Computers in Human Behavior Reports 14 (2024) 100410

Contents lists available at ScienceDirect

Computers in Human Behavior Reports

journal homepage: www.sciencedirect.com/journal/computers-in-human-behavior-reports

Sweeter than honey: Are Gmail accounts associated with greater rewards at a higher risk of hijacking?

Danielle Stibbe [a,b], Stijn Ruiter [a,b,*], Wouter Steenbeek [a], Asier Moneva [a,c]

[a] Netherlands Institute for the Study of Crime and Law Enforcement (NSCR), Netherlands
[b] Universiteit Utrecht, Netherlands
[c] Center of Expertise Cyber Security, The Hague University of Applied Sciences, Netherlands

ARTICLE INFO

Keywords:
Account hijacking
Cybercriminal decision-making
Hacker forums
Honey accounts
Personal data theft
Rational choice perspective
Target selection

ABSTRACT

Objectives: This study investigates the effect of advertised rewards in credential leaks on the likelihood and speed of account hijacking.
Methods: In an online field experiment, we created 176 honey Gmail accounts and randomly assigned them to eight different posts containing account credential leaks. We used a 2 × 2 experimental design, manipulating two key variables within the post titles: the number of accounts (5 K or 1.5 M) and the promise of access to additional platforms (absent or present). We then monitored the accounts for any subsequent activity.
Results: Our findings indicate that the promise of access to additional platforms increased the likelihood and speed of an attempted access. Only 12 accounts were fully accessed, however, because most hijackers did not complete the second-factor authentication (2FA) process required for gaining full access. It seems that the 2FA acted as a deterrent to complete Gmail account hijacking.
Conclusions: The study aligns with the rational choice perspective of crime, showing that the prospect of greater rewards leads to more attempted account accesses.
Pre-registration: https://osf.io/9y26z.

1. Introduction

Technological advances have expanded the landscape of criminal opportunities (Newman & Clarke, 2013), leading to a surge in the frequency and severity of cybercrime (Maimon & Louderback, 2019). In particular, the unauthorized access to personal data, facilitated by the digitization of private information (Wheatley et al., 2016), has become a growing concern. Estimating the losses from personal data breaches is challenging, but reported figures have nearly doubled in the last five years, amounting to approximately $520 Million in 2021 alone (Internet Crime Complaint Center, 2022). In reality, this number is probably much higher, given the substantial underreporting of cybercrime (Cheng, Chau, & Chan, 2018; McMurdie, 2016; Sangari et al., 2022).

One targeted category of personal data that are frequently stolen, exploited, and sold is email credentials, consisting of usernames and passwords. These credentials can be stolen through various means, such as phishing scams (Bursztein et al., 2014; Thomas et al., 2017), malware (Stone-Gross et al., 2009), or data breaches (Chen et al., 2021; Thomas et al., 2017), sometimes exploiting the widespread habit of password reuse (Ives et al., 2004; Missaoui et al., 2018; Poornachandran et al., 2016). Once stolen, credentials can be exploited or sold. If someone tries to exploit the credentials, they can use them to attempt to log in to the account. These attempted accesses can either be successful or fail. Failures could be due to account protection measures like 2-factor authentication (2FA). In this study we distinguish between attempted accesses and full accesses. Alternatively, attackers can share its credentials with others, either for a price or free of charge (Madarie et al., 2019; Thomas et al., 2017). When shared for free, the credentials are

* Corresponding author. Postbus 71304, 1008 BH, Amsterdam, Netherlands.
E-mail address: s.ruiter@uu.nl (S. Ruiter).

https://doi.org/10.1016/j.chbr.2024.100410
Received 8 January 2024; Received in revised form 28 February 2024; Accepted 19 March 2024
Available online 20 March 2024

often posted on surface websites like hacker forums or paste sites.[1] This practice is sometimes executed as a 'proof of content' for a much larger credential leak, where although a large number of credentials is promised, only a small portion is immediately presented (Stone-Gross et al., 2011).

Information stored in email accounts, such as Gmail, often contains details that can be used to access other services. For example, through a Gmail account one can access a wide range of services such as Google Drive, for storing files on the cloud, and Google Calendar, for scheduling appointments, among many others. In addition, some emails contain the login details from other platforms, which could range from streaming and gaming platforms to online shopping or financial services. Therefore, once hijackers gain access to a Gmail account, they can take control of many of the victim's assets. Aside from other functionalities, hijackers may also use the compromised accounts to send spam messages to third parties, search for sensitive information within email correspondences (Onaolapo et al., 2016), or assume the account owner's identity for further exploitation (Bursztein et al., 2014). Gmail accounts thus hold potential intrinsic value for hijackers.

In this study, we tested whether an increase in the implied benefits of Gmail accounts in a credential leak would increase their likelihood of being accessed by hijackers.

1.1. Theory

The notion that perceived benefits increase the likelihood of a target being chosen for a crime derives from the *Rational Choice Perspective*, a theory rooted in the idea that individuals act in ways that maximize their personal benefits while minimizing associated costs (Beccaria, 2009; Becker, 1968; Bentham, 1907; Clarke, 1980, 2016; Clarke & Cornish, 1985; Cornish & Clarke, 2014; Lattimore & Witte, 2017). According to the rational choice perspective, criminal behavior is the result of a cost-benefit analysis in which benefits, like social status or financial gains, are weighed against the costs, like the effort required or the risk involved in committing a crime. Importantly, this form of rational decision-making is bounded by available resources, including time, effort, and available information (Simon, 1990). Situational cues of risks and rewards can also influence decision-making (Clarke, 2016; Cornish & Clarke, 2014). The rational choice perspective has not only been applied to understand why offenders commit crime, but also to criminal target selection (Bernasco et al., 2012; Cornish & Clarke, 2014). When faced with various potential targets, individuals are likely to choose the one they believe will yield the greatest personal gains. The higher the perceived reward, the more likely it is to be selected—provided that all costs remain constant. Note that the rational choice perspective provides a generic framework for understanding offender decision-making (e.g., Clarke, 2016). As such, it is not crime type specific and researchers can apply it to many different types of crime, including cybercrime such as account hijacking. Hence, if a Gmail account is promised to yield increased benefits, it becomes more attractive to account hijackers compared to an account lacking such promises.

Similarly, it can be argued that the higher the perceived value of a target, the sooner it would be selected for a crime. To illustrate this point, consider two shoe stores that are identical in all aspects except for the value of shoes that they sell. If both stores open at the same time, it seems logical that the store offering more expensive shoes is more likely to be selected sooner for shoplifting. Furthermore, it has been argued that the perceived attractiveness of a product is influenced by its novelty and availability (Felson, 1998; Gould, 1969). If we conceptualize account credential information as a hot product (Newman & Clarke, 2013), the more time elapses since it is leaked, the more likely it is to have been exploited by other offenders. In such a case, the perceived benefits are greater for the most recently posted credentials.

While there is limited direct empirical evidence on the effect of utility calculations on the likelihood of target selection for a crime, previous studies provide support for this idea. For instance, Wright and Decker (1997) conducted semi-structured interviews with individuals actively involved in armed robberies and found that when facing financial needs, respondents chose targets perceived as having low risk and high reward. Similarly, Beauregard and Leclerc (2007) interviewed individuals convicted of sex offenses and found a preference for committing crimes in environments with low risk and easily accessible targets. Additionally, Copes (2003) found that individuals who engaged in auto theft mentioned various motivations, including financial gain and hedonism.

However, these studies have several limitations inherent to the use of semi-structured interviews. The interviewed individuals may have been susceptible to memory biases. Interview-based studies are also confined by their ability to reach individuals willing to be interviewed, and limited in their ability to establish causal relationships. To address some of these limitations, researchers have used the hypothetical scenario method in randomized controlled trials. For example, Decker et al. (1993) showed that active residential burglars were more likely to target locations that offered higher monetary rewards and a reduced likelihood of apprehension. Although studies using hypothetical scenario methods can establish causality, they share some of the limitations of interviews, such as biases associated with the presence of researchers, artificial study conditions, and reliance on self-reports, which may limit their generalizability (Delgado-Rodriguez & Llorca, 2004; Exum & Bouffard, 2010).

