






# Not All Victims Are Created Equal: Investigating Differential Phishing Susceptibility

Matthew Canham<sup>1</sup> , Shanée Dawkins<sup>2</sup> , and Jody Jacobs<sup>2</sup> 

<sup>1</sup> Quantum Improvements Consulting, Orlando, FL 32803, USA  
mcanham@quantumimprovements.net

<sup>2</sup> National Institute of Standards and Technology (NIST), Gaithersburg, MD 20899, USA

**Abstract.** Repeat clickers refer to individuals who repeatedly fall prey to phishing attempts, posing a disproportionately higher risk to the organizations they inhabit. This study sought to explore the potential influence of three factors on repeat clicking behavior. First, building from previous research, we examined the impact of individual characteristics such as personality traits (Big 5 and Locus of Control), expertise (security and phishing knowledge), and technology usage. Second, social engineering tactics were considered as a potential factor, based on the specifications of the NIST Phish Scale, a metric for rating an email's human phishing detection difficulty. Third, the impact of contextual factors, such as world events, were investigated. Data was collected from study participants via a survey on their individual differences, followed by campaigns in which they were emailed a total of eight messages (four phishing and four controls) over a four-week period of time. Repeat clickers were found to spend less time working online, check email more often, have a more internally oriented locus of control, and a lower need for cognition than the comparison groups. The Phish Scale resulted in difficulty scores closely corresponding to observed click-rates in phishing emails, suggesting that it is an effective metric of evaluating human phishing detection difficulty in a university environment.

**Keywords:** Repeat Clickers · NIST Phish Scale · Phishing Susceptibility · Security Awareness · Human-centered Cybersecurity

## 1 Introduction

Phishing is a form of email-based social engineering attack [1, 2], which continues to pose the most pressing security challenge to the human attack surface [3]. Repeat clickers, a subset of individuals who repeatedly fall prey to phishing attempts, pose a disproportionately higher risk to the organizations they inhabit [2, 4–8]. This study sought to clarify some of the contributing factors underlying repeat clicking behavior.

The ground truth in real-world phishing attacks can be extremely difficult (if not impossible) to establish, because cyber criminals often undertake significant measures to conceal their presence on a network [2]. Simulated phishing exercises, which send mock phishing email campaigns, are a form of training intended to inoculate users against

phishing susceptibility and help them recognize phishing attacks [9]. While the efficacy of these simulations has been called into question [10], these training exercises currently provide the best source of data for understanding real-world phishing susceptibility. Details on both detrimental user actions, including clicking an embedded hyperlink, downloading an attachment, or replying to the sender, as well as beneficial user actions, such as reporting the simulated phish, are also recorded. Simulated phishing campaigns can serve as effective proxies for studying real-world phishing attacks because they closely mimic actual attacks (sometimes the emails are neutered copies of real attacks), and users are not typically warned in advance that a simulated email campaign is about to be launched.

Here, in Sects. 1.1 through 1.3, we highlight the need for understanding the repeat clicker phenomenon and the underlying causes in their behavior. Section 1.4 goes into deeper detail about the NIST Phish Scale metric and its purpose for use in this investigation.

### 1.1 Differing Patterns of Phishing Susceptibility

Research studies examining these simulated phishing campaigns find that “failures” (clicking a link, responding to the sender, or entering credentials) are Pareto distributed, meaning that most users fail a maximum of one or two campaigns, but that a subset fail three or more within a given timeframe. One study found that this subset, the ‘repeat clickers,’ present approximately three times the risk exposure of other users [5]. These findings were corroborated by another study that observed that a small portion of employees fell for phishing emails multiple times in a similar “long-tail” Pareto distribution of simulation failures. Understanding why these users present a significantly elevated risk exposure is critical to reducing the overall human attack surface for an organization [8].

Little scientific research has examined the underlying causes of this repeat clicker phenomenon, with most studies being industry reports on quantity of occurrence but little explication of user characteristics [10, 11]. Li et al. [7] found that the best predictor of phishing susceptibility was previously falling prey to a phishing email, and that repeat clickers did not respond to training interventions. As part of a study focused on the efficacy of phishing training interventions, researchers explored repeat clickers peripherally, identifying three clicking patterns which they termed as: all-clickers, non-clickers, and everyone else. The researchers found that the all-clickers failed the phishing simulations due to “an interest in the subject matter and lack of careful attention” [4]. Interestingly, most had no memory of identifying anything suspicious in the emails. The all-clickers performance did not improve in response to the training interventions introduced by the researchers. These patterns of responses may suggest that these actions might result from default habituated responses to email, rather than being driven by security knowledge.

### 1.2 Working to Unravel the Mystery of Repeat Clickers

A key consideration in understanding phishing victimization hinges on discerning the individual differences between persons who occasionally fall prey to phishing emails, and those who repeatedly fall prey to them. Within the context of this study, individual

traits refer to personality traits such as locus of control (LOC), a need for cognition, the Big 5, expertise with information technology and security, or prior victimization.

LOC describes the degree to which an individual feels that they can affect change or direct the course of their own fate [15]. The need for cognition refers to an individual's intrinsic need to question and evaluate the information that they are being provided in relation to a message, which may impact their critical evaluation of phishing messaging. The Big 5 personality dimensions (Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism) [28] have been found to influence phishing susceptibility in other studies [29–31] suggesting that certain personality dimensions may predict repeat clicking. While the impact of information technology expertise on phishing susceptibility has been mixed [30–32] some suggest that this may play a role in repeat clicking [6]. Prior victimization by online scammers may also predict repeat clicking behavior [6].

Personality factors. The persistent nature of repeat clicking suggests that personality factors may play a role in driving this behavior [2]. In fact, in the early twentieth century, researchers suggested that some individuals are more susceptible to accidents than other individuals in part due to personality traits [12]. Interestingly, this higher susceptibility to accidents appear to also follow a Pareto distribution [13]. As a result, some suggested that an “accident prone” personality type might exist, although this hypothesis was later discredited [14]. The more likely explanation is that a combination of relevant personality traits and individual differences, such as LOC, collectively influence accident proneness [14]. An internally oriented LOC implies that the individual believes that they can direct life events and steer their life direction. An externally oriented LOC implies that the individual believes that they are powerless to affect change and that events will happen as they are meant to, regardless of their own level of effort. Some researchers have speculated a relationship between an externally oriented LOC and increased accident proneness [14]; several studies support this relationship [16–22].

