THE GERONTOLOGICAL SOCIETY OF AMERICA®

Journals of Gerontology: Psychological Sciences
cite as: *J Gerontol B Psychol Sci Soc Sci*, 2021, Vol. 76, No. 9, 1711–1715
doi:10.1093/geronb/gbaa228
Advance Access publication December 30, 2020

OXFORD

Research Report

Is This Phishing? Older Age Is Associated With Greater Difficulty Discriminating Between Safe and Malicious Emails

Matthew D. Grilli, PhD,[1,2,*] Katelyn S. McVeigh, BA,[1] Ziad M. Hakim,[1,3] Aubrey A. Wank, MA,[1] Sarah J. Getz, PhD,[4,5] Bonnie E. Levin, PhD,[4,5] Natalie C. Ebner, PhD,[4,6] and Robert C. Wilson, PhD[1,7]

[1]Department of Psychology, University of Arizona, Tucson, USA. [2]Department of Neurology, University of Arizona, Tucson, USA. [3]Department of Psychology, University of Florida, Gainesville, USA. [4]Evelyn F. McKnight Brain Institute, University of Miami, Miami, Florida, USA. [5]Department of Neurology, Miller School of Medicine, University of Miami, Florida, USA. [6]Department of Aging and Geriatric Research, Institute on Aging, University of Florida, Gainesville, USA. [7]Cognitive Science Program, University of Arizona, Tucson, USA.

*Address correspondence to: Matthew D. Grilli, PhD, Department of Psychology, University of Arizona, 1503 E. University Blvd., Tucson, AZ 85721-0068, USA. E-mail: mdgrilli@arizona.edu

Received: July 10, 2020; Editorial Decision Date: December 21, 2020

Decision Editor: Derek Isaacowitz, PhD, FGSA

Abstract

Objectives: As our social worlds become increasingly digitally connected, so too has concern about older adults falling victim to "phishing" emails, which attempt to deceive a person into identity theft and fraud. In the present study, we investigated whether older age is associated with differences in perceived suspiciousness of phishing emails.
Methods: Sixty-five cognitively normal middle-aged to older adults rated a series of genuine and phishing emails on a scale from definitely safe to definitely suspicious.
Results: Although older age was not related to a shift in overall perception of email safety, older age was related to worse discrimination between genuine and phishing emails, according to perceived suspiciousness.
Discussion: These findings suggest that cognitively normal older adults may be at particular risk for online fraud because of an age-associated reduction in their sensitivity to the credibility of emails.

Keywords: Aging, Cybersecurity, Decision making, Online scams

Email is part of a digitally connected lifestyle, and for many is essential to daily personal and work-related interactions. However, email comes with a dark side, as online fraudsters have learned to use "phishing" emails to deceive people into providing personal information or downloading malicious software (Carr, 2011). The consequences of clicking on just one phishing email can be severe, including identity theft, fraud, and extortion via malicious encryption of electronic information. Phishing emails may be especially devastating for older adults, for whom there is higher risk of lost independent living and limited earning power for recovery (Templeton & Kirkman, 2007). However, whether older adults are more vulnerable to phishing emails is unclear and may depend on demographic and situational factors, and how well phishing messages mask their true intent (Baillon et al., 2019; Ebner et al., 2020; Gavett et al., 2017; Lin et al., 2019; Sheng et al., 2010).

The present study investigated the effect of older age on the ability to detect the *suspiciousness* of emails. Not surprisingly, whether a person is suspicious of an email is a key factor in determining if they engage with it (Vishwanath et al., 2018). However, it is not clear whether older age is associated with differences in perceived suspiciousness of phishing emails. One possibility is that older age is related to a maladaptive shift in the general perception of email safety. For instance, if the link between greater trustworthiness and older age (Bailey & Leon, 2019) applies to the perception of emails, there could be a tendency to view emails as generally safe, which may lead to more phishing emails being engaged with. Another, nonmutually exclusive possibility is that older age is related to greater difficulty perceiving differences between genuine and phishing emails. For example, episodic memory and executive functioning, which tend to decline with older age (Glisky, 2007), may be needed to decide if an email is malicious or safe, as might prior experience with email. From this view, older adults may judge an email with incomplete information, leading to less precise decisions about suspiciousness. In other words, with older age, not only might many phishing emails be viewed as overly safe but also many genuine emails might be viewed too suspiciously, compromising the value of suspiciousness as a psychological signal of which emails are safe.