Researchers have also used spatial analyses to investigate the relationship between crime patterns and the associated costs and benefits. For example, Bernasco et al. (2012) analyzed street robberies in Chicago by examining the costs and benefits associated with the choice of crime location. Their findings revealed that robbers were more likely to commit offenses in easily accessible locations near their homes and in areas where cash was readily available. Similarly, Townsley et al. (2015) found that areas with higher concentration of potential targets were more likely to be selected for burglary. While these studies are based on objective data and avoid potential biases associated with self-report measures, they are observational and therefore limited in their causal inference.

In summary, empirical research supports the hypothesis that the benefits of a target correlate positively with its likelihood to be selected for a crime. However, many studies have limitations related to external validity and/or causal inference. A solution would be to conduct randomized controlled field experiments, but practical and ethical considerations have made these very rare in criminology (Dezember et al., 2021).

1.2. Honeypots: a research opportunity

Technological advances not only expand crime opportunities, but also research opportunities. Honeypots are computer systems designed to mimic real online targets and detect unauthorized interactions that can be used to attract and monitor cybercriminals (Spitzner, 2003). Honeypots not only serve as tools for cybersecurity experts to identify vulnerabilities in their systems and bolster their defenses, but also as a means for researchers to study cybercriminal decision-making in a realistic environment while maintaining a relatively low profile to avoid detection by the perpetrators. For example, Onaolapo et al. (2016)

[1] Paste sites are platforms for sharing text files and code scripts among internet users. They are recognized as a common space for sharing stolen account credentials, and have been used by researchers in credential theft studies (Bermudez Villalva et al., 2018; Bernard-Jones et al., 2018; Bourke & Grzelak, 2018; Madarie et al., 2019; Onaolapo et al., 2016). By using a designate search engine to look for terms related to stolen account credentials, internet users can find a list of such posts, and click on their titles to view their content. A difference between paste sites and hacker forums is that, in the former, there are normally no restrictions on viewing posts.

created a novel infrastructure of honey Gmail accounts and leaked their login credentials online to study cybercriminal behavior. Their findings revealed that intruders primarily sought out emails containing financial information.

Honeypots have been used to examine the effect of costs on criminal decision-making [for a review on the use of honeypots for cybercrime research, see Perkins and Howell (2021)]. For instance, Maimon et al. (2014) conducted a study to determine whether displaying a warning banner after unauthorized access to a honey computer system had a deterrent effect on the intruder. In their study, they presented half of the intruders with a risk warning that explicitly stated the prohibition of intrusion, the system's active monitoring, and the potential involvement of law enforcement. They found that warning banners significantly reduced the duration of intrusions, although they did not affect the frequency of intrusions. Additionally, Wilson et al. (2015) found that intruders executed fewer commands in honeypots where a warning banner was displayed, but only in intrusions lasting longer than 50 s, where the text of the banners was more likely to be processed by intruders.[2]

However, it is essential to acknowledge that these honeypot studies come with some limitations. Possibly the most important methodological criticism is that the honey computer systems used by the authors were incapable of distinguishing between human activity and automated bot activity, which may have affected the validity of the results (Vetterl, 2020). Additionally, it has been argued that sophisticated attackers can differentiate between a genuine target system and a honeypot through various fingerprinting techniques, which raises the possibility that intruders engaging with these honeypots are either inexperienced attackers or, as mentioned earlier, automated bots (Holt, 2017).

In spite of their inherent limitations, and in line with the rational choice perspective, honeypot studies support the notion that the implication of risk can deter criminal behavior. To the best of our knowledge, no prior research has examined the effect of manipulating the implication of rewards on cybercriminal behavior in a randomized controlled field experiment.

1.3. The present study

In this pre-registered study, we aimed to investigate whether advertising a cybercriminal target as being more rewarding would increase criminal activity directed towards that target. To accomplish this, we carried out a field experiment where we deliberately leaked the account credentials of a set of honey Gmail accounts in a series of posts on a hacker forum. When leaking the credentials, we systematically altered the title of the post advertising the account credentials to manipulate the implied reward value of the accompanied Gmail accounts. Subsequently, we monitored all activity in these accounts using an adapted version of the scripts developed by Onaolapo et al. (2016). This allowed us to determine whether our manipulation of rewards influenced the unauthorized accesses of our honey Gmail accounts.

We manipulated the implied rewards of the Gmail accounts in two ways. First, we varied the number of accounts linked to each post ("*5K*" or "*1.5M*"). Second, we manipulated the promise of rewards ("*present*" or "*absent*"), by explicitly mentioning other online platforms associated with the leaked accounts, encompassing services related to gaming, streaming, music, and financial activities. These platforms are inherently beneficial to hijackers, as they provide access to a broader range of features. When an account is promised to include access to these platforms, it is thus more beneficial than an account without such promise. It should be noted that the number of accounts promised in a post constitutes a reward associated with that post, rather than the accounts themselves. Nevertheless, it is important to recognize that individual accounts may be less likely to be fully exploited by other hijackers in larger leaks. In such cases, account owners are also less likely to have noticed the credentials were stolen and change their password. This, in turn, increases the likelihood that the credentials remain functional. Furthermore, certain accounts can possess cumulative value, particularly in terms of financial gain, meaning that more accounts imply more reward.

According to the rational choice perspective, a target is more likely to be selected if it has a higher reward value (Clarke & Cornish, 1985; Cornish & Clarke, 2014). Therefore, if a target is perceived as more attractive, we can expect it to be attacked more often. Consequently, our first hypothesis posited that reward implications would increase the number of unauthorized accesses to our honey Gmail accounts. Specifically, we expected that.

H1a. Accounts in the *1.5M* condition will receive more unauthorized accesses than accounts in the *5K* condition.

H1b. Accounts in posts where a promise of rewards was *present* would experience more unauthorized accesses compared to accounts in posts where the promise was *absent*.

Additionally, we hypothesized that, holding the exposure time of the accounts constant, more attractive targets would be targeted sooner. Thus, we expected that.

H2a. Accounts in the *1.5M* condition would be accessed sooner than accounts in the *5K* condition.

H2b. Accounts in the *reward present* condition would be accessed sooner than accounts in the *reward absent* condition.

2. Method

We preregistered our research question, hypotheses, method, and analysis plan on the Open Science Framework (OSF) before observing any data (anonymized). Our goal was to study whether the implication of rewards associated with email account credentials affects the likelihood of these accounts being accessed, and the speed with which such access occurs. To do this, we created 176 honey Gmail accounts[4][3] and adapted the scripts published by Onaolapo et al. (2016) to monitor account accesses. We then leaked the account credentials on a hacker forum, manipulating the implication of rewards in the post advertising those leaks. Finally, we monitored the activity in each account for 60 days, with the data collection period spanning from August 2nd and October 10th, 2022.

2.1. Credential leaks

We divided the 176 Gmail accounts equally and randomly into four conditions in a 2×2 design, making 44 accounts in each condition. Accounts within each condition were further subdivided in half and leaked in two separate posts. Consequently, there was a total of eight posts, with each post containing 22 account credentials. For more details on account creation, seeAppendix A.

2.1.1. Experimental manipulations

To manipulate the implied rewards associated with our accounts, we varied the title of the posts in each condition for a total of four conditions (Fig. 1). In posts belonging to the *promise of rewards present* conditions, we included a list of other media platforms implied to be linked to our accounts. For example, "PAYPAL, BANK LOGIN, NETFLIX, SPOTIFY,

[2] For a review of the current state of criminological theory application in cyberspace, see Bossler (2019).

[3] This number was initially chosen based on a power calculation conducted for a different analysis, but was retained because it corresponded to the number of accounts created by the time the analysis plan was formulated.