While there does appear to be a relationship between LOC and accident proneness, studies investigating the relationship between LOC and security behaviors have had mixed results. One study found no relationship between LOC and likelihood of clicking the hyperlink in a phishing email [23]; however, this was measured through a self-report survey, which are not always good proxies for actual behavior [24]. Another study found that individuals who demonstrated a more externally oriented LOC had weaker engagement with information security policies in the workplace [25], while another found that an internally oriented LOC indicated an increased cyber risk perception strongly related to reduced engagement in cyber misbehaviors [26]. Finally, another study found a positive correlation between internally oriented LOC in males and higher susceptibility to online investment scams [27]. A qualitative study using one-on-one semi-structured interviews with repeat clickers and protective stewards (users who repeatedly report potentially malicious emails without falling prey to them), indicated that repeat clickers reported a more internally oriented LOC than the protective stewards [6]. These research studies on repeat clicking suggest that individual differences are in part, if not as a whole, driving these behaviors.

Anyone is susceptible to phishing under the right conditions [1], and to be successfully “phished” once or twice may simply be bad luck. A study by Canham et al. [5] intentionally integrated immediate feedback by presenting recipients with copies of the

phishing emails that they had just fallen prey to, with highlighted cues indicating what the recipient could have used to identify that phishing email. This feedback was provided every time a recipient failed a simulation [5]. To be successfully phished eight or more times indicates that the user is not improving from training or prior experience, consistent with other research findings [4], and their actions may be driven in large part by their individual differences.

**Contextual Factors and Social Engineering.** In addition to a main effect of individual traits, the stability and persistence of these behaviors may also be driven by an interaction with contextual factors or social engineering tactics. Contextual factors refer to situational circumstances that are beyond the immediate control of an email recipient such as world events (a disaster often provides a pretext for cybercriminals to use in emails) or work role (employees who are in constant contact with the public are more exposed). Social engineering tactics refer to the methods employed by cybercriminals in deceiving email recipients into engaging with their phishing emails.

### 1.3 Research Questions

This study investigated whether the persistent aspect of repeat clicking behavior is either being driven entirely as a main effect of individual characteristics or is driven by an interaction between individual traits and social engineering tactics or contextual effects. Most phishing research treats susceptibility as a binary construct without consideration for potential degrees of susceptibility, thus relying on one-time exposures to simulated phishing emails [2]. As such, there is a gap in prior research in understanding the underlying factors for a key segment of the susceptible population, the repeat clickers. This study explored the influential factors for occasional versus repeated human susceptibility to phishing emails by focusing on the differences between repeat clickers (RC), one-time clickers (1C), and zero-clickers (ZC)<sup>1</sup>. Our study was guided by three research questions:

- RQ1: Are there detectable individual differences between the RC, 1C, and ZC groups, in ways that are discoverable through psychometric assessments?
- RQ2: Will RC exhibit similar clicking patterns as the other groups (1C, ZC) in response to different types of phishing emails? For example, will RC uniformly click every link which lands in the email inbox, or will they click in a similar pattern to the other groups?
- RQ3: How will contextual effects impact RC click-rates compare to the other groups?

In social engineering, higher sophistication attacks are less likely to be identified by individuals as potentially malicious, and therefore are more likely to succeed. A challenge posed in investigating RQ2 was how to quantify or control for the difficulty of human detection of phishing emails. The NIST Phish Scale described in the next section was determined to be the best use for this evaluation metric.

---

<sup>1</sup> 1C and ZC were chosen to reduce confusion between the number 0 and the capital O.

## 1.4 Human Phishing Detection Difficulty Using the NIST Phish Scale

The NIST Phish Scale was developed to address the challenge of developing a human phishing detection metric that accounts for user behavior. The Phish Scale integrates two dimensions of email difficulty (see Fig. 1): the number of observable cues which might alert a user that the email is a phish, and the alignment of the message premise with the recipient's user context (e.g., work role, expectations). By accounting for both a phishing email's characteristics (cues) and context (premise alignment), the Phish Scale helps to provide richer insight into human detection difficulty of phishing emails.

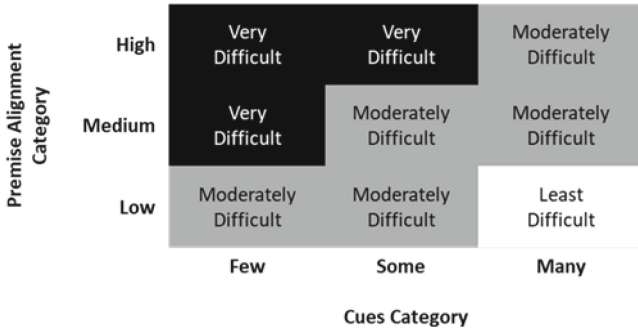
**Message Cues.** Cybercriminals who send malicious emails have different goals than those who send legitimate messages and may contain "cues" in the email that alert the receiver to its potential maliciousness. Examples of such cues include mismatched hyperlinks, oddly worded phrases, and incorrect logos [9]. When the number of cues that an email contains increases, the likelihood that a potential victim will become suspicious correspondingly increases, thus lowering the difficulty of human detection. The Phish Scale currently includes five types of message cues: Error—relating to spelling and grammar errors and inconsistencies contained in the message; Technical indicator—pertaining to email addresses, hyperlinks, and attachments; Visual Presentation indicator—relating to branding, logos, design, and formatting; Language and Content—such as a generic greeting and lack of signer details, use of time pressure, and threatening language; and Common Tactic—use of humanitarian appeals, too good to be true offers, time-limited offers, and poses as a friend, colleague, or authority figure [33].

**Premise Alignment.** Studies analyzing simulated phishing email campaigns reveal that user context contributes significantly to susceptibility to phishing emails [9]. When a phishing email is highly congruent with a user's context (e.g., work role, job tasks, personal interests), that user is more likely to ignore signs that the email is potentially malicious and instead focus on the task-relevant information that is associated with their work role [9]. While contextual knowledge endogenously drives our visual attention toward certain aspects of an email, our attention may also be exogenously "captured" by features contained within the email [34]. This suggests that when an email is less relevant to a user's context, they will be more likely to visually process the message from a data-driven perspective and therefore be more sensitive to cues that the email is malicious.

The Phish Scale method outlines four elements of a phishing email that are related to user context: the email premise 1) mimics a workplace process or practice; 2) has work-place relevance; 3) aligns with other situations or events, including external to the workplace; and 4) engenders concerns over consequences for NOT clicking. A fifth element considers whether the user has been the subject of targeted training, specific warnings, or other exposure [35]. The more the phishing email's premise aligns with the context of the user, the harder it is for the user to detect the email as a phish.

