used Spotify at least once (100% of teens vs. 89.0% of adults in the Spotify condition had used Spotify at least once $p < 0.005$, $\theta = 0.090$). Adult participants were more likely to be frequent users of Spotify. Teen participants were also more likely to have purchased Roblox Robux (38.6% of teens vs. 13.5% of adults in the Roblox condition had purchased Robux; $p < 0.002$, Cramer's V = 0.265). Teens report more weekly usage of mobile games (91.8% of teens vs. 72.2% of adults reported playing mobile games for some time each week; $p = 0.034$ $\theta = 0.240$).

B. RQ1: Participant interactions with giveaway scams

Most participants (84.4%) correctly identified scam stimuli as definitely or probably fraudulent, with 9.2% indicating that the video was definitely or probably legitimate and 6.4% selecting the "I'm not sure" option. Participants' success varied significantly based on which video they saw ($p < 0.004$; Cramer's V = 0.311). 97.8% of participants successfully identified the most easily spotted scam (Spotify Scam 1), while only 68.3% successfully identified the hardest-to-recognize scam (Roblox Scam 1).

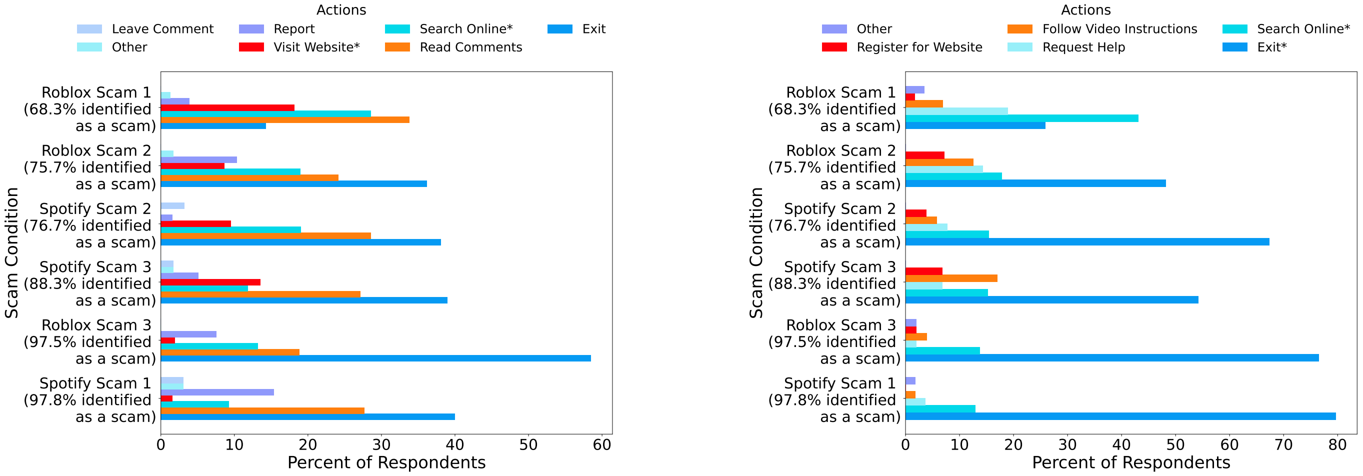
Figure 2 shows what actions participants recommended based on the scam video they saw and the corresponding video of someone visiting the scam website, alongside the percentage of participants who recognized the stimuli as a scam. For the easiest-to-recognize scams, participants were less likely to recommend that their friend visit the website ($p < 0.005$, Cramer's V = 0.289) or search online to learn more about what was shown in the video ($p < 0.005$, Cramer's V = 0.301). Participants viewing the videos of the hardest-to-recognize scam websites were less likely to recommend that their friend exit the website ($p < 0.003$, Cramer's V = 0.360), but were not significantly more or less likely to indicate that their friend should register for the scam website or follow the instructions from the scammer. Rather, users were more likely to want to search online after viewing the website ($p < 0.003$, Cramer's V = 0.326). This suggests that the most obvious scams are easily dismissed by users, while more convincing scams lead users to want to do research to learn more. Across all conditions, only 5.5% of participants recommended that their friend register for the website or follow the instructions from the scam without doing additional research or asking for help.

Many users who suggested visiting the website after viewing the scam video also recommended other actions. For example, 37.1% of participants who recommended that their friend visit the website also suggested that their friend should search online to learn more. In their open-ended responses, many of these users expressed a desire to learn more before proceeding. For example, one adult participant who suggested that their friend should visit the website wrote "I would tell them to spend some time checking the safety of the website before doing any downloads to make sure it is a safe site, in case this is a website or company that is not widely known." Searching online about these scams may prevent victimization, as victims report scams to crowd-sourced review websites like Trustpilot.[4] For example, one of the scam websites we showed participants has 29 negative reviews on TrustPilot as of July 10th, 2024.[5]

More so than searching online, users viewed looking at the comments as a good source of information on other users' experiences. 60% of users who recommended that their friend should visit the website linked in the video also recommended that they should look at the comments. One adult participant explained, "I would check comments to see if it worked for anyone else and if so then visit the website on the video." Similarly, another adult participant wrote, "Comments will determine action to take." While potentially intuitive, relying on comments for information is a bad idea, as the video poster can control what comments are visible or even post fake testimonials to convince others of the scam. Many scam videos featured testimonials in the comments, claiming that the website worked. A few participants recognized the potential unreliability of YouTube comments. For example, one adult participant wrote, "...I would recommend they checked out a

[4] https://www.trustpilot.com

[5] https://www.trustpilot.com/review/boujeecoupons.com



(a) Scam Video

(b) Video of Someone Visiting Scam Website

Fig. 2: Participants' recommended actions in response to the scam stimuli. Participants could select multiple actions. An asterisk in the legend indicates that the rate of recommending this action differed between conditions. The scam conditions are ordered along the y-axis from most to least convincing based on the scam rating task ($n = 260$). Those who did not identify a video as a scam were either unsure or thought the video was legitimate. We use red and orange to indicate the most dangerous actions.

reliable online source about the website (NOT the comments which can be fake)."

For participants who recommended exiting the video or website, the necessity of doing tasks in order to get a reward was the most common feature brought up by participants, with users typically finding these suspicious or not worth the effort. 41.9% of users who recommended that their friend exit the video after viewing the scam video and 31.4% of those who recommended that their friend exit the website after viewing the video of someone visiting the website pointed to the presence of tasks. For example, one teenage participant wrote, after viewing the second Spotify scam, "...having to complete untrustworthy offers seems suspicious." Upon viewing the web video, they added, "Completing online offers most likely never gives you what you want."

Users who mentioned the tasks in their responses did not necessarily realize that the videos they saw were scams. 16.9% of the participants who recommended exiting the video and 8.4% of participants who recommended exiting the website were coded as thinking that following the provided instructions would not be worth the effort. For example, one adult participant who saw the second Roblox scam wrote, "...it is too much hassle for what little bit you get. I'm sure they're getting plenty from you for signing up. Not worth it."