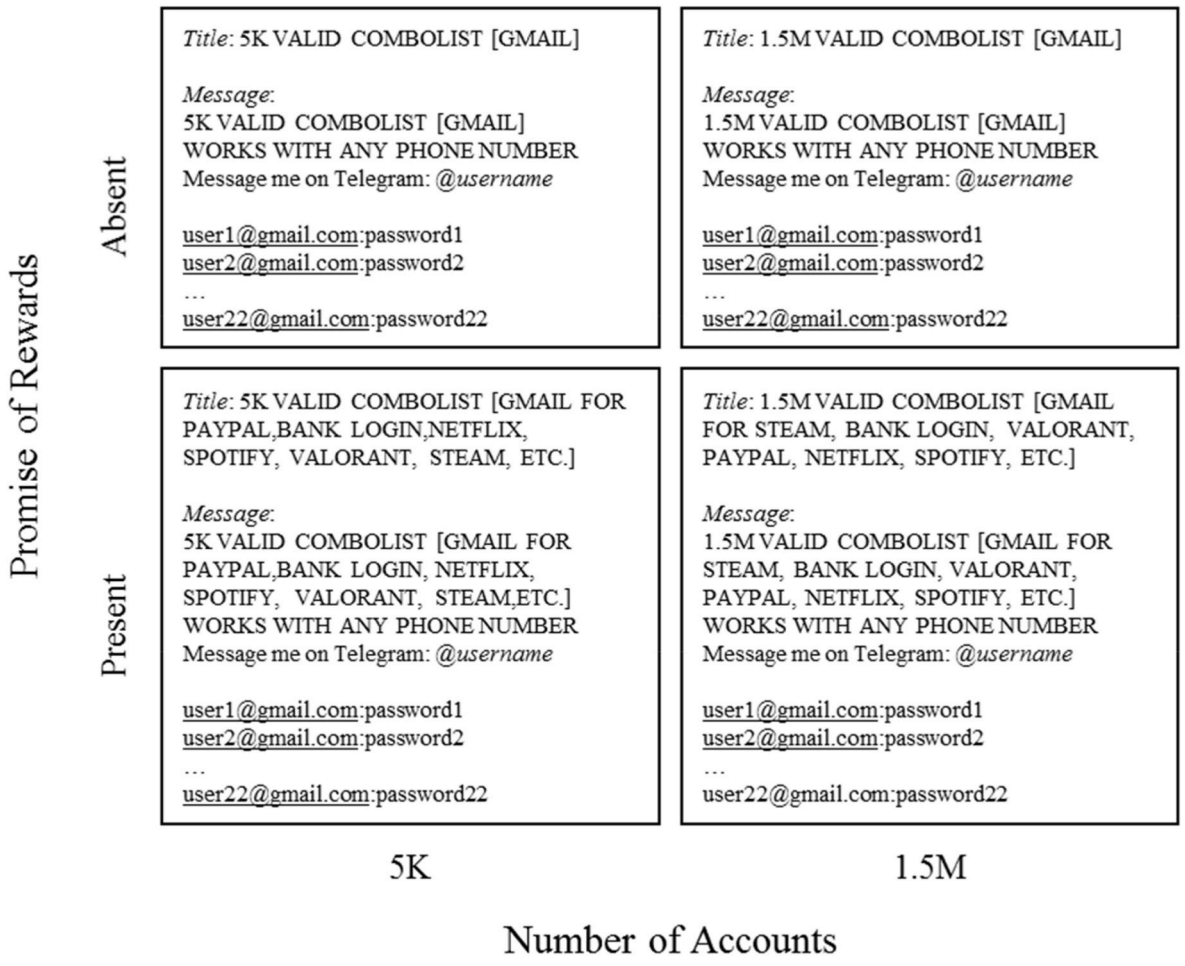

Fig. 1. Example posts for each condition.

VALORANT, STEAM, ETC." To avoid repeating post titles, we randomly shuffled the order of platform names between posts. Posts' content repeated the title and included a note indicating that the accounts "[work] with any phone number" .[4] Finally, to mimic other posts we had observed on the hacking forum, we invited users to request access to the remaining accounts promised by contacting a username that we created specifically for this study on Telegram, but we did not respond to any requests.

2.1.2. Leak platforms

All posts were published on a hacker forum.[5] The hacker forum resembles a conventional forum, featuring tabs that categorize various discussion topics and subtopics, many of which pertain to hacking-related topics, including the leaking of account credentials. Each subtopic presents a list of post titles, along with additional details about each post, such as the author's username and the number of views or replies. Any forum member can create a post, while non-members, can only view posts. We configured our posts to remain hidden, so that only users who had registered on the forum and left a reply, or users who had paid to upgrade their user account on the forum to always be able to see the hidden content, could access them. This acted as a barrier to entry for new forum users, encouraging active participation with our posts while helping to filter out bots and spammers.

2.1.3. Leak schedule

To maximize the exposure to our posts while avoiding the risk of overusing identical post titles in the hacker forum, we divided the credentials in each condition in two and leaked each part at a different time. The first four posts, one of each condition, were randomly assigned to be leaked either in the morning or evening of the same day, resulting in two posts being published on each time frame. The remaining four posts were leaked about a week later, in the opposite time slots they were assigned the week before; posts published in the morning were then published in the evening, and vice versa.

2.1.4. Honeypot infrastructure

The honeypot infrastructure consisted of three components: (1) a Google Apps Script (GAS) embedded within a Google account, (2) a Python monitoring system, and (3) a sinkhole server. The infrastructure can be found on OSF.[6]

[4] We included this note because while we were pilot testing for this study, Google changed its protocol for accessing accounts without recovery phone numbers. Currently, to log into such accounts from a new device, SMS verification is required, which can be done with any phone number. We incorporated this information in the posts because previous studies using the same design or infrastructure (Bermudez Villalva et al., 2018; Bernard-Jones et al., 2018; Onaolapo et al., 2016) were carried out when Google had not yet implemented 2FA, resulting in many more account accesses than we observed during our pilot tests.

[5] To safeguard the continued use of these platforms for future research, we opted to redact the names of these websites.

[6] See https://osf.io/kqtue/?view_only=d80a41f0847b49a39104e74f8a5827c9.

Google App Script. The GAS allows Google account owners to automate actions within their accounts. Using the infrastructure published by Onaolapo et al. (2016),[7] we embedded such a script into all of our honey Google accounts. The script automatically triggered an email to a designated third-party email in our control whenever an email was opened, sent, drafted or starred in the honey Gmail account's mailbox. Additionally, the GAS sent a daily notice indicating that the honey account remained useable for the experiment. If the account's password was changed or if someone disabled the GAS, the account was considered compromised and the notice ceased. Four accounts were compromised in this manner, leading to the suspension of data collection from those accounts. However, the accounts were marked as accessed.

Monitoring system. To track accesses to the accounts, we modified the Python scripts written by Onaolapo et al. (2016) to accommodate changes in Google Chrome and Python. The scripts autonomously logged into the accounts, recording information on accesses. We also added a feature to capture attempted yet incomplete accesses, called "critical security alerts". These alerts were triggered when a user entered correct login credentials on an unknown device, but did not complete the SMS verification process.

Sinkhole Server. For ethical reasons described below, we built a sinkhole mail server based on Onaolapo et al. (2016). Using a Python script, we coded a Simple Mail Transfer Protocol (SMPT) server that acts as a *sinkhole* for incoming emails. The server was designed to reroute all outgoing messages from our honey Gmail accounts. Instead of arriving at their original destinations, the messages were intercepted and saved to disk. This mechanism remained effective even if hijackers disabled the GAS, but not if they disabled the rerouting process.

2.2. Measurements

2.2.1. Accesses to honey accounts

Incomplete accesses. Using our monitoring system, we periodically collected data on critical security alerts received by our honey Gmail accounts, marking the entry of a correct set of credentials without the completion of Google's 2FA. These alerts served as indicators of *incomplete access* to an account the data included a timestamp, a country-level location estimation, and the device, operating system (OS), and browser used. A maximum of three incomplete accesses could be received before Google blocked our accounts and required us to change our passwords.

Full accesses. Our monitoring system also collected data on *full accesses* to our honey accounts. An access, defined as a login event, was recorded individually, even if a hijacker logged in multiple times. The data included a timestamp and unique cookie information for each access, including the Internet Protocol (IP) address, a country-level location estimation, and the device, OS, and browser used. This cookie information served to distinguish unique account hijackers. Given that the accounts were created specifically for this study, any accesses originating from an IP outside our control were deemed unauthorized.

One of our honey Gmail accounts indicated its password had been changed before any information was collected about an unauthorized access to that account, suggesting that a hijacker had changed the account's password. We assumed that the account had been accessed at least once, and counted the moment we lost access to that account as the moment it had been accessed for the first time.

Attempted accesses. To calculate the number of *attempted accesses*, we combined data about full and incomplete accesses.

Time to first attempt/access. We used the timestamp data to determine the first attempted and full access to each account. Subsequently, we calculated the *time to first attempt/access*, representing the number of days between the moment the account credentials were leaked and the first attempt/access.