**Measuring Human Detection Difficulty Using the NIST Phish Scale.** The Phish Scale measures the human detection difficulty of a phishing email by considering both the number of cues and the message's premise alignment [36]. The greater the number of observable cues and the more an email's message is mismatched to its recipient, the less difficult that phishing email is for the recipient to detect. On the other end of the

spectrum, the fewer number of cues and the more closely the phishing email corresponds with the user’s context, the more difficult that phish is to detect. These two components are measured for an email, and categorized according to the procedure outlined in prior publications [35]. The categories for each component are combined to create an overall phishing email human detection difficulty metric (illustrated in Fig. 1).



**Fig. 1.** The NIST Phish Scale (NPS).

The NIST Phish Scale was used as an evaluation metric in this study because it provides a holistic and integrated approach for measuring human phishing detection difficulty [33].

## 2 Methods

We conducted a study which utilized multiple simulated phishing campaigns run over the course of several weeks. In this study, we examined or controlled for all three potentially contributing factors (individual differences, social engineering tactics, and contextual factors). We examined individual level traits by collecting demographic, personality, and email security knowledge prior to sending simulated phishing emails. We examine the impact of social engineering tactics by rating each email using the NIST Phish Scale. Finally, we control for contextual factors by counter-balancing each simulated phishing email with an email sent explicitly by the research team, thus controlling for events (such as the outbreak of a pandemic) which might impact overall email response rates.

### 2.1 Participants

After the study protocol was reviewed and approved by the university human subjects review board, 120 undergraduate students were recruited to participate. These students were enrolled in a university undergraduate psychology course and participating in the psychology department’s research subject pool. Three participants requested to be withdrawn from the study, leaving a total of 117 research participants whose data were analyzed. No explanation was provided for the request to have data removed. None of the participants received any formal phishing or security awareness training as part of this study.

## 2.2 Study Protocol

The study was conducted in two parts: Part I Online Questionnaires and Part II Phishing Email Campaigns. In Part I, informed consent was obtained prior to the start of any research procedures. In the consent form, subjects were informed that they were to be sent additional emails, but they were not explicitly told that these were phishing simulations. Subjects were asked to complete a series of questionnaires to collect their demographic data and their individual differences. These questionnaires queried participants about their gender, age, year in school, anxiety about internet usage, knowledge of phishing emails, confidence in detecting phishing emails, and whether they had previously been the victim of online fraud. After completing these questionnaires, participants were asked to provide their first name and university email address. Finally, they were reminded that the researchers would be contacting them via email in the upcoming weeks. The first part of the study took approximately 50 min including consent.

In Part II of the study, all participants were sent a total of eight emails, four phishing emails (*Phishing Email*) and four survey invitation emails (*Survey Invitation*). The latter was intended to act as a control to obtain click-rate baselines. The number of email campaigns was selected due to time constraints of fitting the study within the academic semester. During each email campaign, participants were randomly split into two equal groups without their knowledge. One group received a simulated phishing email, while the other group received an invitation to complete a study related survey. Group assignments rotated with each subsequent campaign. This design provided a counter-balanced control for contextual factors which might influence click-rates. The order of the *Phishing Emails* was also counterbalanced using a Latin Squares technique [37].

All emails included the participant's first name to create a feeling of personalization. If a participant clicked the embedded link in an email (*Phishing Emails* and *Survey Invitation*) the link directed them to complete a short survey asking about situational factors such as device usage (did they read the email on a mobile device, laptop, or desktop), amount of sleep obtained in the previous night, stress level, workload, consideration of consequences for clicking the link, and consequences for not clicking the link. The linked surveys were identical between the simulated *Phishing Emails* and *Survey Invitation* except for the first line which informed participants that this was a simulated phishing exercise as part of the research study (if participant had clicked the link in a *Phishing Email*), or alternatively the first line thanked them for completing the survey (if the participant had clicked the link in the *Survey Invitation*). This procedure is consistent with phishing simulation training programs commonly used by organizations around the world [11, 38–40].

Approximately two weeks after participants completed the Part I questionnaires, they were sent the first in a series of eight emails (four phishing and four study survey emails). The *Survey Invitation* emails were sent to obtain baseline click-rates and psychological states of the participants. These emails were intended to control for contextual factors and unexpected events which may have influenced click-rates for a particular campaign. All email campaigns were sent between March 16, 2020, and April 20, 2020, using the KnowBe4 phishing simulation platform. This study was conducted during the first few weeks following the university shutdown due to the COVID-19 pandemic, so having control emails (the *Survey Invitation* emails) to obtain baseline responses was critical

to evaluating RQ3 because student response rates may have been influenced due to the shift to an online course format.

Research participants were not explicitly told that the intent of this study was to learn about phishing, nor told how many phishing simulations they were to be sent, until the debriefing at the conclusion of the study. The reason that participants were not explicitly informed of the true nature of this study is that prior research has demonstrated that when subjects are informed that they are participating in phishing research they do not behave in the same way that they naturally would, and that this adversely affects the research outcomes [24]. At the end of the study, all participants were asked to complete a final survey about their experiences and provided with a debriefing statement describing the phishing research objective.

### 2.3 Questionnaires

Since so little research has been conducted on the underlying factors of repeat clicking, this study adopted an exploratory approach with regard to which individual differences might contribute to this behavior. For this reason, the following assessments and questionnaires were administered<sup>2</sup>: Basic Demographics Questionnaire (developed for this study), Internet Anxiety Questionnaire (adapted from [41]), Online Behavior and Online Fraud Questionnaire (developed for this study), Phishing Knowledge Assessment (developed for this study), Phishing Detection Confidence (developed for this study), International Personality Item Pool (Neuroticism, Extraversion, and Openness (IPIP-NEO)-60 [28, 42], Need for Cognition [43], Curiosity [44], Tolerance for Ambiguity [45], Risk Taking [46], Risk Avoidance [47], Distrust [47], Locus of Control [48, 49].