Users reasoning about tasks went beyond just the fear that they would not be compensated or feeling that the compensation was not worth the effort. Some participants (13.2% who recommended exiting the video and 11.5% of those who recommended exiting the website) objected to the need to provide personal information to complete tasks. For example, one adult participant who saw the first Spotify scam video wrote, "If you have to enter or complete other things that you have to input you[r] information I don[']t trust [it]." Some participants pointed to specific security or privacy risks associated with performing the tasks. For example, one teen participant who saw the first Roblox scam video wrote, "Because it says 'Download apps' it's obviously a scam. Everyone with slight common sense should know that downloading random apps and opening links can give you a virus." 6.6% of those who recommended exiting the video and 11% of those who recommended exiting the website were coded as mentioning a security or privacy risk.

Participants often based their decisions on less direct indicators of deception. For example, the second most common feature that participants pointed to was the presenter of the scam video. 22.8% of those who recommended that their friend exit the video mentioned the presenter. The first Spotify scam video, which was one of the easiest to recognize as a scam, is presented by a man in a mask with a concealed voice. Many users found this suspicious; 46.2% of participants who saw this video mentioned the presenter in their justification for their recommended action. One adult participant wrote, "I bet the reason he is dressed like that is... because what he is doing is illegal...." Videos with computer-generated voices rather than real human voices were also viewed as suspicious. One adult participant who viewed the third Spotify scam wrote, "This looks like a scam." They pointed specifically to the fact that it featured an "AI voice over instead of a human" as unconvincing.

The aesthetic quality of the website was another important influence on users' decision-making. 30.4% of those who said that their friend should exit the website pointed to some aesthetic feature. Roblox Scam 3, which was one of the easiest to recognize as a scam, was hosted on a website that was filled with ads that were made to look like download links. 60.4% of participants who saw the video of this website mentioned its poor quality. For example, one adult participant explained their decision to exit the website by stating, “The website is very messy and littered with cheap, scammy ads.” A focus on aesthetic elements can be useful in identifying the most obvious scams, but may not protect users from more polished schemes.

People who fell for the scam (i.e., those who stated that their friend should follow the instructions/register for the scam website) often did not provide specific reasoning beyond the fact that they believed the video. For example, one teen participant who recommenced that their friend follow the instructions in the Roblox Scam 3 wrote, “...if they want their Robux they should do that.” Some participants emphasized that simple instructions were provided. For example, one adult participant who recommended that their friend follow the instructions in Spotify Scam 3 explained their answer by stating, “The steps are easy to follow.” Similarly, a participant who recommended that their friend follow the instructions from the Roblox Scam 2 explained their answer by pointing to the fact that it gave “simple instructions, [and] offer[ed] help.”

Users’ reactions to the legitimate videos may suggest that they were more paranoid than they would be during normal browsing. Only 52% said the legitimate video they saw was probably or definitely legitimate, 12% were unsure, and 36% said it was probably or definitely fraudulent. As with the scam stimuli, participants’ ratings varied significantly based on the video they saw ($p = 0.013$, Cramer’s V = 0.295). 72.2% participants who saw the most convincing legitimate video (the first Roblox video) identified it as legitimate, while only 37.8% of those who saw the least convincing legitimate video (the second Spotify video) were convinced. Lower accuracy in identifying legitimate stimuli as compared to fraudulent stimuli has been noted in other studies, such as Lastdrager et al.’s [49] study of phishing training for children.

Figure 3 shows what actions participants recommended based on the legitimate stimuli they saw. There was less difference between the conditions based on the legitimate web stimuli, with only the rate of participants suggesting that their friend should register for the website varying significantly based on the condition ($p < 0.005$, Cramer’s V = 0.277). This result is unsurprising, as all participants who saw the Roblox stimuli saw the same web video (i.e., the Microsoft rewards website). The first and second Spotify web stimuli showed the same Spotify free trial website.

Participants’ reasoning about the legitimate stimuli was less related to security fears. For example, the most common feature pointed to by people who recommended that their friend exit the video was that the video presented a trial offer and would not be free once the trial period ended. 18.6% of participants who recommended exiting the video mentioned the trial. For example, an adult participant who recommended that his friend exit the third legitimate Spotify video wrote, “...the video is somewhat misleading and is just showing him how to sign up for a free trial, rather than showing ways to get Spotify premium for free.”

For those who believed the legitimate stimuli were presenting a scam, the inability to see URLs was often a factor. Users (reasonably) feared they could be directed to a phishing scheme. For example, a teen participant who saw the second legitimate Spotify video wrote, “...I would avoid going to a direct link provided in a video like this as there are a number of ways it could be malware or a phishing scam.” Presumably, if the participants could see that the link led to the actual Spotify website, they would not have thought that they were being led to a scam.

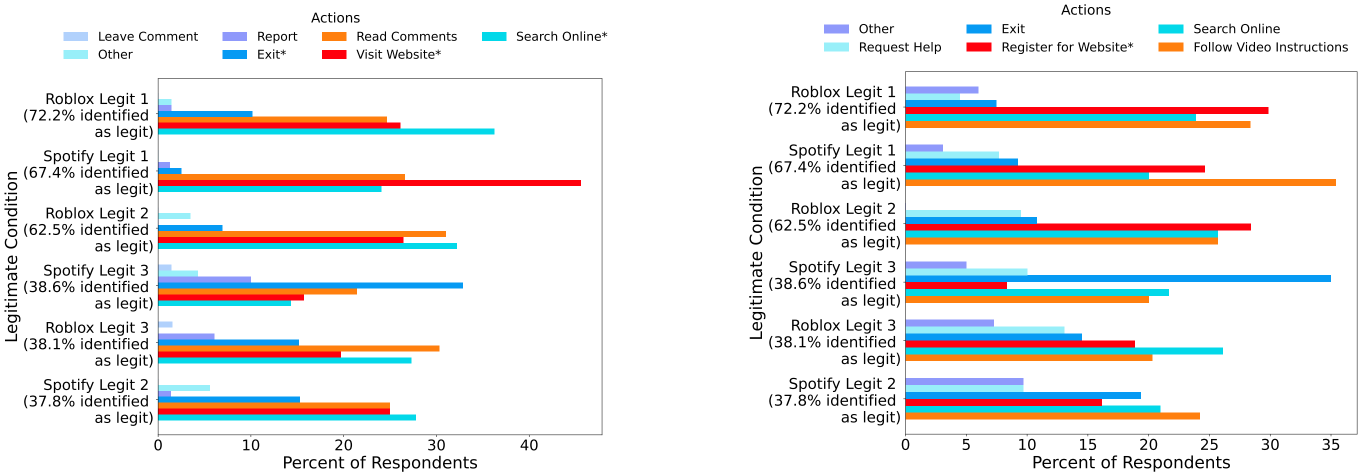
Users generally showed a bias towards legitimate stimuli during the search result selection task. For the Roblox stimuli ($p < 0.001$, $\eta^2_Q = 0.104$), the most frequently selected results were the first and second legitimate videos, which both were selected by 37.4% of participants. The least selected thumbnail was the first scam thumbnail, which was only selected by 7.2%. 64.0% of participants selected at least one legitimate search result, while only 38.1% of participants selected at least one scam search result. For the Spotify stimuli ($p < 0.001$, $\eta^2_Q = 0.0814$), the most selected thumbnail was the first legitimate search result, which was picked by 36.7% of participants. The least selected thumbnail was the third clickbait thumbnail, which was picked by only 6.7% of participants. 57.3% of participants selected at least one legitimate search result, while only 41.3% of participants selected at least one scam search result.