2.2.2. Number of views

Each time a forum user visited our post, whether by clicking its title or entering its URL manually, contributed to a cumulative count of views. Since this metric did not distinguish between unique viewers, the volume of views could be inflated, for example due to repeated views or automated bots employed by forum users. the total *number of views* for each post was compiled by the end of data collection.

2.2.3. Replies

The content of our posts remained hidden for non-paying forum members, requiring them to log into their account and reply to the post to access its content. Therefore, we assumed that users who replied to a post were interested in its content. To gauge this interest, we counted the *number of replies* to each post at the end of data collection.

2.2.4. Number of requests

Our post titles promised either five thousand or one-and-a-half million accounts, but we only displayed 22. To access more accounts, readers were directed to approach a unique Telegram user assigned to each post. These Telegram accounts were created specifically for this experiment. By counting the contacts made with our Telegram accounts, we calculated the *number of requests* for more accounts.

2.3. Ethical considerations

To uphold ethical standards, we sought advice from the (anonymized). We only started data collection after receiving no ethical objections from the committee. We also followed a strict data protocol in accordance with the European General Data Protection Regulation (GDPR). We took several measures during data collection to address ethical concerns. First, by leaking webmail account credentials, we risked them being abused by hijackers to harm a third party. We mitigated this risk by creating a sinkhole server, to which all outgoing messages from the accounts were rerouted. Those messages were then saved to disk and prevented from being forwarded to their intended destination. Second, by the nature of our study, hijackers were deceived to believe that they were interacting with real account credential leaks. Any time spent interacting with our honey Gmail accounts, hijackers could not spend causing potential harm to the owners of real accounts.

3. Analysis

The following section describes the analyses. For each analysis, we report whether it was preregistered or additional. For all statistical tests, we used an alpha level of 0.05. A complete report of the code used for the analysis, assumption checks, and results following the preregistration can be found on OSF (footnote 6). A complete account of the results of the preregistered analysis, including all model coefficients and the results of sensitivity analyses, can be found in Appendix B.

The account credentials leaked in this study were nested within posts, with 22 credentials in each post. To assess the potential impact of this nesting on the results, we fitted each model with the second level variable *post*. We then tested whether this addition improved model fit, and if it did, we analyzed the multilevel model rather than the single-level version.

For each model, we also examined the presence of influential observations on the estimates by computing DFBETA values for each observation, following the method outlined by Belsley et al. (2005). Observations with DFBETA values greater than $\frac{2}{\sqrt{N}}$, where N is the sample size, were considered influential. To address this, we conducted sensitivity analyses by re-running the analyses after sequentially excluding each influential observation individually.

In this study, 48 of our accounts experienced a total of 58 incomplete

[7] The original monitor infrastructure can be found on https://bitbucket.org/gianluca_students/gmail-honeypot/src/master/.

accesses and 12 accounts were fully accessed 16 times. In total, 57 accounts had experienced a total of 74 attempted accesses. Table 1 describes the number of accounts and probability of attempted, incomplete, and full access in each condition.

3.1. Attempted accesses

The analysis was preregistered. Since the number of incomplete accesses cannot exceed three before an account is blocked by Google, we decided to analyze the *probability* of an account experiencing an attempted access rather than the number of attempted accesses.

To determine whether reward messages influenced the probability of an account experiencing an attempted access, we conducted a logistic regression with *number of accounts, promise of rewards,* and their *interaction* as predictors. Adding *post* as a random effect did improve the fit of the model, $\chi^2(1) = 27.19, p < .001$, so we used a multilevel model for this analysis.

3.2. Time to first attempted access

To test the effect of reward messages on the time to first attempted access, we first computed the survival function for each condition. In this context, 'survival' is defined as the absence of unauthorized attempted accesses to an account. While an account survives, it is 'at risk' of experiencing its first unauthorized attempted access. To predict the risk at a specific point in time we used a Cox proportional hazards models (CPH; Cox, 1972).

Using a CPH model, we tested the effect of the reward conditions on the time Gmail accounts survived until their first attempted access. Since adding post as a second-level factor improved the fit of the model [$\chi^2(1) = 48.74, p < .001$], we used a multilevel CPH model for this analysis.

3.3. Fully accessed accounts

In total, 12 accounts were fully accessed, including four that we lost access to. Following our preregistration, we conducted a logistic regression to test the first two hypotheses. Rather than testing the effect of the reward conditions on the number of accesses per account, we tested their effect on the probability of an account being accessed. To test these hypotheses, we fitted a binomial generalized linear model predicting the probability account access from the number of accounts (*5K* vs. *1.5M*) associated with its post, the promise of rewards (*present* vs. *absent*) associated with the accounts, and the interaction between these two factors.

Table 1
Descriptive statistics of Accesses, incomplete access and total attempted accesses to honey Gmail accounts.

Condition	Full accesses		Incomplete accesses		Attempted accesses	
	Number of accounts	Prob. of event	Number of accounts	Prob. of event	Number of accounts	Prob. of event
5 K accounts						
Promise absent	0	0.00	1	0.02	1	0.02
Promise present	5	0.11	23	0.52	25	0.57
1.5 M accounts						
Promise absent	5	0.11	3	0.03	8	0.18
Promise present	2	0.05	21	0.48	23	0.52

Note. The total number of accounts per condition was 44.

3.4. Time to first full access

To test our last two hypotheses, predicting that our reward conditions would affect the time to a honey Gmail account's first access, we fitted a CPH model as described above. The model predicted the hazard of a full access from our two reward conditions, namely number of accounts and promise of rewards, and their interaction.

4. Results

Based on the figures reported by Onaolapo et al. (2016), we expected our honey Gmail accounts to experience many more accesses than they had. The proportion of accessed accounts in our study (6.8%) was notably lower compared to Onaolapo and colleagues' (2016) study (90%). It is important to note that logistic regression analyses can substantially underestimate the likelihood of a rare event (King & Zeng, 2001), leading to lower analytical power than initially anticipated.

This challenge is further illustrated by the sensitivity analyses. Since few accounts had been hijacked, all accessed accounts were flagged as influential observations. Even when the results of the main analysis indicated a statistically significant effect, excluding any influential observation (with one exception) rendered the effect no longer statistically significant. For this reason, the following section focuses on the main findings of the analysis, and any results of the sensitivity analyses are fully described in Appendix B.

4.1. Attempted accesses

Before observing the data, we pre-registered our intent to examine the effect of reward implication on the likelihood of *attempted access*.

The likelihood of attempted access was positively affected by the *promise of rewards*. When rewards were promised, the odds of an *attempted access* increased almost 137 times ($z = 2.21, p$ .027). In other words, accounts advertised to include access to multiple other platforms were more likely to receive an attempted access than accounts with no such promise. However, excluding one of the influential observations rendered this effect no longer statistically significant.

4.2. Time to first attempted access

While not pre-registered, for comprehensive analysis, we decided to conduct an additional CPH analysis, predicting the time until the first attempted access based on our reward conditions.

As shown in Fig. 2, the hazard of a honey account receiving an attempted access was positively affected by the presence of promised rewards. When rewards were *present* (median survival time = 5 days), the hazard of account access was more than 40 times greater ($z = 2.01$, $p = .045$) than when rewards were not promised (median survival time = 60 days). In other words, accounts promised to be associated with multiple other online platforms received attempted accesses sooner than accounts advertised with on such promise.

4.3. Accessed accounts

The following two pre-registered analyses aimed to test whether a higher number of accounts and the presence of promised rewards would increase the likelihood of an account being accessed, and reduce the time required for the first account access.

Table 1 shows the number of honey accounts accessed in each condition. The results of the logistic regression can be found in Table B1 in Appendix B. Since no statistically significant effect was found, the null hypothesis in H_{1a} and H_{2b} could not be rejected. Thus, we did not find evidence supporting the notion that a higher number of accounts or a promise of reward lead to an increased likelihood of accessing accounts.