### 2.4 Email Detection Difficulty

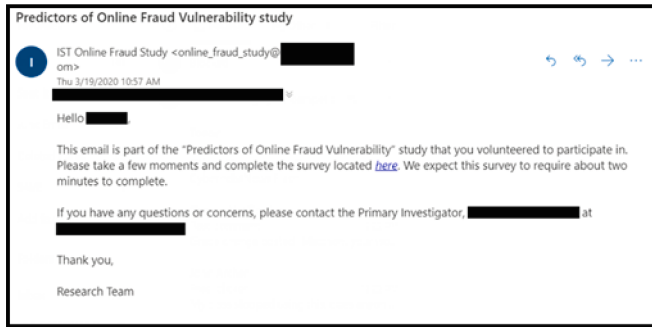
The four Phishing Emails were independently rated by two researchers familiar with the Phish Scale. During the assessment of scoring conflicts, a third researcher with knowledge of the target audience (student population) provided input towards the evaluation. A consensus was reached among all three parties, resulting in the difficulty ratings of ‘very difficult’ for three of the four Phishing Emails and ‘moderately difficult’ for one of the four Phishing Emails. The four Survey Invitation emails were identical.

The *Survey Invitation* email (presented in Fig. 2) informed participants that the email was part of the study they had volunteered for and asked for them to follow the link to complete a short survey.

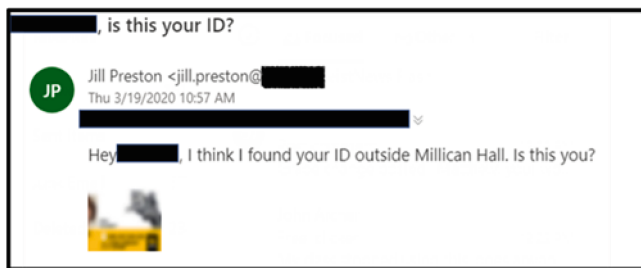
The *Lost ID Phishing Email* (see Fig. 3) claimed that the sender had found the recipient’s student ID and was attempting to confirm that the receiver was the same person as the ID. This email included an obscured image that appeared to be a thumbnail image of a student ID. To view the image, the participant needed to click on the image to view it. If the participant clicked on this image, the embedded link directed them to the study survey. This email had a very difficult human detection difficulty rating, with few cues and a medium premise alignment.

---

<sup>2</sup> A pre-print version of this manuscript is available which contains the full versions of all scales in the appendix.

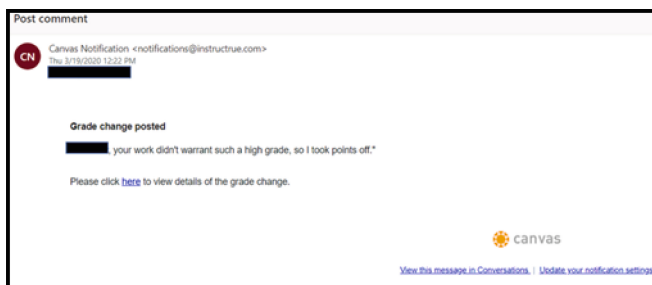


**Fig. 2.** The *Survey Invitation Email* (18.6% Response Rate across all groups)



**Fig. 3.** Lost ID *Phishing Email* (53.8% Click-Rate)

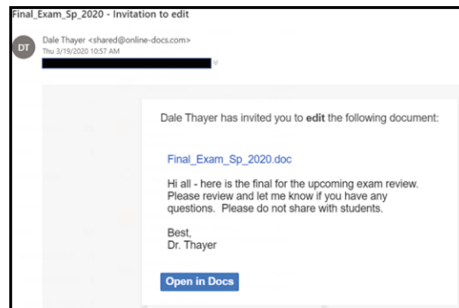
The Grade Change *Phishing Email* (see Fig. 4) claimed to notify the recipient that their grade had been changed. This email spoofed an online course management platform that was used by the university at the time of the study. A link was included in this email that claimed to take the participant to the course management site but in fact directed them to the *Phishing Email* survey landing page. The Grade Change email had a very difficult human detection difficulty rating, with some cues and a high premise alignment.



**Fig. 4.** Grade Change *Phishing Email* (54.7% Click-Rate)

The Final Exam *Phishing Email* (see Fig. 5) employed the pretext of being sent to the wrong recipient and implied that it was meant for someone with a similar name or

email address as the participant. This email claimed to have a copy of an upcoming final exam for a teaching assistant to review and advised “do not share with students.” If a participant clicked on the attached document, they were redirected to the *Phishing Email* survey landing page. This email had a very difficult human detection difficulty rating, with some cues and a high premise alignment. This *Phishing Email* was sent during the last month of the semester, which raised its premise alignment score. However, not all participants received this email at the exact same point in the semester. No order effects were observed; the participants who received this email closer to finals week did not respond at a higher rate than students who received the email earlier in the month.



**Fig. 5.** Final Exam *Phishing Email* (29.1% Click-Rate)

The final *Phishing Email* (see Fig. 6), Free iClicker, had a premise of a student no longer needing an iClicker (a wireless response device used by students in some classes). This email was created with the assistance of undergraduate research assistants who advised that this pretext would garner much interest since students are required to purchase iClickers for several classes as transient course requirements, and finding a device being offered for a low price or free (as described in this phishing email) would be perceived as highly desirable. If a participant clicked the embedded hyperlink, it directed them to the *Phishing Email* survey landing page. The iClicker email had a moderately difficult human detection difficulty rating, with some cues and a medium premise alignment.



**Fig. 6.** Free iClicker *Phishing Email* (5.1% Click-Rate)

## 2.5 Data Analysis

Findings from the initial survey were analyzed using a one-way analysis of variance (ANOVA) with the response group (ZC, 1C, or RC) as the comparative factor. Differences in the click-rates of the phishing email templates, and any potential order effects, were analyzed using a chi-square test with the control emails as the baseline response comparison.

## 3 Results

Due to the limited number of phishing emails which were sent, for the purposes of this analysis, repeat clickers are defined as someone who clicked on two or more hyperlinks within the simulated phishing emails. In this study, 27 (23.1%) subjects did not click on any of the phishing hyperlinks (ZC), 32 (27.4%) clicked on exactly one phishing hyperlink (1C), and 58 (49.6%) clicked on two or more phishing hyperlinks and thus met the definition of being a repeat clicker (RC) in this study.

### 3.1 Individual Differences, RQ1

**Demographics.** None of these assessments differed significantly between the groups. Surprisingly, there were no significant differences between the 1C and RC groups on any of the Knowledge Assessment questions, which included the self-assessed level of confidence to detect phishing emails.