To evaluate the possibilities of age effects in response bias (e.g., the tendency to perceive emails as generally safe) or discrimination toward phishing and genuine emails, we turned to a new laboratory task that we developed (Hakim et al., 2020), known as the Phishing Email Suspicion Test or "PEST." In PEST, participants make judgments about their suspiciousness toward a series of emails, which includes real phishing emails that have been deployed (by others) to defraud people. In a group of middle-aged and older adults, we hypothesized that if older age is associated with worse detection of phishing emails because of a general shift in email perception, there would be an age-related stronger response bias toward judging emails as safe. We also hypothesized that if older age is related to worse detection of phishing emails because of difficulty perceiving the differences between safe and malicious emails, there would be an age-related reduction in discrimination regarding email suspiciousness ratings.

Method

Participants

The present study, approved by the Institutional Review Board at University of Arizona, was largely run over the course of the 2019–2020 academic year. Participants, who were recruited from the Tucson, Arizona community via an email and website advertisement, had to be middle-aged or older (50+ years) and cognitively unimpaired according to a neuropsychological profile approach (Grilli et al., 2018, see Supplementary Materials for details and cognitive scores). Participants were screened for depression, were living independently, and reported no relevant neurologic history (see Supplementary Materials).

We recruited approximately 80 individuals, based on a power analysis for a separate intervention study. Fifteen individuals were excluded because their cognition was not unimpaired ($n = 7$), they scored high on depression ($n = 6$), or there were technical issues with PEST ($n = 2$). Therefore, 65 individuals were included. A power analysis, which was conducted with G*Power 3.1 (Faul et al., 2007) before any of the analyses, indicated that with a sample size of 65, $\alpha = 0.05$, and power $(1 - \beta) = 0.80$, we could detect medium relationships ($r = 0.33$) between age and email performance (two-tailed significance tests).

The sample was older ($M = 69$, $SD = 6.9$, range = 54–83), highly educated ($M = 17.2$, $SD = 2.1$, range = 12–22), and the majority were female (47 females, 18 males) and non-Hispanic White (59 non-Hispanic White, six Hispanic).

Procedures

Phishing Email Suspicion Test

PEST presents participants with emails in a realistic inbox environment with the instructions to review emails and judge their credibility until the inbox that they are viewing is empty. The participant judges the credibility of each email on a 4-point scale: 1 = *definitely safe* to 4 = *definitely suspicious*. In total, participants are presented with 40 real phishing emails, 40 real genuine emails, 40 simulated phishing emails, and 40 simulated genuine emails. PEST includes simulated emails to increase the diversity of emails that are judged. PEST selects all emails from a library of 348 emails made up of 140 real phishing emails, 40 real genuine emails, 84 simulated phishing emails, and 84 simulated genuine emails. Therefore, each participant is not presented with the exact same emails, because as part of a larger project, we are interested in evaluating why some emails may be more effectively deceptive than others. Also, the order of emails in the task is randomized for each participant. To minimize learning, and for ecological validity, participants do not receive feedback about their performance until the end of the experiment. PEST takes approximately 1 hr to complete. Please see Supplementary Materials for a visualization of the PEST email environment and details about PEST.

Analyses

We analyzed the effect of age on response bias and discrimination in two ways. In one approach, we used standard signal detection theory/analyses to calculate bias and d' (Green & Swets, 1966). For these calculations, phishing emails were considered "signal present." A "definitely" or

"possibly" rating in the correct direction for phishing and genuine emails were hits and correct rejections, respectively, whereas a "definitely" or "possibly" rating in the incorrect direction for genuine and phishing emails were false alarms and misses, respectively. We calculated bias and d' separately for real and simulated emails. With these metrics, we used Pearson correlations to examine the relationship between age and bias toward judging emails, and between age and discrimination ability (d').