It is unclear why participants gravitated toward selecting legitimate search results. When asked which features of the thumbnail they considered, participants most often indicated the video’s title (64.5%), the thumbnail image (61.7%) and the number of views (30%). The open-ended questions provided little additional information on users’ selections, with participants often just reiterating which parts of the image they looked at. The most common form of reasoning was vague, with users expressing that they found some thumbnails more reasonable than others. For example, one participant who selected the first legitimate Spotify search result and the second Spotify scam search result wrote, “I wanted information that didn’t look shady or underhanded.” For the Roblox search results, some participants seemed to recognize the creators of the legitimate videos. For example, one participant who only selected the first legitimate thumbnail wrote, “The guy in the video is popular in gaming.” The legitimate stimuli generally had a greater number of views as compared to the scam stimuli, which may have guided participants’ selections.

C. RQ2: Differences in teenager and adult behavior

We noted several significant differences between adult and teenage participants. However, none suggested that teens were more susceptible to the scams.

Teen participants were not significantly better or worse at identifying scams or legitimate stimuli. They also did not behave differently during the search results selection task. During the video review task, they were significantly more likely to indicate that their friend should report scam videos than adults (21.2% of teen participants vs. 4.4% of



(a) Legitimate Video

(b) Video of Someone Visiting Legitimate Website

Fig. 3: Participants' recommended actions in response to the legitimate stimuli. An asterisk in the legend indicates that the rate of recommending this action differed between conditions. The legitimate conditions are ordered along the y-axis from most to least convincing based on the scam rating task ($n = 260$). We use the same colors for actions as we did in Figure 2, but here all actions are safe.

adult participants; $p < 0.001$, Cramer's V $= 0.250$). Teen participants who recommended reporting the video seemed to have a desire to protect others. For example, one teen participant who indicated that their friend should report the second Roblox scam video explained that they wanted "to lessen the percentage of people to see and fall for it." Similarly, a teen participant who recommended that their friend report the first Spotify scam video wrote, "They should do this to help others from making bad decisions...."

Very few participants addressed why they did not report a scam video, so it is difficult to draw conclusions about their motivation for not reporting. One possible explanation for the difference in reporting between teens and adults is that adult participants may be more jaded about the result of reporting videos. An adult participant who saw the second Spotify scam wrote, "...I wouldn't bother reporting it because I feel actions like that are fighting an uphill battle that have little or no effect."

Teen participants were also significantly more likely to recommend that their friend ask for help in response to the legitimate website ($p = 0.025$, Cramer's V $= 0.192$). In reasoning why they think their friend should ask for help, teens often pointed to the fact that personal information or a credit card was required. For example, one teen participant who saw the second legitimate Roblox video justified asking for help by saying, "make sure you dont give any information without an adult." Similarly, a teen participant who saw the third legitimate Spotify video wrote, "They may not have a card and [the website] might need to be fact checked."

Concerningly, teen participants were significantly more likely to report previous experience searching for one of the phrases we used to find scams, "free Roblox robux," or something similar ($p < 0.001$, Cramer's V $= 0.385$). 38.6% of teens vs. 6.3% of adults reported previously searching for the term. This is at least partially, if not wholly explained by teens having more exposure to Roblox than adults (see section V-A). Searching for free Robux was more strongly associated with the number of times a person played Roblox (8.3% of participants who played Roblox two times or fewer vs 43.8% of participants who played Roblox three or more times reported searching for free Robux at least once; $p < 0.001$, $\theta = 0.545$) and whether or not participants had purchased Robux before (43.3% of those who had purchased Robux vs. 9.1% who had not purchased Robux had previously searched for free Robux; $p < 0.001$, Cramer's V $= 0.356$). There was no significant relationship between any tested independent variable and past searching for "free Spotify premium" or something similar. 20.7% of participants who saw Spotify stimuli reported previously searching for free Spotify Premium.

As a result of searching for these terms, some participants shared that they saw similar results to those in the survey. For example, one adult participant shared that searching for free Spotify Premium "...showed me a bunch of websites that contained verification steps which never worked in giving me access." Similarly, a teen participant wrote that they found "Another... scammy website that make[s] you do surveys or download apps" after searching for free Roblox Robux. Some participants, especially in the Spotify condition, found genuine offers. For example, one adult participant shared, "I used the 3 month free trial from Paypal." Overall, there were 33 participants who reported negative experiences, 10 neutral or unclear experiences, and 8 positive experiences using similar search terms.

In addition to describing past experiences when explicitly asked, some participants' explanations for recommended actions reflected some level of prior experience with these scams.

For example, one adult who recommended exiting out of the third Roblox scam video said the video was "Definitely one of those generators where you do tasks, like watching ads or filling out a form, to get your Robux that never works." Participants' prior experience also influenced their response to the legitimate videos. One teen user who recommended that their friend visit Microsoft Rewards after viewing the third legitimate Roblox video explained , "...i have actually done this. they give free points towards free robux." Participants' experience could also be misleading. Some participants believed the scam websites were real crowdworking websites. One adult participant who recommended that their friend follow all of the instructions on the second Roblox scam website justified the answer by writing, "Obviously I also use survey sites for extra income." Overall, 14.1% of participants referenced their prior knowledge in some way in their responses.

D. RQ3: Other factors impacting scam victimization

Women were less likely to correctly identify scam videos (92.8% of men vs. 75.8% of women correctly identified scam videos; $p < 0.004$, Cramer's V $= 0.223$), but were not significantly better or worse at correctly identifying legitimate videos. They also did not behave significantly differently in any of the tasks. This could partially result from women having lower confidence in their answers, as almost half (44.8%) of the women who failed to identify the scam video selected "I'm not sure." Women were also more likely than men to select "I'm not sure" when rating the legitimate stimuli; however, this difference was not significant. Prior studies have observed that men will show higher confidence in responses than women [72].

Participants who reported more daily time on digital entertainment activities were more likely to correctly identify scams (90.0% of participants reporting more than 4 hours vs. 77.3% of people reporting less than 4 hours correctly identified scams; $p < 0.004$, $\theta = 0.322$), but were not significantly better or worse at identifying legitimate stimuli. This result may be confounded by the fact that women reported significantly less daily time on digital entertainment activities (56.3% of women vs 38.5% of men reported four hours of digital entertainment per day or less; $p = 0.016$, $\theta = 0.200$). This effect remains weakly significant when looking at women alone ($p_{uncorrected} < 0.006$, $\theta = 0.181$), but not men alone.

We found some statistically significant relationships between the tested independent variables and the probability of selecting particular search results, but there were no factors that were consistently associated with users picking search results. The number of times the participant had previously played Roblox was most frequently associated with search result selection. Users who had played Roblox more frequently were significantly more likely to indicate that they would not recommend that their friend pick any of the search results (17% of those who played Roblox five times or less vs. 45.5% of those who played Roblox more than five times selected no thumbnails; $p < 0.05$, $\theta = 0.294$). These participants may have a greater understanding of the risk of this search term, due to their greater experience searching for "free Roblox Robux" or something similar.