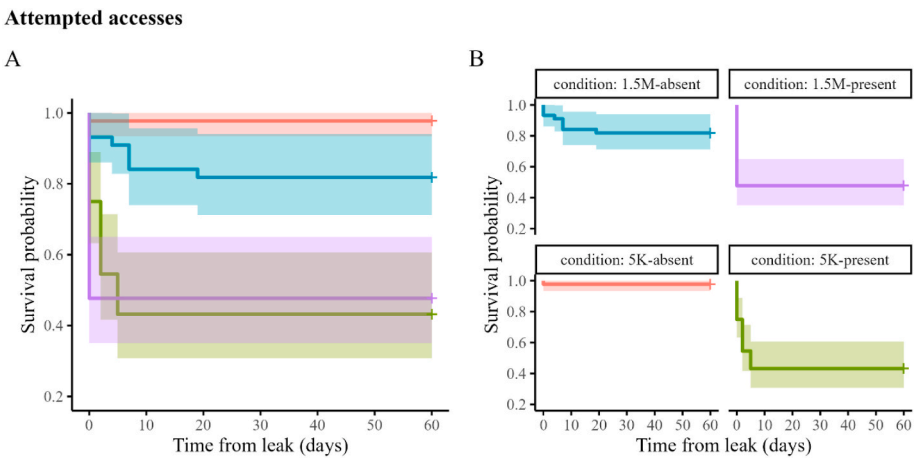

Fig. 2. The survival probability of unauthorized attempted accesses to honey Gmail accounts in each condition as a function of time. The lines symbolize the mean survival probability while the lighter borders illustrate the standard deviation of survival probability within each condition. Fig. 2A represents all conditions, while Fig. 2B represents the different conditions for better visualization.

4.4. Time to first access

Fig. 3 illustrates the survival probability of a honey account not being fully accessed. Since most accounts were not accessed, the median survival time across all conditions was 60 days, corresponding to the data collection period. Similar to the previous analysis, no statistically significant effect was identified. Thus, we did not find evidence to suggest support for hypotheses H_{2a} and H_{2b}.

4.4.1. Views, replies, and requests

Regarding the interaction of users with our posts, Table 2 outlines the number of views, replies, and Telegram requests for more accounts that each post received. Note that two of our posts (one from the *1.5M-absent* condition and the other from the *1.5M-present* condition) were removed from the forum at an unknown date. As this information was collected at the end of data collection, we lack details on the views and replies associated with those two posts, or the duration they were accessible to forum readers.

Although we are unable to statistically analyze the difference between conditions in terms of views and replies, posts that promised rewards had more interactions than posts that did not (197–295 vs. 55–118 views and 14–37 vs. 4–9 replies, accordingly). Notably, the replies-to-views ratio remained constant across all conditions, and the number of requests on Telegram remained close to zero.

Table 2
The number of views, replies, and requests on Telegram to receive more account credentials associated with each post.

Condition	Post number	Views	Replies	Requests
5 K accounts				
Promise absent	3	55	4	0
	5	74	4	0
Promise present	1	232	19	0
	7	295	37	2
1.5 M accounts				
Promise absent	2	118	9	0
	8	*Unknown*	*Unknown*	0
Promise present	4	197	14	1
	6	*Unknown*	*Unknown*	0

5. Discussion

This study examined whether the advertised reward value in posts containing Google account credentials increased their likelihood and speed of access. We experimentally varied the accounts' reward value in the accompanying post titles through two factors: the number of accounts leaked (*5K* vs. *1.5M*) and the promise that the accounts were associated with multiple other platforms (*absent* vs. *present*). Our findings reveal that honey accounts with a promise of rewards were more

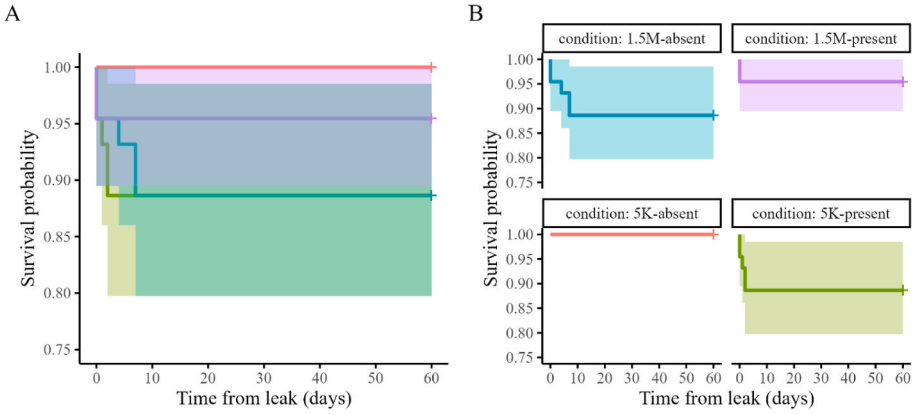

Fig. 3. The survival probability of unauthorized full accesses to honey Gmail accounts in each condition as a function of time. The lines symbolize the mean survival probability while the lighter borders illustrate the standard deviation of survival probability within each condition. Fig. 3A represents all conditions, while Fig. 3B represents the different conditions for better visualization.

prone to attempted accesses and experienced them sooner, even though we saw very few access attempts getting past Google's second factor authentication process (2FA). This supports the notion that the increased effort demanded by 2FA dissuaded potential hijackers from completing unauthorized accesses, as the promise of rewards associated with the accounts was no longer sufficient to outweigh the added effort. This finding highlights the effectiveness of 2FA as a preventive measure against account hijacking. In a similar finding, Google reported that the introduction of 2FA blocked 76% of targeted attacks (Thomas & Moscicki, 2019). While promising for cybersecurity, this posed a challenge in our study, leading to low statistical power to test our hypotheses.

Our accounts received very few full accesses, much fewer than in prior research that used the same design but at a time Google had not yet implemented 2FA (Bermudez Villalva et al., 2018; Bernard-Jones et al., 2018; Onaolapo et al., 2016). We observed only 7% access rate (16 accesses to 12 accounts and another 58 incomplete accesses to 48 accounts), while Onaolapo et al. (2016), for example, recorded a 90% access rate for their accounts.

The target hardening effect of 2FA may explain the overall low likelihood of accessing our honey accounts. The reputation of the 2FA policy among hijacking communities may have dissuaded forum users from attempting to access our accounts. But there are other possible explanations for this low likelihood. Firstly, we only provided a fraction of the promised accounts, a strategy based on actual credential leaks found on a paste site, which may not have been as effective on the hacker forum. Although we observed posts offering a list of free accounts with a link to a Telegram account, we did not encounter the strategy of offering a few accounts while guaranteeing more on another platform. This approach might have been effective with a higher reputation on the forum but could have raised suspicion when originating from novice forum accounts. Indeed, research indicates that the social reputation of vendors on stolen data markets plays a significant role in establishing trust among users (e.g. Décary-Hétu & Laferrière, 2015; Holt, 2013; Holt et al., 2016). The limited exposure of our posts and the removal of two of them could also have contributed to the low access likelihood. Finally, publishing many account credentials together may have led to cross-contamination in the forum users' perception, who may have believed them similar in security measures. Although we used a multilevel regression model to examine this, it did not significantly improve the single-level model fit, likely due to the low likelihood of accesses to all our accounts, resulting in limited statistical power. Therefore, this explanation remains plausible. Regularly scraping information from the posts and posting fewer accounts in more posts could address these issues in future studies.

We found that the number of accounts advertised in the post title did not affect the likelihood or speed of attempted or fully accessing the accounts within that post. Several factors may contribute to this lack of effect. Firstly, users may have harbored suspicions regarding the discrepancy between the advertised and presented account numbers, potentially interpreting it as a scam. If this were the sole reason, we would expect a higher number of views for posts advertising a larger number of accounts, but that was not the case. Alternatively, the removal of two posts (presumably by forum administrators), both belonging to the *1.5M accounts* condition, could explain the lack of effect. This could have been resolved had we identified the removal during data collection, and published new sets of credentials instead. Future studies should consider periodic scraping of forum posts to mitigate such issues. Lastly, it is possible that our manipulation of the number of accounts did not sufficiently influence the perceived reward value of the account credentials. We posited that a larger pool of accounts would be more valuable, as it would be less prone to full exploitation or detection by others. Additionally, we assumed that a higher reward, regardless of its type, would be more enticing. Despite our choice of disparate values for advertised account credentials, forum users might have perceived both 5000 and 1,500,000 as high enough. It is therefore possible that they did not subjectively differentiate the reward value of those numbers.