In another approach, we used a computational model of decision making for email suspiciousness that we recently developed (Hakim et al., 2020; Wilson 2018). This computational model shows that a variety of factors contribute to judging the suspiciousness of an email, including response bias, a person's recent decisions, and the content of the email currently viewed. One advantage of this new computational modeling approach is it can better account for the full range of judgments on a 4-point scale, as opposed to dichotomizing decisions as correct versus incorrect. Another advantage is that this model accounts for the effect of recent decisions on current decisions, allowing us to examine the impact of the random presentation order of emails on overall performance. See Supplementary Materials for a full description of the computational modeling approach and evidence that it taps into the same underlying construct as the standard signal detection theory analyses while providing nonredundant information. For the present study, two parameters from our computational model were of particular interest: (a) a person's response bias, which similar to signal detection theory, captures their overall propensity to say that an email is safe or not; and (b) their discrimination, which captures their overall ability to distinguish between phishing and genuine emails. We examined the relationship between age and these parameters using Pearson correlations.

Although age was not significantly related to education, nor did it differ by gender (*p*s ≥ .08), we repeated all analyses with education and gender as nuisance covariates, given that they can influence age effects and/or phishing detection (Gavett et al., 2017; Lin et al., 2019). All analyses and graphs were conducted in or created with RStudio (R Core Team, 2019) and JASP (JASP Team, 2020).

Results

Ability to Detect Suspiciousness of Emails

The ability to discriminate phishing from genuine emails, as measured by d', was higher for real emails (M = 1.21, SD = 0.50) relative to simulated emails (M = 0.56, SD = 0.29), $t(64)$ = 11.43, p < .001, d = 1.42, 95% CI [1.07, 1.76]. In regard to bias, there was a stronger tendency to judge simulated emails as suspicious (M = −0.22, SD = 0.49) relative to real emails (M = −0.11, SD = 0.43), $t(64)$ = 2.2, p = .04, d = 0.27, 95% CI [0.02, 0.51].

Relationship Between Older Age and Suspiciousness Detection

As shown in Figure 1, for our standard signal detection metrics, older age was not significantly related to response bias toward simulated emails, $r(63)$ = 0.16, p = .20, 95% CI [−0.09, 0.39], or real emails, $r(63)$ = −0.06, p = .63, 95% CI [−0.31, 0.19]. In contrast, older age was significantly related to lower d' for simulated phishing and genuine emails, $r(63)$ = −0.34, p = .005, 95% CI [−0.54, −0.11], and real phishing and genuine emails, $r(63)$ = −0.48 p < .001, 95% CI [−0.65, −0.27].

As shown in Figure 2, the outcomes of our computational modeling approach tell a similar story. Specifically, older age was not significantly related to our computational modeling metric of response bias, $r(63)$ = −0.08, p = .51, 95% CI [−0.32, 0.16]. However, older age was significantly related to worse discrimination between phishing and genuine emails in this model, $r(63)$ = −0.38, p = .002, 95% CI [−0.57, −0.15].

Repeating all of the age-related analyses controlling for gender and education did not alter the significance of the outcomes. Also, repeating the analyses as Spearman correlations did not change the outcomes, indicating that the few participants with very low discrimination did not disproportionately drive significant relationships.

Given that emails were presented in random order, we examined whether age was related to sequential effects in which current ratings are biased by previous stimuli and ratings, using our computational modeling approach (see Supplementary Materials). Neither decision-making relationship was significantly related to age: $r(63)$ = 0.09, p = .48 for prior stimulus and $r(63)$ = 0.04, p = .73 for prior rating. Finally, given that participants viewed emails randomly sampled from a large pool, we examined whether older age was, by chance, associated with seeing emails that are more difficult to judge. Supplementary Figure 2 shows that there was no age-associated bias in difficulty of emails.