We noted several relationships that are likely spurious. Participants who reported more experience doing online tasks for money were less likely to indicate that they would report the scam video ($p = 0.015$, $\theta = 0.338$) or the legitimate video ($p = 0.022$, $\theta = 0.448$). They were also less likely to report searching for "free Roblox Robux" or something similar ($p < 0.001$, $\theta = 0.520$). As discussed in section V-A, teens were much less likely to have prior experience with online tasks. When looking at teenagers or adults alone, these relationships cease to be significant in either sample. Participants who reported more weekly time playing video games on a computer or game console were more likely to correctly identify scam stimuli ($p = 0.022$, $\theta = 0.275$). Men reported significantly more weekly time spent playing video games on a computer or console than women ($p < 0.001$, $\theta = 0.419$), and the rate of successfully identifying scams does not vary with time spent playing console or computer video games when looking at men or women alone.

VI. Discussion

In this section, we summarize the key results from our experiment and discuss how these results compare to other studies with different types of internet fraud. We then provide recommendations for how to reduce giveaway scams on YouTube and other similar platforms.

A. Overview of key takeaways

We found that most participants were able to recognize that the YouTube giveaway scams were fraudulent, with only 9.2% of participants indicating that they believed the scam they saw was legitimate (RQ1). While most participants came to this conclusion based on the necessity to do tasks, others used potentially unreliable heuristics, such as checking YouTube comments or focusing on aesthetic elements of the scam website. Prior studies on phishing have also observed that users may place too much weight on the perceived aesthetic quality of fraud messages and the presence of elements like company logos [9], [61], [62]. This type of reasoning is inherently subjective and may lead users to trust polished scams and distrust poorly crafted but genuine communications.

We also found no significant difference in scam victimization between adults and teens. Teens were more interested in both of the topics we investigated and more likely to have previously searched for "free Roblox Robux," (RQ2). These results suggest that, while minor teenagers are not necessarily worse at scam identification than adults, their greater interest in the topic may lead them to more often encounter and be victimized by giveaway scams. We noted no significant differences in behavior due to other demographic or experiential factors, although women were slightly less likely to correctly identify scams (RQ3). As discussed in section V-D, this seems more attributable to lower expressed confidence than a lower ability to identify scams among women.

As the requirement to complete tasks was the most salient aspect of the scam stimuli, these results are likely most applicable to other giveaway scams that involve tasks.

B. Recommendations to reduce victimization

Where possible, it is important to prevent users from encountering fraud in the first place. AI methods are often employed to detect specific types of fraud, such as phishing messages [17], in order to block them before they reach a user. YouTube already employs machine-learning-based screening methods to identify and block content that violates their community guidelines [90]. It may be possible to tune these tools to better identify giveaway scam content; however, fraudsters will continually adapt their scams to avoid moderation. One simple approach to prevent users from encountering giveaway scams on YouTube and other video-sharing platforms would be to intentionally block search terms associated with victimization. Some platforms have already employed this approach. Users on TikTok, for example, are blocked from searching for "free Roblox Robux" or "free Robux." Inputting either term into the platform leads to an empty results page that states, "This phrase may be associated with behavior or content that violates our guidelines." This approach is heavy-handed, possibly resulting in some legitimate content being less easily found by users. Moreover, it is impossible to foresee all the potential terms that scammers may use to defraud users. Still, this may be effective for the most commonly abused search terms.

Education can be highly effective at reducing fraud victimization, especially when embedded within fraud messages [76]. While videos that are definitely identified as scams should be removed, YouTube and other video-sharing platforms could add warnings about scams to videos that appear to be ways of earning free goods. YouTube already uses a similar approach to counter misinformation. YouTube automatically adds links to Wikipedia articles to contextualize videos that speak about topics that are prone to misinformation (e.g., vaccines, the moon landing, etc.) [38]. Future work should evaluate how effective these types of interventions are at reducing giveaway scam victimization.

The difference in propensity to report between adults and teens may suggest an opportunity to promote reporting through platform design. Increasing reporting of giveaway scams could allow YouTube to find and eliminate those that are not automatically screened more quickly. For example, YouTube could add "gamified" [27] elements to reward successful reporting, such as giving users a profile badge for submitting a certain number of valid reports. Any feature to promote reporting must be tested to ensure that it does not lead to an increase in false reports.

Looking beyond YouTube, many giveaway scams depend on cost-per-action (CPA) advertising networks that provide scammers with compensation in exchange for having users complete tasks. While CPA can be an ethical way to monetize content, many networks turn a blind eye to fraud or even actively aid scammers. For example, in 2023, WIRED reported that the network CPABuild provided templates for scams [15]. Similarly, the blog of OGAds[6], another CPA network frequently used by scammers, features advice on how to avoid YouTube take-downs [60]. Eliminating these services is challenging, as they are generally run anonymously and can easily change names. CPABuild reincorporated as AdBlue-Media[7] following the scrutiny they received from WIRED's reporting [15]. Still, pressure should be placed on Cloudflare, Amazon Web Services, and other critical infrastructure providers to deplatform these services in order to eliminate scammers' ability to monetize giveaway scams.

VII. Conclusion

We conducted an online experiment evaluating the factors that impact adult and teen users' susceptibility to giveaway scams on YouTube. We began by exploring the landscape of these giveaway scams by analyzing a dataset from 2020-2021, estimating that there are at least 1,199 of these scams available for Roblox Robux alone. We also searched organically to find examples of giveaway scams for obtaining free Roblox Robux and free Spotify Premium for our online experiment. We found that most participants recognized the fraudulent nature of these scams, with only 9.2% of participants identifying them as legitimate. Teens were not more likely to fall for the scams but showed a higher level of familiarity with search terms that could lead to victimization and were more likely to suggest reporting scam videos.

Acknowledgment

This research was supported in part by the Innovators Network Foundation, a Carnegie Mellon University CyLab Presidential Fellowship, and the National Science Foundation under grant CNS-2150217. We would also like to thank Andrea Miller for helping conduct the qualitative investigation of YouTube scams.

References

[1] O. Akgul, R. Roberts, M. Namara, D. Levin, and M. L. Mazurek, "Investigating influencer VPN ads on YouTube," in *2022 IEEE Symposium on Security and Privacy (SP)*. San Francisco, CA: IEEE, May 2022, pp. 876–892. [Online]. Available: https://doi.ieeecomputersociety.org/10.1109/SP46214.2022.9833633

[2] S. Ali, H. A. Haykal, and E. Y. M. Youssef, "Child sexual abuse and the internet—a systematic review," *Human Arenas*, vol. 6, no. 2, pp. 404–421, 2023. [Online]. Available: https://doi.org/10.1007/s42087-021-00228-9

[3] M. I. Alwanain, "How do children interact with phishing attacks?" *International Journal of Computer Science & Network Security*, vol. 21, no. 3, pp. 127–133, 2021. [Online]. Available: http://dx.doi.org/10.22937/IJCSNS.2021.21.3.17

[4] M. Anderson, M. Faverio, and J. Gottfried, "Teens, Social Media and Technology 2023," Pew Research Center, Tech. Rep., 2023. [Online]. Available: https://www.pewresearch.org/internet/2023/12/11/teens-social-media-and-technology-2023/

[5] P. Armitage, "Tests for linear trends in proportions and frequencies," *Biometrics*, vol. 11, no. 3, pp. 375–386, 1955. [Online]. Available: https://doi.org/10.2307/3001775