Though the promise of rewards increased the likelihood and speed of attempted access to our honey Gmail accounts, it did not for full access. A plausible explanation mentioned above relates to a change in Google's security policy. The mandatory implementation of 2FA a few months prior to the experiment (Kudikala, 2021) made our honey Gmail accounts less accessible and potentially less appealing to unauthorized parties. The requirement for supplying a phone number and SMS verification may have dissuaded forum users from completing the access. This change aligns with target hardening techniques in situational crime prevention (Clarke, 1980; Cornish & Clarke, 2014; Ho et al., 2022). Although we informed forum users that they could use any phone number for 2FA, distrust or concern about tracking by Google or authorities might have influenced their behavior. Alternatively, forum users may have misunderstood how the credentials are used, and attempted to use them on the platforms advertised in the post titles instead of to access the Gmail accounts. The advertised platforms may not have been appealing to forum users. Future studies could explore alternative rewards, such as financial value or rich personal information, which were identified by Onaolapo et al. (2016) as the goal of some Gmail hijackers. In this study, we chose the manipulation based on actual forum posts, in order to align with existing content and not raise suspicion.

Our findings raise questions about the target selection strategies employed by potential hijackers when confronted with a list of accounts. One might expect that accounts appearing at the top would be accessed more often than accounts further down the list. However, this was not the case in our sample. Future research could investigate this matter by exploring how the order of appearance in a credential leak affects the likelihood of a compromised account being targeted for hijacking.

Despite the unexpected results, this study provides insights into the crime script of Gmail account hijacking through credential acquisition on the hacker forum. The process involves multiple steps, including visiting the hacker forum, registering or logging in, selecting a post offering account credentials, replying to the post to view its content (paying members can skip this step), selecting an account to hijack, dealing with Google's 2FA, and finally accessing and exploiting the account. While this study manipulated post titles at the post selection step, it did not measure their effect until the last two stages. We observed an effect on attempted accesses, but not on actual accesses. This means that rewards had a greater impact on earlier stages in criminal decision-making compared to later ones. This aligns with the rational choice perspective, which suggests that increasing the effort, like with the inclusion of 2FA, reduces the likelihood of a crime. Future studies examining the effect of rewards on account hijacking in platforms with 2FA should focus on early manipulations and assess if the effect extends beyond the 2FA stage. Questions also arise regarding the decision-making of paying and non-paying members: Do paying members exhibit different tendencies when attempting or completing an access? If the requirement to reply before accessing content is removed, would users invest more effort in the 2FA stage to fully access an account? Addressing these questions would require researchers to access the full server log of the forum to analyze unique interactions and member statuses, which was beyond the scope of this experiment.

5.1. Limitations

In addition to the above, this study has some limitations that should be mentioned. Firstly, the low number of accesses across all conditions led to low statistical power, which prevented a robust statistical analysis of the effect of our reward manipulations on the likelihood and speed of accesses to our honey accounts. This was, however, mainly a result of the study rather than a study design issue. We believe this finding is related to Google's new 2FA policy. Even though we informed forum users about the requirement for a phone number, the 2FA may have still

dissuaded potential hijackers. Future studies could consider using alternative platforms or collaborating with Google to exempt the accounts from the 2FA mandate, thus overcoming this barrier.

Furthermore, the activity observed in the study may have involved automated bots in addition to human interaction. Although the forum's security measures make this scenario unlikely due to extensive use of human verification, sophisticated users or bypass techniques could have allowed bots to alter the forum metrics. However, it is less likely that bots exploited the account lists, as Google's 2FA likely hindered bot attacks (they reported 100% success at preventing those attacks; Thomas & Moscicki, 2019). This offers another explanation for the higher number of attempted accesses compared to completed accesses. Nonetheless, accounts associated with promised rewards experienced a higher number of attempted accesses, indicating a preference for Gmail accounts linked to other platforms regardless of bot involvement. It should be noted that there were suspiciously regular intervals between accesses to one of our accounts on Google's records, suggesting automated activity, but these repetitive accesses only accounted for approximately half of the total accesses to that account. In conclusion, it is possible that accesses had been attempted using bot activity, but those attempts were blocked by Google's 2FA. Nevertheless, the preference for accounts with promised rewards over accounts with no such promise is apparent at the attempt stage, suggesting that rewards influence target selection at least at an early stage.

Regarding the generalizability of the results, it should be noted that since the accounts from which the account credentials were leaked were new, the results only generalize to account credential leaks by new forum users. Future work could explore whether the results change when the accounts that leak credentials have higher levels of trust. The study also faced limitations in drawing conclusions about the popularity of the posts publishing our honey account credentials. While we could rely on usernames to distinguish between those who reply to posts and those who request credentials via Telegram, we could not distinguish those who viewed the posts. We were not able to link views, replies, requests, and incomplete accesses to hijackers who completed their access either, as cookie information was only recorded in the latter case.

These limitations highlight the challenges of conducting field experiments online. Future researchers should strive for larger sample sizes with higher statistical power than what is typically recommended by a power analysis to accommodate any potential challenges that may arise during data collection. Given enough time and perhaps collaboration with the account platform itself (for example, to obtain exemption from the 2FA stage), researchers could improve the study design by including more posts and accounts, allowing for a comprehensive examination of the effect of reward implications on different stages of the crime script.

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This research was funded by the Dutch Research Council (NWO), grant 406.17.562

CRediT authorship contribution statement

Danielle Stibbe: Writing – review & editing, Writing – original draft, Visualization, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Stijn Ruiter:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization. **Wouter Steenbeek:** Supervision, Methodology, Funding acquisition, Conceptualization. **Asier Moneva:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data and code are available through OSF.

Acknowledgements

will be added after review process.

Appendix A. Account creation

For the purpose of this study, we created 176 Gmail accounts. To create these accounts, Google required us to provide certain information in a profile, including a first name, a last name, a password, a username, and a date of birth. To ensure diversity in these accounts, we leveraged the API of a random user generator (RandomAPI, 2017). For the sake of consistency, the profiles were constrained by an age restriction ranging from 18 to 64, and they featured US-based names. Since the passwords generated by the RandomAPI seemed suspiciously similar, we decided to use the leaked password database RockYou instead, which contains over 14 million unique passwords (Mutalik et al., 2021). We filtered passwords to meet the length requirements (a minimum of 8 and a maximum of 16 characters), and randomly sampled 176 passwords using R (R Core Team, 2021). If Google rejected the password, typically because it was easy to guess (e.g., '12,345,678'), we replaced it with a new password.

To ensure that the accounts did not appear empty upon access, we took four steps. First, each profile was assigned three interests appropriate to their age and gender, using stratified randomization. For example, a male born in 1986 could be assigned interests like cooking, sports, and media [a comprehensive list of categories by gender and age group is available in the Supplementary Materials (footnote 6)]. Second, we used Google's search engine to identify 100 US-based newsletters and mailing lists within each of the specified interest categories. Third, we randomly selected 50 newsletters and mailing lists for each profile. And fourth, we subscribed each account was subscribed to those services.

Appendix B. Full statistical analysis as pre-registered

B1. Pre-registered hypothesis testing

B1.2. Accessed accounts

To assess the effect of our manipulations on the likelihood of an account to be accessed, we fitted a logistic regression model. The full model estimates and those from the corresponding sensitivity analyses are described in Table B1. We found no statistically significant effect of *number of accounts* or *promise of rewards* on the number of accessed accounts, nor was the effect of their interaction statistically significant. We therefore could

not reject the null hypotheses or find evidence supporting H_{1a} and H_{1b} in our study.

Table B1
The effect of reward implication (promise of rewards, number of accounts) on full accesses to honey Gmail accounts.

	Model parameters of full sample			Model parameter ranges sequentially removing 12 influential observations					
	B (SE)	Odds Ratio [95% CI]	*p*	B (SE)		Odds Ratio		*p*	
				Min	Max	Min	Max	Min	Max
Intercept	−19.57 (1621.23)	0 [0, inf]	0.990	−19.57 (1621.23)	−19.57 (1621.23)	0	0	0.990	0.990
PRewards (present)	17.51 (1621.23)	4.03e+07 [0, inf]	0.992	17.29 (1621.23)	17.51 (1621.23)	3.23e+07	4.02e+07	0.990	0.991
NAccounts (1.5 M)	17.51 (1621.23)	4.03e+07 [0, inf]	0.992	17.29 (1621.23)	17.51 (1621.23)	3.23e+07	4.02e+07	0.990	0.991
PRewards (present) * NAccounts (1.5 M)	−18.50 (1621.23)	0 [0, inf]	0.991	−19.20 (1621.23)	−18.28 (1621.23)	0	0	0.991	0.991

Note. The analysis was performed using generalized linear logistic regression model using a binomial distribution.
Naccounts = number of accounts implied in the title. PRewards = promise of rewards associated with accounts. CI = confidence interval. SE = standard error. Min = minimum. Max = maximum.