**Internet Usage.** As measured by the online behavior questionnaire, the RC group spent less time (on a weekly basis) working online ( $M = 1.64$ ), than the 1C ( $M = 7.06$ ), or the ZC ( $M = 7.12$ ) groups did;  $F(2, 113) = 4.972$ ,  $p < .01$ ,  $\eta^2 = 0.81$ . A Tukey post-hoc test confirmed this difference was between RC and the other groups ( $p < .05$ ), rather than between the 1C and ZC. The questionnaire specified doing *paid* work as a different question from doing *schoolwork* online, meaning that the only significant difference was observed for paid time online. There were no differences between the groups in the total amount of time spent online, weekly time spent on any of the other online activities, nor internet anxiety. There was a difference in the number of times each group checked email throughout the day  $F(2, 95) = 4.965$ ,  $p < .01$ ,  $\eta^2 = 0.095$ , with the RC group checking more often ( $M = 4.04$ ), than the 1C ( $M = 3.59$ ), or ZC (3.17), with this separation being accounted for by the difference between the ZC and RC groups ( $p < .01$ ).

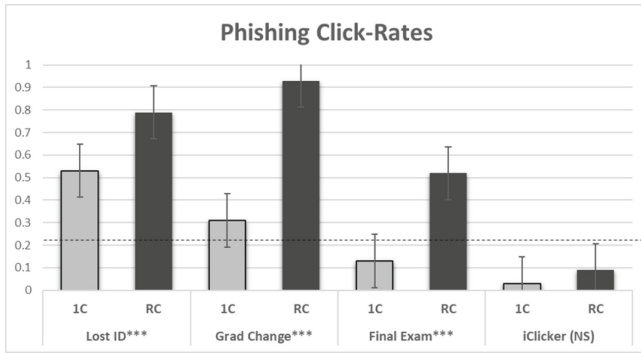
**Big 5 Personality.** There were no differences found between the groups on any of the Big 5 Personality dimensions.

**Locus of Control.** There was a significant difference,  $F(2, 114) = 2.536$ ,  $p < .05$ ,  $\eta^2 = 0.71$ , between the groups on Locus of Control, with the RC group mean ( $M = 2.93$ ), the 1C group ( $M = 3.325$ ), and the ZC group ( $M = 3.379$ ). This indicates that RC group was more internally oriented in their Locus of Control than were the other two groups. A Tukey post-hoc test did not indicate a significant difference at the 0.05 level between the RC and 1C groups ( $p = .058$ ), but significance was observed between the RC and ZC groups ( $p = .037$ ). No significant difference was observed between the between the 1C and ZC groups ( $p = .960$ ).

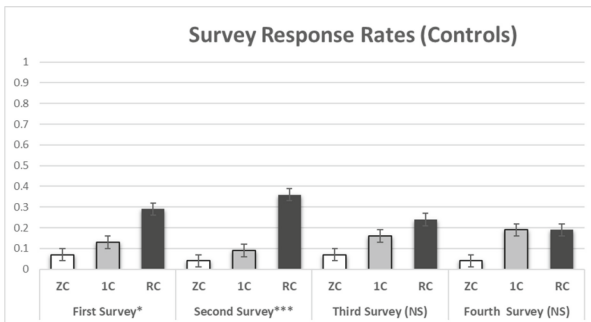
**Other Individual Differences.** There was a difference in the Need for Cognition between the groups,  $F(2, 114) = 3.382, p < .05, \eta^2 = 0.056$ , with the ZC and 1C groups reporting a higher need than the RC group. There were no differences between the groups on the other questionnaires (Curiosity, Tolerance for Ambiguity, Risk Taking, Risk Avoidance, Distrust).

### 3.2 Social Engineering Tactics, RQ2

As depicted in Figs. 7 and 8, three of the *Phishing Emails* received higher click-rates than the mean response rate to the *Survey Invitation* emails (for the 1C and RC groups). The mean click-rate of 20.6% across these two groups is depicted as the dashed line in Fig. 7); the four simulated phishing emails received click-rates of 53.8% (Lost ID), 54.7% (Grade Change), 29.1% (Final Exam), and 5.1% (Free iClicker). It should be noted that while the order that the phishing emails were sent was counter balanced, Fig. 8 depicting Survey Invitation emails follows the order of receipt and response.



**Fig. 7.** Click-Rates for Each Phishing Email for Each Clicker Group (1C and RC)<sup>3</sup>



**Fig. 8.** Response Rates for Each Survey Invitation Email for Each Group (ZC, 1C, and RC)

<sup>3</sup> \* Indicates  $p < .05$ , \*\*  $p < .01$ , and \*\*\*  $p < .001$ .

A matched comparison of the click-rates for each *Phishing Email* against the click-rate for each respective corresponding *Survey Invitation* email is listed in Table 1. The click-rates for the Lost ID and Grade Change *Phishing Emails* were significantly greater than the controls; the Final Exam *Phishing Email* was not statistically significant, and there was no difference between the Free iClicker *Phishing Emails* and the control.

**Table 1.** Chi-squared comparison of phishing emails to control emails.

Phishing Email	$\chi^2$	p-value
Lost ID	9.42	.002
Grade Change	5.82	.016
Final Exam	2.88	.089
Free iClicker	0.001	.977

### 3.3 Contextual Factors, RQ3

As described previously, the *Survey Invitation* emails were sent to obtain baseline click-rates as a means of comparison with the phishing emails, controlling for unexpected events that may have influenced click-rates for a particular email campaign. Although the response rates for the *Survey Invitation* emails decreased over time (as depicted in Fig. 8), there were no differences between any of the email campaigns (the batch of emails including both phishing and control) over time, indicating that there were no influential circumstances impacting the overall likelihood of engaging with emails.

## 4 Discussion

### 4.1 Individual Differences, RQ1

The first research question in this study asked whether detectable differences exist between the RC group compared to the 1C and ZC groups. Differences were found in both the orientation of LOC and in the Need for Cognition.

Previous work suggested that an internally oriented LOC might be associated with repeat clicking behavior [6]. The difference in LOC between the 1C and RC groups, but not the ZC and 1C groups suggests that this factor is differentially associated with repeat clicking behavior, rather than general phishing susceptibility (otherwise a difference would have been observed between ZC and 1C). This factor should be explored through more extensive research; if it replicates, a potential explanation may be that repeat clickers are more responsive to security awareness training than those with an externally oriented LOC, but that they need to be convinced of the efficacy of security practices. This finding is also consistent with other research findings [26, 27]. A potential direction for future research might be exploring whether certain LOC scale items are more predictive of repeat clicking than others.