[6] https://ogads.com

[7] https://adbluemedia.com/

[6] E. Badawi, G.-V. Jourdan, G. Bochmann, I.-V. Onut, and J. Flood, “The “game hack” scam,” in *Proceedings of the 19th International Conference on Web Engineering, (ICWE 2019)*. Daejeon, Republic of Korea: Springer, 2019, p. 280–295. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-19274-7_21

[7] A. Baillon, J. De Bruin, A. Emirmahmutoglu, E. Van De Veer, and B. Van Dijk, “Informing, simulating experience, or both: A field experiment on phishing risks,” *PloS one*, vol. 14, no. 12, p. e0224216, 2019. [Online]. Available: https://doi.org/10.1371/journal.pone.0224216

[8] 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. [Online]. Available: https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

[9] M. Blythe, H. Petrie, and J. A. Clark, “F for fake: four studies on how we fall for phish,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ’11)*. Vancouver, BC, Canada: Association for Computing Machinery, 2011, p. 3469–3478. [Online]. Available: https://doi.org/10.1145/1978942.1979459

[10] E. Bouma-Sims, L. Klucinec, M. Lanyon, L. F. Cranor, and J. Downs, “Recruiting teenage participants for an online security experiment: A case study using peachjar,” in *9th Workshop on Inclusive Privacy and Security (WIPS 2024)*, 2024. [Online]. Available: https://arxiv.org/abs/2408.00864

[11] E. Bouma-Sims and B. Reaves, “A first look at scams on youtube,” in *the 2021 Workshop on Measurements, Attacks, and Defenses for the Web (MADWEB 2021)*, San Diego, CA, Feb. 2021. [Online]. Available: https://dx.doi.org/10.14722/madweb.2021.23001

[12] D. Branley-Bell, L. Coventry, M. Dixon, A. Joinson, and P. Briggs, “Exploring age and gender differences in ict cybersecurity behaviour,” *Human Behavior and Emerging Technologies*, vol. 2022, no. 1, p. 2693080, 2022. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1155/2022/2693080

[13] C. Bravo-Lillo, L. F. Cranor, J. Downs, and S. Komanduri, “Bridging the gap in computer security warnings: A mental model approach,” *IEEE Security & Privacy*, vol. 9, no. 2, pp. 18–26, 2011. [Online]. Available: https://doi.org/10.1109/MSP.2010.198

[14] P. Burda, L. Allodi, and N. Zannone, “Cognition in social engineering empirical research: A systematic literature review,” *ACM Trans. Comput.-Hum. Interact.*, vol. 31, no. 2, Jan. 2024. [Online]. Available: https://doi.org/10.1145/3635149

[15] M. Burgess, “A huge scam targeting kids with roblox and fortnite ‘offers’ has been hiding in plain sight,” *WIRED*, 2023, retrieved from https://www.wired.com/story/poison-pdf-scam-fortnite-roblox/ on February 3, 2025.

[16] M. Button, C. Lewis, and J. Tapley, “Fraud typologies and the victims of fraud: Literature review,” National Fraud Authority, Tech. Rep., 2009. [Online]. Available: https://researchportal.port.ac.uk/en/publications/fraud-typologies-and-the-victims-of-fraud-literature-review

[17] C. Catal, G. Giray, B. Tekinerdogan, S. Kumar, and S. Shukla, “Applications of deep learning for phishing detection: a systematic literature review,” *Knowledge and Information Systems*, vol. 64, no. 6, pp. 1457–1500, 2022. [Online]. Available: https://doi.org/10.1007/s10115-022-01672-x

[18] H. Chen, C. E. Beaudoin, and T. Hong, “Securing online privacy: An empirical test on internet scam victimization, online privacy concerns, and privacy protection behaviors,” *Computers in human behavior*, vol. 70, pp. 291–302, 2017. [Online]. Available: https://doi.org/10.1016/j.chb.2017.01.003

[19] A. Chu, A. Arunasalam, M. O. Ozmen, and Z. B. Celik, “Behind the tube: Exploitative monetization of content on YouTube,” in *31st USENIX Security Symposium (USENIX Security 22)*. Boston, MA: USENIX Association, Aug. 2022, pp. 2171–2188. [Online]. Available: https://www.usenix.org/conference/usenixsecurity22/presentation/chu

[20] J. W. Clark and D. McCoy, “There are no free iPads: An analysis of survey scams as a business,” in *6th USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET 13)*. Washington, D.C.: USENIX Association, Aug. 2013. [Online]. Available: https://www.usenix.org/conference/leet13/workshop-program/presentation/clark

[21] W. G. Cochran, “The comparison of percentages in matched samples,” *Biometrika*, vol. 37, no. 3/4, pp. 256–266, 1950. [Online]. Available: https://doi.org/10.2307/2332378

[22] H. Copes, K. R. Kerley, K. A. Mason, and J. Van Wyk, “Reporting behavior of fraud victims and black’s theory of law: An empirical assessment,” *Justice Quarterly*, vol. 18, no. 2, pp. 343–363, 2001. [Online]. Available: http://dx.doi.org/10.1080/07418820100094931

[23] H. Cramér, *Mathematical methods of statistics*, ser. Princeton Landmarks in Mathematics. Princeton University Press, 1999.

[24] L. F. Cranor, A. L. Durity, A. Marsh, and B. Ur, “Parents’ and teens’ perspectives on privacy in a technology-filled world,” in *10th Symposium On Usable Privacy and Security (SOUPS 2014)*, Menlo Park, CA, Jul. 2014, pp. 19–35. [Online]. Available: https://www.usenix.org/conference/soups2014/proceedings/presentation/cranor

[25] S. Das, L. A. Dabbish, and J. I. Hong, “A typology of perceived triggers for End-User security and privacy behaviors,” in *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*, Santa Clara, CA, aug 2019, pp. 97–115. [Online]. Available: https://www.usenix.org/conference/soups2019/presentation/das

[26] G. Desolda, L. S. Ferro, A. Marrella, T. Catarci, and M. F. Costabile, “Human factors in phishing attacks: A systematic literature reviewd,” *ACM Comput. Surv.*, vol. 54, no. 8, oct 2021. [Online]. Available: https://doi.org/10.1145/3469886

[27] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, “From game design elements to gamefulness: defining ”gamification”,” in *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments (MindTrek ’11)*. Tampere, Finland: Association for Computing Machinery, 2011, p. 9–15. [Online]. Available: https://doi.org/10.1145/2181037.2181040

[28] C. Doherty, J. Kiley, A. Tyson, and B. Jameson, “The whys and hows of generations research,” Pew Research Center, Tech. Rep., 2015. [Online]. Available: https://www.pewresearch.org/politics/2015/09/03/the-whys-and-hows-of-generations-research/

[29] Entertainment Software Association, “Video games remain america’s favorite pastime with more than 212 million americans playing regularly,” 2023. [Online]. Available: www.theesa.com/video-games-remain-americas-favorite-pastime-with-more-than-212

[30] FBI Internet Crime Complaint Center, “Internet crime report,” March 2023. [Online]. Available: https://www.ic3.gov/Media/PDF/AnnualReport/2023_IC3Report.pdf

[31] R. A. Fisher, “On the interpretation of χ^2 from contingency tables, and the calculation of p,” *Journal of the Royal Statistical Society*, vol. 85, no. 1, pp. 87–94, 1922. [Online]. Available: https://doi.org/10.1111/j.2397-2335.1922.tb00768.x

[32] L. C. Freeman, “A further note on freeman’s measure of association,” *Psychometrika*, vol. 41, pp. 273–275, 1976. [Online]. Available: https://doi.org/10.1007/BF02291845

[33] A. K. Ghosh, K. Badillo-Urquiola, S. Guha, J. J. LaViola Jr, and P. J. Wisniewski, “Safety vs. surveillance: What children have to say about mobile apps for parental control,” in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI ’18)*. Montreal QC, Canada: Association for Computing Machinery, 2018, p. 1–14. [Online]. Available: https://doi.org/10.1145/3173574.3173698

[34] F. L. Greitzer, W. Li, K. B. Laskey, J. Lee, and J. Purl, “Experimental investigation of technical and human factors related to phishing susceptibility,” *ACM Transactions on Social Computing*, vol. 4, no. 2, pp. 1–48, 2021.