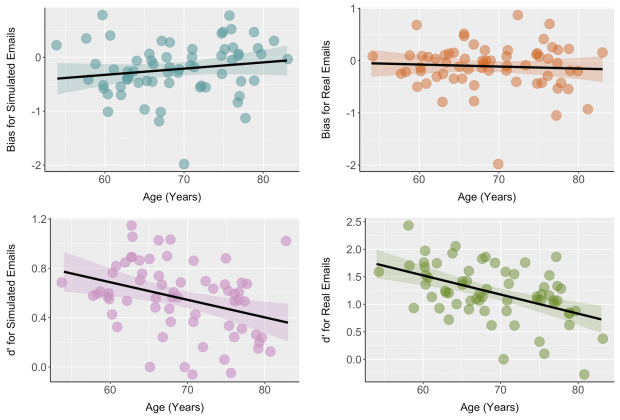

Figure 1. Relationships between older age and signal detection theory metrics of response bias and discrimination (d'). Correlation between age (years) and response bias and d' for simulated emails and real emails. The shaded regions around the regression lines reflect the 95% confidence interval.

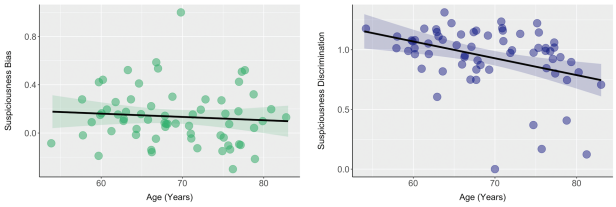

Figure 2. Relationships between older age and computational modeling metrics of response bias and discrimination. Correlation between age (years) and suspiciousness response bias (beta weight) and discrimination bias (beta weight). The shaded regions around the regression lines reflect the 95% confidence interval.

Discussion

Prior work suggests that older adults might be more vulnerable to falling for phishing emails under certain circumstances, although outcomes are mixed (Baillon et al., 2019; Ebner et al., 2020; Gavett et al., 2017; Lin et al., 2019; Sheng et al., 2010). To better understand one factor that might increase vulnerability to phishing messages, we investigated the relationship between older age and suspiciousness toward phishing and genuine emails. On a difficult email judgment task, we found that older age was not related to perceiving emails as generally safer to engage with (i.e., response bias). However, older age was associated with a reduction in the ability to discriminate between genuine and phishing emails, meaning that safe emails were considered in an overly suspicious light, and phishing emails were given too much credibility. These findings, which held whether we dichotomized suspiciousness scores as correct or incorrect or examined for potential subtler shifts in suspiciousness using our computational modeling approach, indicate that with older age, suspiciousness is a less valid signal for judging email intent.

Given that we used a new laboratory-based task with initial ecological validity (Hakim et al., 2020), we provide novel evidence that cognitively healthy older adults may be more susceptible to real-world email scams because of their difficulty discriminating phishing from genuine emails. At a time when more people are online, this study gives important direction for future research on identifying profiles of higher vulnerability to online fraud. For example, it would be interesting to pair perceived suspicious judgments with decision-making behavior to reveal at what threshold of suspiciousness individuals are willing to engage with an email (e.g., is "possibly safe" safe enough to click a link?), and/or whether age affects this threshold. Future research could take a more individualized approach and examine the impact of contextual factors on suspiciousness and decision making (e.g., whether a person uses online banking). Also, we note that our sample was largely non-Hispanic White, well-educated, and most were women. A future study will need to examine a more demographically diverse group. Finally, a larger study could examine whether the age-associated reduction in email discrimination reflects differences in certain cognitive domains, such as executive functioning and episodic memory, or other factors, such as cohort effects, including online literacy. These directions aside, our findings support the idea that older age is linked to greater difficulty judging the credibility of emails, which may lead to both unfortunate clicks and dismissal of genuine messages.

Supplementary Material

Supplementary data are available at *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences* online.

Funding

This research was supported by the Evelyn F. McKnight Brain Research Foundation (AGR DTD 04-11-2018); the National Science Foundation (SBE-1450624); the National Institutes of Health (1R01AG057764); and the Arizona Department of Health Services.

Conflict of Interest

None declared.

Acknowledgments

The present study was not preregistered. We have made PEST freely available to the research community, with the Matlab code here: https://github.com/zmhakim/PEST. We thank Anna Robertson, Christopher Griffith, and Sean Thayer for help administering PEST.