Sensitivity analysis. We identified 12 influential observations [i.e. with a DFBETA for at least one parameter larger than $2/\sqrt{N}$]. Subsequently, we repeated the above analysis 12 times, excluding these influential observations sequentially. The results of the analysis were largely consistent, with the parameter estimates remaining high (see Table B1). Each of those twelve cases represented an accessed account, highlighting the issue of low power in this study. The low number of accounts accessed in this study was so low that each accessed account was identified by DFBETA as an influential case.

B1.3. Time to first access

We fitted a Cox Proportional Hazard (CPH) model to predict the hazard function of an account being accessed from our two reward factors (*promise of rewards*, *number of accounts*) and their interaction. The results of this model are described in Table B2. Neither effect in the model was statistically significant. The average time to first access did not statistically differ between conditions. We could thus not reject the null hypotheses and collect evidence to support hypotheses H_{2a} and H_{2b}.

Sensitivity analysis. We identified two influential observations. The results of the CPH model excluding each of those two influential observations are described in Table B2. Though this exclusion affected the model's coefficients, the p-values remain the same and the conclusion of no statistically significant effects is unchanged.

B2. Preregistered exploratory analysis

Table B2
The effect of reward implication (promise of rewards, number of accounts) on the hazard function of full accesses to honey Gmail accounts.

	Model parameters of full sample			Model parameter ranges sequentially removing 2 influential observations					
	B (SE)	Hazard Ratio [95% CI]	*p*	B (SE)		Hazard Ratio [95% CI]		*p*	
				Min	Max	Min	Max	Min	Max
PRewards (present)	19.81 (8719.32)	4.03e+08 [0, ∞]	0.998	19.99 (9486.73)	19.99 (9486.73)	4.80e+08	4.80e+08	0.998	0.998
NAccounts (1.5 M)	19.80 (8719.32)	3.95e+08 [0, ∞]	0.998	19.97 (9486.73)	19.97 (9486.73)	4.71e+08	4.71e+08	0.998	0.998
PRewards (present) * NAccounts (1.5 M)	−20.74 (8719.32)	0 [0, ∞]	0.998	−21.60 (9486.73)	−21.60 (9486.73)	0	0	0.998	0.998

Note. The analysis was performed using a Cox Proportional Hazards model.
Naccounts = number of accounts implied in the title. PRewards = promise of rewards associated with accounts. Min = minimum. Max = maximum.

B2.2. Attempted accesses

To test whether there was an effect of the reward condition (*number of accounts*, *promise of rewards*) on the probability of an account to experience an attempted access, we fitted a logistic general multilevel model. The results of this analysis are presented in Table B3 When a post included a promise of rewards, the odds of attempted accesses to honey Gmail accounts in this post were nearly 137 times higher, $p = 0.027$. In contrast, the effect of the number of accounts associated with the post was not statistically significant, nor was its interaction with the promise of rewards.

Table B3
The effect of reward implication (promise of rewards, number of accounts) on the probability of attempted accesses to honey Gmail accounts.

	Model parameters of full sample			Model parameter ranges sequentially removing 4 influential observations					
	B (SE)	Odds Ratio [95% CI]	*p*	B (SE)		Odds Ratio		*p*	
				Min	*Max*	*Min*	*Max*	*Min*	*Max*
Intercept	−4.60 (1.80)	0.01 [0, 0.34]	0.011 *	−20.68 (1.92)	−4.70 (332.23)	0	0.01	0.014 *	0.950
PRewards (present)	4.92 (2.22)	136.96 [1.75, 10735.05]	0.027 *	5.02 (2.39)	21 (332.25)	151.14	1.32e+09	0.035 *	0.950
NAccounts (1.5 M)	2.12 (2.29)	8.35 [0.09, 743.31]	0.354	2.11 (2.45)	18.09 (332.24)	8.25	7.18e+07	0.390	0.957
PRewards (present) * NAccounts (1.5 M)	−2.15 (2.97)	0.12 [0, 39.64]	0.470	−18.09 (3.22)	−1.46 (332.26)	0.02	0.23	0.454	0.957

Note. The analysis was performed using logistic generalized linear multilevel model using a binomial distribution. Naccounts = number of accounts implied in the title. PRewards = promise of rewards associated with accounts. SE = standard error. Min = minimum. Max = maximum.
*p < 0.05.
**p < 0.01.
***p < 0.001.

Sensitivity analysis. We identified 4 influential observations. The results of the analyses excluding each observation are also presented in Table B3. For two of the four influential observations, the exclusion had a substantial impact, entirely altering the results and rendering the effect of *promise of reward* no longer statistically significant.

References

Beauregard, E., & Leclerc, B. (2007). An application of the rational choice approach to the offending process of sex offenders: A closer look at the decision-making. *Sexual Abuse, 19*(2), 115–133. https://doi.org/10.1177/107906320701900204

Beccaria, C. (2009). *On crimes and punishments and other writings*. Transaction Publishers. https://doi.org/10.1017/cbo9780511802485.006 (Original work published 1764).

Becker, G. S. (1968). Crime and punishment: An economic approach. In *The economic dimensions of crime* (pp. 13–68). Springer. https://doi.org/10.1007/978-1-349-62853-7_2.

Belsley, D. A., Kuh, E., & Welsch, R. E. (2005). *Regression diagnostics: Identifying influential data and sources of collinearity*. John Wiley & Sons. https://doi.org/10.1002/0471725153

Bentham, J. (1907). *An introduction to the principles of morals*. Clarendon Press (Original work published 1789).

Bermudez Villalva, D. A., Onaolapo, J., Stringhini, G., & Musolesi, M. (2018). Under and over the surface: A comparison of the use of leaked account credentials in the dark and surface web. *Crime Science, 7*(1), 1–11. https://doi.org/10.1186/s40163-018-0092-6

Bernard-Jones, E., Onaolapo, J., & Stringhini, G. (2018). *BABELTOWER: How Language Affects Criminal Activity in Stolen Webmail Accounts*, 991–999. https://doi.org/10.1145/3184558.3191529

Bernasco, W., Block, R., & Ruiter, S. (2012). Go where the money is: Modeling street robbers' location choices. *Journal of Economic Geography*, 119–143. https://doi.org/10.1093/jeg/lbs005

Bourke, D., & Grzelak, D. (2018). Breach detection at scale with aws honey tokens. *Blackhat Asia*, 20–23. https://i.blackhat.com/briefings/asia/2018/asia-18-bourke-grzelak-breach-detection-at-scale-with-aws-honey-tokens-wp.pdf.

Bursztein, E., Benko, B., Margolis, D., Pietraszek, T., Archer, A., Aquino, A., … Savage, S. (2014). Handcrafted fraud and extortion: Manual account hijacking in the wild. In *Proceedings of the 2014 conference on internet measurement conference* (pp. 347–358). https://doi.org/10.1145/2663716.2663749

Chen, D., Chowdhury, M. M., & Latif, S. (2021). Data breaches in corporate setting. In *2021 international conference on electrical, computer, communications and mechatronics engineering (ICECCME)* (pp. 1–6). https://doi.org/10.1109/ICECCME52200.2021.9590974

Cheng, C., Chau, M. C.-L., & Chan, M. L. (2018). A social psychological analysis of the phenomenon of underreporting cybercrimes and the concomitant underlying factors: Three real local case studies. *Communications Association of Hong Kong*, 59–66.

Clarke, R. V. (1980). Situational crime prevention: Theory and practice. *Brit. J. Criminology, 20*, 136.

Clarke, R. V. (2016). Situational crime prevention. In *Environmental criminology and crime analysis* (pp. 305–322). Routledge. https://doi.org/10.4324/9781315709826.

Clarke, R. V., & Cornish, D. B. (1985). Modeling offenders' decisions: A framework for research and policy. *Crime and justice, 6*, 147–185. https://doi.org/10.4324/9781439817803-17

Copes, H. (2003). Streetlife and the rewards of auto theft. *Deviant Behavior, 24*(4), 309–332. https://doi.org/10.1080/713840224

Cornish, D. B., & Clarke, R. V. (2014). The reasoning criminal: Rational choice perspectives on offending. https://doi.org/10.4324/9781315134482.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B, 34*(2), 187–202.