[35] M. D. Grilli, K. S. McVeigh, Z. M. Hakim, A. A. Wank, S. J. Getz, B. E. Levin, N. C. Ebner, and R. C. Wilson, “Is this phishing? older age is associated with greater difficulty discriminating between safe and malicious emails,” *The Journals of Gerontology: Series B*, vol. 76, no. 9, pp. 1711–1715, 2021. [Online]. Available: https://dx.doi.org/10.1093/geronb/gbaa228

[36] P. Grimm, *Wiley international encyclopedia of marketing*. Chichester, United Kingdom: John Wiley & Sons, Ltd, 2010, ch. Social desirability bias, p. 1.

[37] H. Habib, J. Colnago, V. Gopalakrishnan, S. Pearman, J. Thomas, A. Acquisti, N. Christin, and L. F. Cranor, “Away from prying eyes: Analyzing usage and understanding of private browsing,” in *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*. Baltimore, MD: USENIX Association, Aug. 2018, pp. 159–175. [Online]. Available: https://www.usenix.org/conference/soups2018/presentation/habib-prying

[38] E. Hussein, P. Juneja, and T. Mitra, “Measuring misinformation in video search platforms: An audit study on youtube,” *Proc. ACM Hum.-Comput. Interact.*, vol. 4, no. CSCW1, may 2020. [Online]. Available: https://doi.org/10.1145/3392854

[39] M. Jafari and N. Ansari-Pour, "Why, when and how to adjust your p values?" *Cell Journal (Yakhteh)*, vol. 20, no. 4, p. 604, 2019. [Online]. Available: https://dx.doi.org/10.22074/cellj.2019.5992

[40] H. N. Jeffrey Gottfried, Monica Anderson, "Americans' social media use," Pew Research Center, Tech. Rep., Jan. 2024. [Online]. Available: https://www.pewresearch.org/internet/2024/01/31/americans-social-media-use/

[41] H. Jia, P. J. Wisniewski, H. Xu, M. B. Rosson, and J. M. Carroll, "Risk-taking as a learning process for shaping teen's online information privacy behaviors," in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*. Vancouver, BC, Canada: Association for Computing Machinery, 2015, p. 583–599. [Online]. Available: https://doi.org/10.1145/2675133.2675287

[42] P. G. Kelley, L. F. Cranor, and N. Sadeh, "Privacy as part of the app decision-making process," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. Paris, France: Association for Computing Machinery, 2013, p. 3393–3402. [Online]. Available: https://doi.org/10.1145/2470654.2466466

[43] S. Kemp, "Fraud reporting in catalonia in the internet era: Determinants and motives," *European Journal of Criminology*, vol. 19, no. 5, pp. 994–1015, 2022. [Online]. Available: https://doi.org/10.1177/1477370820941405

[44] K. R. Kerley and H. Copes, "Personal fraud victims and their official responses to victimization," *Journal of Police and Criminal Psychology*, vol. 17, no. 1, pp. 19–35, 2002. [Online]. Available: https://doi.org/10.1007/BF02802859

[45] M. Kezer, B. Sevi, Z. Cemalcilar, and L. Baruh, "Age differences in privacy attitudes, literacy and privacy management on facebook," *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, vol. 10, no. 1, 2016. [Online]. Available: https://doi.org/10.5817/CP2016-1-2

[46] A. Kharraz, W. Robertson, and E. Kirda, "Surveylance: Automatically detecting online survey scams," in *2018 IEEE Symposium on Security and Privacy (SP)*, San Francisco, CA, May 2018, pp. 70–86. [Online]. Available: https://dx.doi.org/10.1109/SP.2018.00044

[47] P. Kumaraguru, J. Cranshaw, A. Acquisti, L. Cranor, J. Hong, M. A. Blair, and T. Pham, "School of phish: A real-world evaluation of anti-phishing training," in *Proceedings of the 5th Symposium on Usable Privacy and Security (SOUPS 2009)*. Mountain View, California, USA: Association for Computing Machinery, 2009. [Online]. Available: https://doi.org/10.1145/1572532.1572536

[48] P. Kumaraguru, Y. Rhee, A. Acquisti, L. F. Cranor, J. Hong, and E. Nunge, "Protecting people from phishing: the design and evaluation of an embedded training email system," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. San Jose, CA: Association for Computing Machinery, 2007, p. 905–914. [Online]. Available: https://doi.org/10.1145/1240624.1240760

[49] E. Lastdrager, I. C. Gallardo, P. Hartel, and M. Junger, "How effective is Anti-Phishing training for children?" in *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*. Santa Clara, CA: USENIX Association, Jul. 2017, pp. 229–239. [Online]. Available: https://www.usenix.org/conference/soups2017/technical-sessions/presentation/lastdrager

[50] E. E. Lastdrager, "Achieving a consensual definition of phishing based on a systematic review of the literature," *Crime Science*, vol. 3, pp. 1–10, 2014. [Online]. Available: https://doi.org/10.1186/s40163-014-0009-y

[51] A. B. M. Laura E Berk, *Infants, Children, and Adolescents*, 8th ed. Pearson, 2016.

[52] T. Lin, D. E. Capecci, D. M. Ellis, H. A. Rocha, S. Dommaraju, D. S. Oliveira, and N. C. Ebner, "Susceptibility to spear-phishing emails: Effects of internet user demographics and email content," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 26, no. 5, pp. 1–28, 2019. [Online]. Available: https://doi.org/10.1145/3336141

[53] S. Livingstone, M. Stoilova, and R. Nandagiri, "Children's data and privacy online: growing up in a digital age: an evidence review," *LSE Research Online*, 2019. [Online]. Available: https://eprints.lse.ac.uk/101283/1/Livingstone_childrens_data_and_privacy_online_evidence_review_published.pdf

[54] A. E. Marwick and danah boyd, "Networked privacy: How teenagers negotiate context in social media," *New Media & Society*, vol. 16, no. 7, pp. 1051–1067, 2014. [Online]. Available: https://doi.org/10.1177/1461444814543995

[55] A. Mathur, J. Vitak, A. Narayanan, and M. Chetty, "Characterizing the use of Browser-Based blocking extensions to prevent online tracking," in *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*. Baltimore, MD: USENIX Association, Aug. 2018, pp. 103–116. [Online]. Available: https://www.usenix.org/conference/soups2018/presentation/mathur