Author Contributions

M. D. Grilli, R. C. Wilson, N. C. Ebner, S. J. Getz, and B. E. Levin conceived the study. R. C. Wilson conducted the computational modeling of the PEST data. K. S. McVeigh curated and analyzed the data with supervision from M. D. Grilli. Z. M. Hakim built PEST with supervision from R. C. Wilson and M. D. Grilli. A. A. Wank developed the study procedures and oversaw data collection, with supervision from M. D. Grilli. M. D. Grilli wrote the paper with input from all other authors. All authors approved the manuscript.

Data Availability

Data for the present study are deposited here: https://osf.io/bvc7k/?view_only=2c348e5c099044cdb225ee759a843a54.

References

Bailey, P. E., & Leon, T. (2019). A systematic review and meta-analysis of age-related differences in trust. *Psychology and Aging*, **34**(5), 674–685. doi:10.1037/pag0000368

Baillon, A., de Bruin, J., Emirmahmutoglu, A., van de Veer, E., & van Dijk, B. (2019). Informing, simulating experience, or

both: A field experiment on phishing risks. *PLoS One*, **14**(12), e0224216. doi:10.1371/journal.pone.0224216

Carr, J. (2011). *Inside cyber warfare* (2nd ed.). M. Loukides (Ed.). O'Reilly Media.

Ebner, N. C., Ellis, D. M., Lin, T., Rocha, H. A., Yang, H., Dommaraju, S., Soliman, A., Woodard, D. L., Turner, G. R., Spreng, R. N., & Oliveira, D. S. (2020). Uncovering susceptibility risk to online deception in aging. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, **75**(3), 522–533. doi:10.1093/geronb/gby036

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, **39**(2), 175–191. doi:10.3758/bf03193146

Gavett, B. E., Zhao, R., John, S. E., Bussell, C. A., Roberts, J. R., & Yue, C. (2017). Phishing suspiciousness in older and younger adults: The role of executive functioning. *PLoS One*, **12**(2), e0171620. doi:10.1371/journal.pone.0171620

Glisky, E. L. (2007). Changes in cognitive function in human aging. In D. R. Riddle (Ed.), *Brain aging: Models, methods, and mechanisms* (pp. 3–20). Taylor & Francis.

Green, D., & Swets, J. (1966). *Signal detection theory and psychophysics*. Wiley.

Grilli, M. D., Wank, A. A., Bercel, J. J., & Ryan, L. (2018). Evidence for reduced autobiographical memory episodic specificity in cognitively normal middle-aged and older individuals at increased risk for Alzheimer's disease dementia. *Journal of the International Neuropsychological Society*, **24**(10), 1073–1083. doi:10.1017/S1355617718000577

Hakim, Z. M., Ebner, N. C., Oliveira, D. S., Getz, S. J., Levin, B. E., Lin, T., Lloyd, K., Lai, V.T., Grilli, M.D., & Wilson, R. C. (2020). The Phishing Email Suspicion Test (PEST) a lab-based task for evaluating the cognitive mechanisms of phishing detection. *Behavioral Research Methods*, 1–11. doi:10.3758/s13428-020-01495-0

JASP Team. (2020). *JASP (version 0.12.2) [computer software]*.

Lin, T., Capecci, D. E., Ellis, D. M., Rocha, H. A., Dommaraju, S., Oliveira, D. S., & Ebner, N. C. (2019). Susceptibility to spear-phishing emails: Effects of internet user demographics and email content. *ACM Transactions on Computer-Human Interaction*, **26**(5), Article 32. doi:10.1145/3336141

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Sheng, S., Holbrook, M., Kumaraguru, P., Cranor, L., & Downs, J. (2010). Who falls for phish? A demographic analysis of phishing susceptibility and effectiveness of interventions. In CHI'10: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 373–382). doi:10.1145/1753326.1753383

Templeton, V. H., & Kirkman, D. N. (2007). Fraud, vulnerability, and aging: Case studies. *Alzheimer's Care Today*, **8**(3), 265–277. doi:10.1097/01.ALCAT.0000281875.55721.0f

Vishwanath, A., Harrison, B., & Ng, Y. J. (2018). Suspicion, cognition, and automaticity model of phishing susceptibility. *Communication Research*, **45**(8), 1146–1166. doi:10.1177/0093650215627483

Wilson, R. C. (2018). Sequential choice effects predict prevalence-induced concept change. doi:10.31234/osf.io/75bpy