An intriguing finding was that repeat clickers were lower in their Need for Cognition scores than the other groups. The need for cognition refers to the intrinsic desire for a person to comprehend and structure environmental information [43, 50]. Thus, this appears to be a reasonable factor influencing repeat clicking behavior. Little research has been conducted on the relationship between need for cognition and phishing susceptibility [50], with no studies looking at the relationship with repeat clicking that the authors are aware of.

## 4.2 Social Engineering Tactics, RQ2

The second research question asked whether the RC group would exhibit similar clicking patterns as the other groups in response to different social engineering tactics (as measure by the Phish Scale). The rank ordering of Phishing Email click-rates for the RC group did not follow the same pattern as the 1C group. The rank order for the 1C click-rate was Lost ID, Grade Change, Final Exam, and iClicker. The RC group had a similar order, with Grade Change having a higher click-rate than Lost ID. The response rates for the surveys declined over time for the RC group, as they did for the other groups; however, these were significantly higher for the first two survey emails sent, and then without significant difference for the last two. Interestingly, the ZC group had a consistently lower survey response rate than the other groups. This suggests that perhaps they (like the RC group) are engaging with email habitually, which is consistent with other research [4]. These findings suggest that an interaction between social engineering tactics and repeat clicking may be occurring.

## 4.3 Contextual Factors, RQ3

Our final research question focused on the contextual effects of external circumstances that influence repeat clicking behavior. It is likely that the low click-rates for the Free iClicker email were driven by current events, since all email campaigns were sent between March 16, 2020, and April 20, 2020, during the first few weeks following the university shutdown due to the COVID-19 pandemic. When this *Phishing Email* was developed (pre-pandemic) the student research assistants advised that iClicker devices were highly sought after by students. When the university shutdown due to the COVID-19 pandemic, all classes were taught exclusively online, thus greatly reducing the value of iClickers, which were not used in online courses. This drastically reduced the appeal of these phishing emails for the target population. While this did not directly address RQ3, the severely low click-rates for the iClicker email does demonstrate the effect for susceptibility to specific messages sent within matched contexts, further warranting the relevance of the Phish Scale as an effective evaluation metric.

## 4.4 The NIST Phish Scale and Training Implications

At the time of this study, the Phish Scale was a novel method which had not yet been tested on a population outside of the U.S. Government; it was developed and tested based on data from a single U.S. Federal Government agency [35]. Although not the larger

purpose of this study, tangentially, we sought to contribute to the ongoing testing of the Phish Scale by applying the Phish Scale in a university environment. The overall click-rates corresponded with the Phish Scale ratings, suggesting that it effectively predicted human phishing email detection difficulty outside of the development context.

While the study presented in this paper represents a smaller sample of emails than the original study [33], its results set the stage for further exploration into the differences between human phishing detection in professional and non-professional target audiences. It may be that since students and professional employees approach email from different contexts, the differences in their environment and tasks lead to diverging email behaviors. These populations may also be driven by diverse psychological motives. Phishing susceptibility factors could have important training implications as they demonstrate that organizations may need their security awareness training to account for a user's role and how it influences context.

**Premise Alignment Impact on Overall Click-Rates.** Robert Cialdini discusses 'magnetizers' of attention, the self-relevant, the unfinished, and the mysterious [51] which may have played a role in the click-rates for the various *Phishing Emails* in this study. The Lost ID email utilized all three magnetizers. Self-relevance was established by the email's claim of having the recipient's identification card and the message was addressed to participants personally. The email was unfinished because it included the embedded image of an identification card which was obscured to the degree that the receiver could identify it as a student ID card, but not to the degree that they could resolve whether it was their ID card. Finally, it created a mystery. In fact, several participants reported that they had they currently possessed their student ID but wanted to know whose ID card had been found.

Applying these same magnetizers to the Grade Change *Phishing Email*, we observe a similar pattern in that it was self-relevant (it was a participant's grade that was claimed to have been changed), unfinished (the participant needed to click the link to read the complete message), and mysterious (the participant was not informed why the grade was being changed). Consistent with the Phish Scale difficulty ratings, these two emails received substantially higher click-rates than the other phishing emails (both over 50%), suggesting that user context, or an email's premise alignment, does play a significant role in likelihood of clicking. This is further supported when we consider that the Final Exam *Phishing Email* was not self-relevant unless the student happened to be in the class it purported to be from, not mysterious because it told the receiver exactly what the attachment contained, but it was unfinished since the receiver simply need to open the attachment to obtain the contents. The iClicker was neither mysterious, unfinished, nor self-relevant (due to the COVID-19 shutdown of in-person classes). Both the Final Exam and iClicker emails garnered substantially fewer clicks, further supporting the potential influence of attentional magnetizers and the influence of context on phishing email engagement.

**High Click-Rates Might Be Helpful? Security departments often rely on phishing simulation click-rates (failures) as a singular metric of human security performance.** Click-rates alone do not factor in user context and are absent of the detection difficulty of the emails being sent. Rather than focusing solely on click-rates and reducing them to zero, considering the users in an organization together with the

difficulty they experience in identifying phishing emails may reduce vulnerabilities to highly sophisticated phishing emails in the wild. Additionally, considering user context in phishing awareness training programs can provide qualitative information on the topics or pretexts users might be more susceptible to (premise alignment). This information is key to understanding user behaviors in detecting phishing emails, potentially enhancing the training return on investment of phishing simulations.

Not widely discussed are the potential benefits of failing simulated phishing campaigns. Over 100 years of learning research demonstrates that receiving feedback on actions taken (both succeeding and failing) is mandatory for effective learning to occur [52]. The key for individuals to improve is to not continuously fail for the same reasons or at the same level of difficulty. Vygotsky discussed a zone of proximal development which represents the gap between what an individual is currently capable of, and what an individual is (currently) incapable of achieving [53]. It is critical to guide performers through this zone of proximal development to ensure that they do not become overwhelmed in the learning process. It is this process of guiding through the zone where a tailored approach to helping educate repeat clickers might be most effective.

## 5 Conclusion

This study represents a potentially paradigm shifting perspective on phishing susceptibility in that our findings suggest that differences exist between occasional and repeated victims of phishing email scams. This has significant implications for protecting users online and defending the human attack surface within organizations. Additional work should also seek to understand the underlying and contributing factors for these differences. While this work focused on the impact of individual differences on repeat clicking, future studies should also explore role-based influences on this behavior. One challenge here is that to study this phenomenon requires a long-term engagement with the study population. This challenge withstanding, significant benefits potentially exist for those companies or agencies which uncover the mystery behind repeat clickers.

## 6 Disclaimer

Any mention of commercial products or companies is for information only and does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products are necessarily the best available for the purpose.