[56] A. McCormac, T. Zwaans, K. Parsons, D. Calic, M. Butavicius, and M. Pattinson, "Individual differences and information security awareness," *Computers in Human Behavior*, vol. 69, pp. 151–156, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0747563216308147

[57] B. McNally, P. Kumar, C. Hordatt, M. L. Mauriello, S. Naik, L. Norooz, A. Shorter, E. Golub, and A. Druin, "Co-designing mobile online safety applications with children," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Montreal QC, Canada: Association for Computing Machinery, 2018, p. 1–9. [Online]. Available: https://doi.org/10.1145/3173574.3174097

[58] J. Nicholson, Y. Javed, M. Dixon, L. Coventry, O. D. Ajayi, and P. Anderson, "Investigating teenagers' ability to detect phishing messages," in *2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*. Virtual: IEEE, 2020, pp. 140–149. [Online]. Available: https://doi.org/10.1109/EuroSPW51379.2020.00027

[59] G. Norris, A. Brookes, and D. Dowell, "The psychology of internet fraud victimisation: a systematic review," *J Police Crim Psych*, vol. 34, pp. 231—-245, 2019. [Online]. Available: https://dx.doi.org/10.1007/s11896-019-09334-5

[60] OGAds, "The takedown problem," 2023, retrieved from https://web.archive.org/web/20231001161836/https://ogads.com/blog/youtube-takedowns on February 3, 2025.

[61] K. Parsons, M. Butavicius, M. Pattinson, D. Calic, A. Mccormac, and C. Jerram, "Do users focus on the correct cues to differentiate between phishing and genuine emails?" in *Proceedings of the Australasian Conference on Information Systems (ACIS 2015)*, Adelaide, SA, Australia, 2016. [Online]. Available: https://aisel.aisnet.org/acis2015/6

[62] K. Parsons, A. McCormac, M. Pattinson, M. Butavicius, and C. Jerram, "Phishing for the truth: A scenario-based experiment of users' behavioural response to emails," in *Proceedings of the 28th IFIP TC 11 International Information Security and Privacy Conference (SEC 2013)*. Auckland, New Zealand: Springer, Jul. 2013, pp. 366–378. [Online]. Available: https://doi.org/10.1007/978-3-642-39218-4_27

[63] M. Pattinson, M. Butavicius, K. Parsons, A. McCormac, and D. Calic, "Factors that influence information security behavior: An australian web-based study," in *Proceedings of the Third International Conference on Human Aspects of Information Security, Privacy, and Trust (HAS 2015)*. Los Angeles, CA: Springer, Aug. 2015, pp. 231–241. [Online]. Available: https://doi.org/10.1007/978-3-319-20376-8_21

[64] M. R. Pattinson, C. Jerram, K. Parsons, A. McCormac, , and M. A. Butavicius, "Managing phishing emails: a scenario-based experiment," in *Proceedings of the Fifth International Symposium on Human Aspects of the Information Security & Assurance (HASIA 2011)*, London, United Kingdom, Jul. 2011. [Online]. Available: https://www.cscan.org/?page=openaccess&eid=7&id=22

[65] D. Raval, "Who is victimized by fraud? evidence from consumer protection cases," *Journal of Consumer Policy*, vol. 44, no. 1, pp. 43–72, 2021. [Online]. Available: https://dx.doi.org/10.1007/s10603-020-09466-w

[66] 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 (CCS '16)*. Vienna, Austria: Association for Computing Machinery, 2016, p. 666–677. [Online]. Available: https://doi.org/10.1145/2976749.2978307

[67] M. D. Reisig and K. Holtfreter, "Shopping fraud victimization among the elderly," *Journal of Financial Crime*, vol. 20, pp. 324 – 337, 2013. [Online]. Available: dx.doi.org/10.1108/JFC-03-2013-0014

[68] V. Rideout, A. Peebles, S. Mann, and M. B. Robb, "The common sense census: Media use by tweens and teens," Common Sense Media, Tech. Rep., 2021. [Online]. Available: https://www.commonsensemedia.org/research/the-common-sense-census-media-use-by-tweens-and-teens-2021

[69] C. Robb and S. Wendel, "Who can you trust? assessing vulnerability to digital imposter scams," *Journal of Consumer Policy*, vol. 46, no. 1, pp. 27–51, 2023. [Online]. Available: https://doi.org/10.1007/s10603-022-09531-6

[70] D. M. Sarno, J. E. Lewis, C. J. Bohil, and M. B. Neider, "Which phish is on the hook? phishing vulnerability for older versus younger adults," *Human factors*, vol. 62, no. 5, pp. 704–717, 2020. [Online]. Available: https://doi.org/10.1177/0018720819855570

[71] D. M. Sarno, J. E. Lewis, C. J. Bohil, M. K. Shoss, and M. B. Neider, "Who are phishers luring?: A demographic analysis of those susceptible to fake emails," in *Proceedings of the human factors and ergonomics society annual meeting*, vol. 61, no. 1. Austin, TX: SAGE Publications, 2017, pp. 1735–1739. [Online]. Available: https://doi.org/10.1177/1541931213601915

[72] H. Sarsons and G. Xu, "Confidence men? gender and confidence: Evidence among top economists," *Harvard University, Department of Economics, Littauer Center*, pp. 1–26, 2015. [Online]. Available: https://scholar.harvard.edu/files/sarsons/files/confidence_final.pdf

[73] R. C. Serlin, J. Carr, and L. A. Marascuilo, "A measure of association for selected nonparametric procedures." *Psychological Bulletin*, vol. 92, no. 3, p. 786, 1982. [Online]. Available: https://doi.org/10.1037/0033-2909.92.3.786

[74] Y. Shang, Z. Wu, X. Du, Y. Jiang, B. Ma, and M. Chi, "The psychology of the internet fraud victimization of older adults: A systematic review," *Frontiers in Psychology*, vol. 13, 2022. [Online]. Available: https://doi.org/10.3389/fpsyg.2022.912242

[75] S. Sheng, M. Holbrook, P. Kumaraguru, L. F. Cranor, and J. Downs, "Who falls for phish? a demographic analysis of phishing susceptibility and effectiveness of interventions," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. Atlanta, GA: Association for Computing Machinery, 2010, p. 373–382. [Online]. Available: https://doi.org/10.1145/1753326.1753383

[76] S. Sheng, B. Magnien, P. Kumaraguru, A. Acquisti, L. F. Cranor, J. Hong, and E. Nunge, "Anti-phishing phil: The design and evaluation of a game that teaches people not to fall for phish," in *Proceedings of the 3rd Symposium on Usable Privacy and Security (SOUPS 2007)*. Pittsburgh, PA: Association for Computing Machinery, 2007, p. 88–99. [Online]. Available: https://doi.org/10.1145/1280680.1280692

[77] J. Tang, E. Birrell, and A. Lerner, "Replication: How well do my results generalize now? the external validity of online privacy and security surveys," in *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*. Boston, MA: USENIX Association, Aug. 2022, pp. 367–385. [Online]. Available: https://www.usenix.org/conference/soups2022/presentation/tang

[78] M. Theofanos, Y.-Y. Choong, and O. Murphy, "'passwords keep me safe' – understanding what children think about passwords," in *30th USENIX Security Symposium (USENIX Security 21)*. Virtual: USENIX Association, Aug. 2021, pp. 19–35. [Online]. Available: https://www.usenix.org/conference/usenixsecurity21/presentation/theofanos

[79] United States Census Bureau, "American comunity survey data releases," 2023, https://www.census.gov/programs-surveys/acs/news/data-releases.2022.html.