Décary-Hétu, D., & Laferrière, D. (2015). Discrediting vendors in online criminal markets. In *Disrupting criminal networks: Network analysis in crime prevention* (pp. 129–152).

Decker, S., Wright, R., & Logie, R. (1993). Perceptual deterrence among active residential burglars: A research note. *Criminology, 31*(1), 135–147. https://doi.org/10.1111/j.1745-9125.1993.tb01125.x

Delgado-Rodriguez, M., & Llorca, J. (2004). Bias. *Journal of Epidemiology & Community Health, 58*(8), 635–641. https://doi.org/10.1136/jech.2003.008466

Dezember, A., Stoltz, M., Marmolejo, L., Kanewske, L. C., Feingold, K. D., Wire, S., … Maupin, C. (2021). The lack of experimental research in criminology—evidence from Criminology and Justice Quarterly. *Journal of Experimental Criminology, 17*, 677–712.

Exum, M. L., & Bouffard, J. A. (2010). Testing theories of criminal decision making: Some empirical questions about hypothetical scenarios. In *Handbook of quantitative criminology* (pp. 581–594). Springer. https://doi.org/10.1007/978-0-387-77650-7_28.

Felson, M. (1998). *Crime and everyday life* (2nd ed.). Pine Forge Press.

Gould, L. C. (1969). The changing structure of property crime in an affluent society. *Social Forces, 48*(1), 50–59. https://doi.org/10.2307/2575468

Ho, H., Ko, R., & Mazerolle, L. (2022). Situational crime prevention (SCP) techniques to prevent and control cybercrimes: A focused systematic review. *Computers & Security, 115*, Article 102611. https://doi.org/10.1016/j.cose.2022.102611

Holt, T. J. (2013). Examining the forces shaping cybercrime markets online. *Social Science Computer Review, 31*(2), 165–177. https://doi.org/10.1177/0894439312452998

Holt, T. J. (2017). On the value of honeypots to produce policy recommendations. *Criminology & Pub. Pol'y, 16*, 737. https://doi.org/10.1111/1745-9133.12315

Holt, T. J., Smirnova, O., & Hutchings, A. (2016). Examining signals of trust in criminal markets online. *Journal of Cybersecurity, 2*(2), 137–145. https://doi.org/10.1093/cybsec/tyw007

Internet Crime Complaint Center. (2022). Internet crime report 2021. https://www.ic3.gov/Media/PDF/AnnualReport/2021_IC3Report.pdf.

Ives, B., Walsh, K. R., & Schneider, H. (2004). The domino effect of password reuse. *Communications of the ACM, 47*(4), 75–78. https://doi.org/10.1145/975817.975820

King, G., & Zeng, L. (2001). Logistic regression in rare events data. *Political analysis, 9*(2), 137–163.

Kudikala, C. (2021). Google to mandate two-step verification from november 9. November 3 https://telecomtalk.info/google-to-mandate-two-step-verification-from/477288/.

Lattimore, P., & Witte, A. (2017). Models of decision making under uncertainty: The criminal choice. In *The reasoning criminal* (pp. 129–155). https://doi.org/10.4324/9781315134482-9

Madarie, R., Ruiter, S., Steenbeek, W., & Kleemans, E. (2019). Stolen account credentials: An empirical comparison of online dissemination on different platforms. *Journal of Crime and Justice, 42*(5), 551–568. https://doi.org/10.1080/0735648x.2019.1692418

Maimon, D., Alper, M., Sobesto, B., & Cukier, M. (2014). Restrictive deterrent effects of a warning banner in an attacked computer system. *Criminology, 52*(1), 33–59. https://doi.org/10.1111/1745-9125.12028

Maimon, D., & Louderback, E. R. (2019). Cyber-dependent crimes: An interdisciplinary review. *Annual Review of Criminology, 2*(1), 191–216. https://doi.org/10.1146/annurev-criminol-032317-092057

McMurdie, C. (2016). The cybercrime landscape and our policing response. *Journal of Cyber Policy, 1*(1), 85–93. https://doi.org/10.1080/23738871.2016.1168607

Missaoui, C., Bachouch, S., Abdelkader, I., & Trabelsi, S. (2018). Who is reusing stolen passwords? An empirical study on stolen passwords and countermeasures. In *Cyberspace safety and security: 10th international symposium, CSS 2018, amalfi, Italy, october 29–31, 2018, proceedings* (pp. 3–17). https://doi.org/10.1007/978-3-030-01689-0_1, 10.

Mutalik, R., Chheda, D., Shaikh, Z., & Toradmalle, D. (2021). Rockyou. https://doi.org/10.21227/gzcg-yc14.

Newman, G. R., & Clarke, R. V. (2013). Superhighway robbery. *Willan*. https://doi.org/10.4324/9781843924876

Onaolapo, J., Mariconti, E., & Stringhini, G. (2016). What happens after you are pwnd: Understanding the use of leaked webmail credentials in the wild. In *Proceedings of the 2016 internet measurement conference* (pp. 65–79). https://doi.org/10.1145/2987443.2987475

Perkins, R. C., & Howell, C. J. (2021). Honeypots for cybercrime research. In *Researching cybercrimes: Methodologies, ethics, and critical approaches* (pp. 233–261). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-74837-1_12.

Poornachandran, P., Nithun, M., Pal, S., Ashok, A., & Ajayan, A. (2016). Password reuse behavior: How massive online data breaches impacts personal data in web. In *Innovations in computer science and engineering: Proceedings of the third ICICSE, 2015* (pp. 199–210). https://doi.org/10.1007/978-981-10-0419-3_24

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Version 4.2. .

RandomAPI. (2017). *Random user generator*. Retrieved 2021-09-18 from https://randomuser.me/api.

Sangari, S., Dallal, E., & Whitman, M. (2022). Modeling under-reporting in cyber incidents. *Risks, 10*(11), 200. https://doi.org/10.3390/risks10110200

Simon, H. A. (1990). Bounded rationality. *Utility and probability*, 15–18.

Spitzner, L. (2003). *Honeypots: Tracking hackers* (Vol. 1). Reading: Addison-Wesley.

Stone-Gross, B., Cova, M., Cavallaro, L., Gilbert, B., Szydlowski, M., Kemmerer, R., … Vigna, G. (2009). Your botnet is my botnet: Analysis of a botnet takeover. In *Proceedings of the 16th ACM conference on Computer and communications security* (pp. 635–647). https://doi.org/10.1145/1653662.1653738

Stone-Gross, B., Holz, T., Stringhini, G., & Vigna, G. (2011). The underground economy of spam: A botmaster's perspective of coordinating large-scale spam campaigns. *LEET, 11*, 4, 4.

Thomas, K., Li, F., Zand, A., Barrett, J., Ranieri, J., Invernizzi, L., … Moscicki, A. (2017). Data breaches, phishing, or malware? Understanding the risks of stolen credentials.

In *Proceedings of the 2017 ACM SIGSAC conference on computer and communications security* (pp. 1421–1434). https://doi.org/10.1145/3133956.3134067

Thomas, K., & Moscicki, A. (2019). *New research: How effective is basic account hygiene at preventing hijacking*. Google security blog. May 17, 2019 https://security.googleblog.com/2019/05/new-research-how-effective-is-basic.html.

Townsley, M., Birks, D., Bernasco, W., Ruiter, S., Johnson, S. D., White, G., & Baum, S. (2015). Burglar target selection: A cross-national comparison. *Journal of Research in Crime and Delinquency, 52*(1), 3–31. https://doi.org/10.1177/0022427814541447

Vetterl, A. (2020). *Honeypots in the age of universal attacks and the Internet of Things* [Doctoral dissertation, University of Cambridge].

Wheatley, S., Maillart, T., & Sornette, D. (2016). The extreme risk of personal data breaches and the erosion of privacy. *The European Physical Journal B, 89*, 1–12. https://doi.org/10.1140/epjb/e2015-60754-4

Wilson, T., Maimon, D., Sobesto, B., & Cukier, M. (2015). In *The effect of a surveillance banner in an attacked computer system: Additional evidence for the relevance of restrictive deterrence in cyberspace* (pp. 829–855). https://doi.org/10.1177/0022427815587761, 52.

Wright, R. T., & Decker, S. H. (1997). In *Armed robbers in action: Stickups and street culture. UPNE*. https://doi.org/10.2307/2653906