**Acknowledgements.** The authors wish to thank Dr. Ben D. Sawyer, Dr. Erica Castilho Grao, Dr. Clay Posey, Michael Constantino, and Delainey Strickland for their assistance in collecting and analyzing the data for this study.

This research was conducted with the support of the National Institute of Standards and Technology (NIST) under Financial Assistance Award Number: 60NANB19D123. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of NIST or the U.S. Government.

## References

1. Hadnagy, C.: *Social Engineering: The Science of Human Hacking*, 1st ed. Wiley (2018). <https://doi.org/10.1002/9781119433729>
2. Canham, M., Fiore, S.M., Constantino, M., Caulkins, B., Reinerman-Jones, L.: *The Enduring Mystery of the Repeat Clickers* (2019)
3. Verizon. 2023 Data Breach Investigations Report (DBIR). Verizon Enterprise Solutions. <https://www.verizon.com/business/resources/reports/2023-data-breach-investigations-report-dbir.pdf>. Accessed 19 Jun 2023
4. Caputo, D.D., Pflieger, S.L., Freeman, J.D., Johnson, M.E.: Going spear phishing: exploring embedded training and awareness. *IEEE Secur. Priv.* **12**(1), 28–38 (2013). <https://doi.org/10.1109/MSP.2013.106>
5. Canham, M., Posey, C., Strickland, D., Constantino, M.: Phishing for long tails: examining organizational repeat clickers and protective stewards. *SAGE Open* **11**(1), 215824402199065 (2021). <https://doi.org/10.1177/2158244021990656>
6. Canham, M.: Repeat Clicking: A Lack of Awareness Is Not the Problem. PsyArXiv, preprint (2023). <https://doi.org/10.31234/osf.io/36eqn>
7. Li, W., Lee, J., Purl, J., Greitzer, F., Yousefi, B., Laskey, K.: Experimental Investigation of Demographic Factors Related to Phishing Susceptibility (2020). <https://doi.org/10.24251/HICSS.2020.274>
8. Lain, D., Kostiaainen, K., Čapkun, S.: Phishing in Organizations: Findings from a Large-Scale and Long-Term Study. In: *2022 IEEE Symposium on Security and Privacy (SP)*, pp. 842–859 (2022). <https://doi.org/10.1109/SP46214.2022.9833766>
9. Greene, K., Steves, M., Theofanos, M., Kostick, J.: User context: an explanatory variable in phishing susceptibility. In: *Proceedings 2018 Workshop on Usable Security*, San Diego, CA: Internet Society (2018). <https://doi.org/10.14722/usec.2018.23016>
10. Elevate Security. High Risk Users and Where to Find Them (2023)
11. PhishMe. Enterprise Phishing Susceptibility Report (2015). [https://cofense.com/wp-content/uploads/2017/10/PhishMe\\_EnterprisePhishingSusceptibilityReport\\_2015\\_Final.pdf](https://cofense.com/wp-content/uploads/2017/10/PhishMe_EnterprisePhishingSusceptibilityReport_2015_Final.pdf)
12. Vernon, H.M.: An investigation of the factors concerned in the causation of industrial accidents. *J. Manag. Hist.* **1**(2), 65–78 (1918)
13. Hogan, R.: The accident-prone personality. *People Strategy* **39**(1), 20–24 (2016)
14. Hansen, C.P.: Personality characteristics of the accident involved employee. *J. Bus. Psychol.* **2**(4), 346–365 (1988)
15. Rotter, J.B.: Rotter’s Internal-External Control Scale. *Psychological Monographs: General and Applied* (1966)
16. Bridge, R.G.: “Internal-external control and seat-belt use”, presented at the Western Psychological Association. American Psychological Association, San Francisco (1971)
17. Hoyt, M.F.: Internal-external control and beliefs about automobile travel. *J. Res. Pers.* **7**, 288–293 (1973)
18. Denning, D.L.: Correlates of employee safety performance. In: Presented at the Southeastern I/O Psychology Association Meeting, Atlanta, Georgia (1983)
19. Wichman, H., Ball, J.: Locus of control, self-serving biases, and attitudes towards safety in general aviation pilots. *Aviat. Space Environ. Med.* **54**(6), 507–510 (1983)
20. Jones, J.W.: *The Safety Locus of Control Scale*. St. Paul, MN: The St. Paul Companies (1984)
21. Jones, J.W., Wuebker, L.: Development and validation of the Safety Locus of Control (SLC) scale. *Percept. Mot. Skills* **61**, 151–161 (1985)
22. Mayer, R.E., Treat, J.R.: Psychological, social, and cognitive characteristics of high-risk drivers: a pilot study. *Accid. Anal. Prev.* **9**, 1–8 (1977)