[80] United States Department of Agriculture, "Child nutrition programs: Income eligibility guidelines," in *Federal Register, Volume 87*. National Archives and Records Administration, 2022. [Online]. Available: https://www.federalregister.gov/documents/2022/02/16/2022-03261/child-nutrition-programs-income-eligibility-guidelines

[81] United States Federal Trade Commision, "Consumer sentinel network data book 2023," Feb. 2024. [Online]. Available: https://www.ftc.gov/reports/consumer-sentinel-network-data-book-2023

[82] I. Vakilinia, "Cryptocurrency giveaway scam with youtube live stream," in *2022 IEEE 13th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*. New York, NY: IEEE, Oct. 2022, pp. 0195–0200. [Online]. Available: https://doi.org/10.1109/UEMCON54665.2022.9965686

[83] A. van der Heijden and L. Allodi, "Cognitive triaging of phishing attacks," in *28th USENIX Security Symposium (USENIX Security 19)*. Santa Clara, CA: USENIX Association, Aug. 2019, pp. 1309–1326. [Online]. Available: https://www.usenix.org/conference/usenixsecurity19/presentation/van-der-heijden

[84] J. Wang, Y. Li, and H. R. Rao, "Coping responses in phishing detection: an investigation of antecedents and consequences," *Information Systems Research*, vol. 28, no. 2, pp. 378–396, 2017. [Online]. Available: https://doi.org/10.1287/isre.2016.0680

[85] M. Wei, J. Mink, Y. Eiger, T. Kohno, E. M. Redmiles, and F. Roesner, "Sok (or solk?): On the quantitative study of sociodemographic factors and computer security behaviors," in *33rd USENIX Security Symposium (USENIX Security 24)*. Philadelphia, PA: USENIX Association, Aug. 2024. [Online]. Available: https://www.usenix.org/conference/usenixsecurity24/presentation/wei-miranda-solk

[86] M. Wei, E. Zeng, T. Kohno, and F. Roesner, "Anti-Privacy and Anti-Security advice on TikTok: Case studies of Technology-Enabled surveillance and control in intimate partner and Parent-Child relationships," in *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*. Boston, MA: USENIX Association, Aug. 2022, pp. 447–462. [Online]. Available: https://www.usenix.org/conference/soups2022/presentation/wei

[87] H. Whittle, C. Hamilton-Giachritsis, A. Beech, and G. Collings, "A review of online grooming: Characteristics and concerns," *Aggression and Violent Behavior*, vol. 18, no. 1, pp. 62–70, 2013. [Online]. Available: https://doi.org/10.1016/j.avb.2012.09.003

[88] M. T. Whitty, "Do you love me? psychological characteristics of romance scam victims," *Cyberpsychology, Behavior, and Social Networking*, vol. 21, no. 2, pp. 105–109, 2018, pMID: 28657792. [Online]. Available: https://doi.org/10.1089/cyber.2016.0729

[89] R. T. Wright, M. L. Jensen, J. B. Thatcher, M. Dinger, and K. Marett, "Research note: Influence techniques in phishing attacks: An examination of vulnerability and resistance," *Information Systems Research*, vol. 25, no. 2, pp. 385–400, 2014. [Online]. Available: http://www.jstor.org/stable/24700179

[90] YouTube, "How does youtube manage harmful content?" retrieved from https://www.youtube.com/howyoutubeworks/our-commitments/managing-harmful-content/ on February 3, 2025.

[91] C. Zhu, S. Huang, R. Evans, and W. Zhang, "Cyberbullying among adolescents and children: A comprehensive review of the global situation, risk factors, and preventive measures," *Frontiers in Public Health*, vol. 9, 2021. [Online]. Available: https://doi.org/10.3389/fpubh.2021.634909

[92] S. Zhuo, R. Biddle, Y. S. Koh, D. Lottridge, and G. Russello, "Sok: Human-centered phishing susceptibility," *ACM Trans. Priv. Secur.*, vol. 26, no. 3, apr 2023. [Online]. Available: https://doi.org/10.1145/3575797

[93] 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 (CHI' 20)*. Honolulu, HI: Association for Computing Machinery, Aug. 2020, p. 1–15. [Online]. Available: https://doi.org/10.1145/3313831.3376570

Appendix A
Artifact Appendix

A. Description & Requirements

This artifact provides the materials necessary to replicate the behavioral experiment presented in this paper. It also provides the data we collected in the experiment and the analysis code used to obtain statistical results. The artifact is composed of the following parts:

1) **Survey instruments**: The survey instruments used to conduct the study, including the informed consent forms, main experiment survey, and teen compensation survey.
2) **Media**: the pictures and videos presented to participants during the study.
3) **Data**: the data obtained from the experiment and metadata explaining the meaning of each field.
4) **Recruitment materials**: the flier sent to parents of teenage participants and the advertisement posted on Prolific to recruit adult participants.
5) **Analysis code**: the analysis script used to perform statistical testing for the experiment. The analysis script is distributed in the form of an R notebook, with explanations provided alongside the code.
6) **README**: a file explaining the file structure of the artifact.

1) How to access: The artifact is available on Zendodo[8] and Github[9]

2) Hardware dependencies: None

3) Software dependencies: Analysis was performed using R version 4.3.3. However, it is not expected that the version of R will affect the results. The analysis script is distributed in the form of an R Markdown (.Rmd) file, requiring RStudio[10] to view and run. Data is stored in Excel format (.xlsx), which requires Microsoft Excel or compatible spreadsheet software to view. The survey is in Microsoft Word format (.docx), which requires Microsoft Word or compatible document editing software to view. Media is distributed as PNG files for images and MP4 files for videos. Finally, the following readily available R packages are used in the analysis script: `RVAideMemoire`,[11] `dplyr`,[12] `rstatix`,[13] `readxl`,[14] `rcompanion`,[15] `DescTools`[16] and `stringr`.[17]

B. Artifact Installation & Configuration

After installing the dependencies above, no additional configuration steps are necessary to run the analysis script on Windows systems. On non-Windows systems, the format of the path to the data file may need to be changed (i.e., on UNIX systems, the correct relative path is `data/df_analysis.xlsx`). The variable `data_path` in line 41 of `analysis.Rmd` should be updated accordingly.

[8] https://doi.org/10.5281/zenodo.13910629

[9] https://github.com/elijahbs/The-Kids-Are-All-Right

[10] https://posit.co/download/rstudio-desktop/

[11] https://cran.r-project.org/web/packages/RVAideMemoire/index.html

[12] https://cran.r-project.org/web/packages/dplyr/index.html

[13] https://cran.r-project.org/web/packages/rstatix/index.html

[14] https://cran.r-project.org/web/packages/readxl/index.html

[15] https://cran.r-project.org/web/packages/rcompanion/index.html

[16] https://cran.r-project.org/web/packages/DescTools/index.html

[17] https://cran.r-project.org/web/packages/stringr/index.html