23. Ayaburi, E., Andoh-Baidoo, F.K.: Understanding phishing susceptibility: an integrated model of cue-utilization and habits. In: ICIS 2019 Proceedings (2019). [https://aisel.aisnet.org/ici2019/cyber\\_security\\_privacy\\_ethics\\_IS/cyber\\_security\\_privacy/43](https://aisel.aisnet.org/ici2019/cyber_security_privacy_ethics_IS/cyber_security_privacy/43)
24. Parsons, K., McCormac, A., Pattinson, M., Butavicius, M., Jerram, C.: The design of phishing studies: Challenges for researchers. *Comput. Secur.* **52**, 194–206 (2015). <https://doi.org/10.1016/j.cose.2015.02.008>
25. Hadlington, L., Popovac, M., Janicke, H., Yevseyeva, I., Jones, K.: Exploring the role of work identity and work locus of control in information security awareness. *Comput. Secur.* **81**, 41–48 (2018)
26. Johnson, K.: Better Safe than Sorry: The Relationship between Locus of Control, Perception of Risk, and Cyber Misbehaviors – ProQuest. In: Doctoral dissertation, University of South Florida (2018). <https://www.proquest.com/openview/42ccd20fc5e2b6403ecce12dff9686055/1?pq-origsite=gscholar&cbl=18750>. Accessed 30 Dec 2023
27. Whitty, M.T.: Is there a scam for everyone? Psychologically profiling cyberscam victims. *Eur. J. Crim. Policy Res.* **26**(3), 399–409 (2020). <https://doi.org/10.1007/s10610-020-09458-z>
28. McCrae, R.R., Costa, P.T.: Validation of the five-factor model of personality across instruments and observers. *J. Pers. Soc. Psychol.* **52**(1), 81–90 (1987). <https://doi.org/10.1037/0022-3514.52.1.81>
29. Lawson, P., Zielinska, O., Pearson, C., Mayhorn, C.B.: Interaction of personality and persuasion tactics in email phishing attacks. *Proc. Hum. Factors Ergon. Soc. Ann. Meet.* **61**(1), 1331–1333 (2017). <https://doi.org/10.1177/1541931213601815>
30. Pattinson, M., Jerram, C., Parsons, K., McCormac, A., Butavicius, M.: Why do some people manage phishing e-mails better than others? *Inf. Manag. Comput. Secur.* **20**(1), 18–28 (2012). <https://doi.org/10.1108/09685221211219173>
31. Sudzina, F., Pavlicek, A.: Propensity to click on suspicious links: impact of gender, of age, and of personality traits. In: Digital Transformation – From Connecting Things to Transforming Our Lives, University of Maribor Press, pp. 593–601 (2017). <https://doi.org/10.18690/978-961-286-043-1.41>
32. Workman, M.: Wisecrackers: A theory-grounded investigation of phishing and pretext social engineering threats to information security. *J. Am. Soc. Inform. Sci. Technol.* **59**(4), 662–674 (2008). <https://doi.org/10.1002/asi.20779>
33. Steves, M.P., Greene, K.K., Theofanos, M.F.: A phish scale: rating human phishing message detection difficulty. In: Proceedings 2019 Workshop on Usable Security, San Diego, CA: Internet Society (2019). <https://doi.org/10.14722/usec.2019.23028>
34. Canham, M., Hegarty, M.: Effects of knowledge and display design on comprehension of complex graphics. *Learn. Instr.* **20**(2), 155–166 (2010). <https://doi.org/10.1016/j.learninstruc.2009.02.014>
35. Steves, M., Greene, K., Theofanos, M.: Categorizing human phishing difficulty: a Phish Scale. *J. Cybersecurity* **6**(1), 1–16 (2020). <https://doi.org/10.1093/cybsec/tyaa009>
36. Dawkins, S., Jacobs, J.: NIST Phish Scale User Guide. National Institute of Standards and Technology, Gaithersburg, MD, NIST TN 2276 (2023). <https://doi.org/10.6028/NIST.TN.2276>
37. Shah, K.R., Sinha, B.K.: 4 Row-Column Designs. *Theory of Optimal Designs*. In: Lecture Notes in Statistics, no. 54. Springer-Verlag (1989). <https://doi.org/10.1007/978-1-4612-3662-7>
38. Carella, A., Kotsoev, M., Truta, T.M.: Impact of security awareness training on phishing click-through rates. In: 2017 IEEE International Conference on Big Data (Big Data), pp. 4458–4466 (2017). <https://doi.org/10.1109/BigData.2017.8258485>

39. Halevi, T., Memon, N., Nov, O.: Spear-phishing in the wild: a real-world study of personality, phishing self-efficacy and vulnerability to spear-phishing attacks. In: Social Science Research Network, Rochester, NY, SSRN Scholarly Paper ID 2544742 (2015). <https://doi.org/10.2139/ssrn.2544742>
40. Moody, G.D., Galletta, D.F., Dunn, B.K.: Which phish get caught? An exploratory study of individuals' susceptibility to phishing. *Eur. J. Inf. Syst.* **26**(6), 564–584 (2017). <https://doi.org/10.1057/s41303-017-0058-x>
41. Joiner, R., Brosnan, M., Duffield, J., Gavin, J., Maras, P.: The relationship between Internet identification, Internet anxiety and Internet use. *Comput. Hum. Behav.* **23**(3), 1408–1420 (2007). <https://doi.org/10.1016/j.chb.2005.03.002>
42. Maples-Keller, J.L., Williamson, R.L., Sleep, C.E., Carter, N.T., Campbell, W.K., Miller, J.D.: Using item response theory to develop a 60-item representation of the NEO PI-R using the international personality item pool: development of the IPIP-NEO-60. *J. Pers. Assess.* **101**(1), 4–15 (2019). <https://doi.org/10.1080/00223891.2017.1381968>
43. Cacioppo, J.T., Petty, R.E.: The need for cognition. *J. Pers. Soc. Psychol.* **42**(1), 116–131 (1982). <https://doi.org/10.1037/0022-3514.42.1.116>
44. Collins, R.P., Litman, J.A., Spielberger, C.D.: The measurement of perceptual curiosity. *Personality Individ. Differ.* **36**(5), 1127–1141 (2004). [https://doi.org/10.1016/S0191-8869\(03\)00205-8](https://doi.org/10.1016/S0191-8869(03)00205-8)
45. Herman, J.L., Stevens, M.J., Bird, A., Mendenhall, M., Oddou, G.: The tolerance for ambiguity scale: towards a more refined measure for international management research. *Int. J. Intercult. Relat.* **34**(1), 58–65 (2010). <https://doi.org/10.1016/j.ijintrel.2009.09.004>
46. Nicholson, N., Soane, E., Fenton-O'Creevy, M., Willman, P.: Personality and domain-specific risk taking. *J. Risk Res.* **8**(2), 157–176 (2005). <https://doi.org/10.1080/1366987032000123856>
47. Tellegen, A.: *Multidimensional Personality Questionnaire-276 (MPQ-276) Test Booklet*, 1st ed., vol. 1. University of Minnesota Press, Minneapolis (1995)
48. Levenson, H.: Differentiating among internality, powerful others, and chance. In: *Research with the Locus of Control Construct*, Lefcourt, H.M., Ed., Academic Press, pp. 1–15 (1981)
49. Oregon Research Institute. *Locus of Control, Single Construct Scoring Keys, International Personality Item Pool* (2022). <https://ipip.ori.org/newSingleConstructsKey.htm>
50. (Robert) Luo, X., Zhang, W., Burd, S., Seazzu, A.: Investigating phishing victimization with the Heuristic–Systematic Model: A theoretical framework and an exploration. *Comput. Secur.* **38**, 28–38 (2013). <https://doi.org/10.1016/j.cose.2012.12.003>
51. Cialdini, R.B.: *Pre-Suasion: A Revolutionary Way to Influence and Persuade*, Reprint edition. Simon & Schuster, New York (2016)
52. Mayer, R.E., Alexander, P.A.: *Handbook of Research on Learning and Instruction*. Taylor & Francis, Florence (2016)
53. Vygotsky, L.S., Cole, M.: *Mind in Society: Development of Higher Psychological Processes*. Harvard University Press, Cambridge (1978)