Human Factors in Cybersecurity, Vol. 53, 2022, 151–160
https://doi.org/10.54941/ahfe1002212

Non-Experts' Perceptions Regarding the Severity of Different Cyber-Attack Consequences: Implications for Designing Warning Messages and Modeling Threats

Natalie Lodinger, Keith Jones, Akbar Siami Namin, and Benjamin Widlus

Texas Tech University, Lubbock, TX 79406, USA

ABSTRACT

Cyber-defenders must account for users' perceptions of attack consequence severity. However, research has yet to investigate such perceptions of a wide range of cyber-attack consequences. Thus, we had users rate the severity of 50 cyber-attack consequences. We then analyzed those ratings to a) understand perceived severity for each consequence, and b) compare perceived severity across select consequences. Further, we grouped ratings into the STRIDE threat model categories and c) analyzed whether perceived severity varied across those categories. The current study's results suggest not all consequences are perceived to be equally severe; likewise, not all STRIDE threat model categories are perceived to be equally severe. Implications for designing warning messages and modeling threats are discussed.

Keywords: Consequences, Severity, Warning messages, Threat modeling

INTRODUCTION

Cyber-defenders must account for users' perceptions of attack consequence severity. For example, they must consider those perceptions when designing warning messages and when modeling potential threats.

The cybersecurity warning message literature has produced three key recommendations: warnings should: 1) describe attack consequences (Bartsch et al., 2013; Hardee et al., 2006), 2) convey attack severity (Bartsch, et al., 2013; Bauer et al., 2013), and 3) align with how users think (Bartsch & Volkamer, 2013; Blythe & Camp, 2012). To do so, attack consequence severity must be described in a way that aligns with how users think. Otherwise, users may not trust the warning message (Bartsch et al., 2013; Ibrahim et al., 2010). Thus, cybersecurity warning message designers must account for users' perceptions of attack consequence severity.

Threat modeling identifies ways a given system can be compromised (Xiong & Lagerström, 2019). Users make decisions that threaten system security (Arief & Besnard, 2003), and their perceptions of severity influence those

decisions (Bartsch et al., 2013; Dodel & Mesch, 2017; Ng et al., 2009). For example, users who consider a consequence to be severe are more likely to try to prevent it than those who do not consider it to be severe (Ng et al., 2009). Thus, cybersecurity threat modelers must account for users' perceptions of attack consequence severity.

Research has investigated how users think about topics related to cybersecurity (e.g., Wash, 2010) and how users perceive the severity of phishing attack consequences (Foster et al., 2021). However, research has yet to investigate how users perceive the severity of a wide range of cyber-attack consequences. Without that knowledge, cyber-defenders cannot account for users' perceptions of attack consequence severity when designing warning messages or modeling threats.

To address that limitation, we had users rate the severity of 50 cyber-attack consequences. We analyzed those ratings to a) understand perceived severity for each consequence, and b) compare perceived severity across select consequences. Further, we grouped ratings into the STRIDE threat model categories (Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, & Elevation of Privilege; Xiong & Lagerström, 2019) and c) analyzed whether perceived severity varied across categories. The current study is the first to provide specifics regarding non-experts' perceptions regarding a wide range of cyber-attack consequences.

METHOD

Participants

Two hundred and one students participated for course credit. Thirty-two participants were removed from the sample because they did not complete the study. Two more participants were removed because they responded with the same answer to every question, which suggested careless responding (Johnson, 2005). The resultant sample included 167 participants (99 female, 67 male, 1 did not report). Their ages ranged from 17 to 41 years ($M = 20.77$, $SD = 3.33$). None reported working or having taken a college-level course in any field related to computer security or privacy.

Surveys

The first survey consisted of instructions and a set of 50 questions. Each question included a description of one cyber-attack consequence and a 7-point Likert scale that ranged from "not severe" (1) to "severe" (7). Participants were instructed to base their rating on the worst possible outcome of that consequence. Appendix A provides consequence descriptions. Each was written in non-technical language and described how users would be affected. The second survey concerned demographics and included questions about the participant's age, gender, and experience working in or taking classes about computer security.

Procedure

The research complied with the APA Code of Ethics, and was approved by the Texas Tech Institutional Review Board. Each participant 1) provided

informed consent, 2) completed the perceived severity survey, 3) completed the demographics survey, and 4) received partial course credit.

RESULTS

Perceived Severity of Individual Consequences

We used bootstrapping (1000 samples; sampled with replacement; sample size = 167) to compute a mean and confidence interval for perceived severity for each consequence (see Figure 1). We did so to provide the best possible estimate of the population mean for each consequence.

Figure 1 reveals that all but one CI are above the severity scale midpoint (4). The lower limit of the exception is only slightly below the midpoint (Lower limit = 3.98). Thus, essentially all consequences were perceived to be at least moderately severe.

Nine of the fifty consequences (18%) had CIs that fell below 5. Most concerned the cyber-attacker affecting the functioning of a device or Web site (Consequences 5, 7, 8, 16, 17, 48). This suggests participants were concerned, but not particularly so, about such disruptions or inconveniences.

Only one consequence (Consequence 49) had a CI above 6. It concerned financial consequences, which suggests participants considered financial consequences to be fairly severe. This is consistent with studies that found protecting financial information is important to non-experts (Bartsch & Volkamer, 2013; Hardee et al., 2006), and extends that work by quantifying the level of perceived severity.

Comparing Perceived Severity of Individual Consequences

We did not compare each consequence against every other consequence. That would have required 1225 comparisons. We did not think it would be meaningful to analyze and interpret so many comparisons, especially given Type I error inflation.

Instead, we selected three comparisons based on Figure 1. These comparisons were performed with the raw, rather than bootstrapped, data. Non-bootstrapped means and their associated CIs can be found in Appendix A. Wilcoxon signed-rank tests were employed. The Bonferroni corrected p-value was 0.017 (i.e., .05/3).

The first test compared perceived severity ratings for the highest rated consequence, i.e., the consequence related to finances (Consequence 49), to those for the next highest rated consequence (Consequence 42). The former were significantly greater than the latter ($V = 621$, $p < .001$). Therefore, participants perceived financial consequences as more severe than all other consequences.

The second test compared perceived severity ratings for an attacker accessing information on an Internet site to those for an attacker deleting such information (Consequence 14). Multiple consequences concern an attacker accessing information online. We selected one of those consequences (Consequence 3) to conduct this test because it most directly states the attacker accesses information online without stating other behaviors the attacker

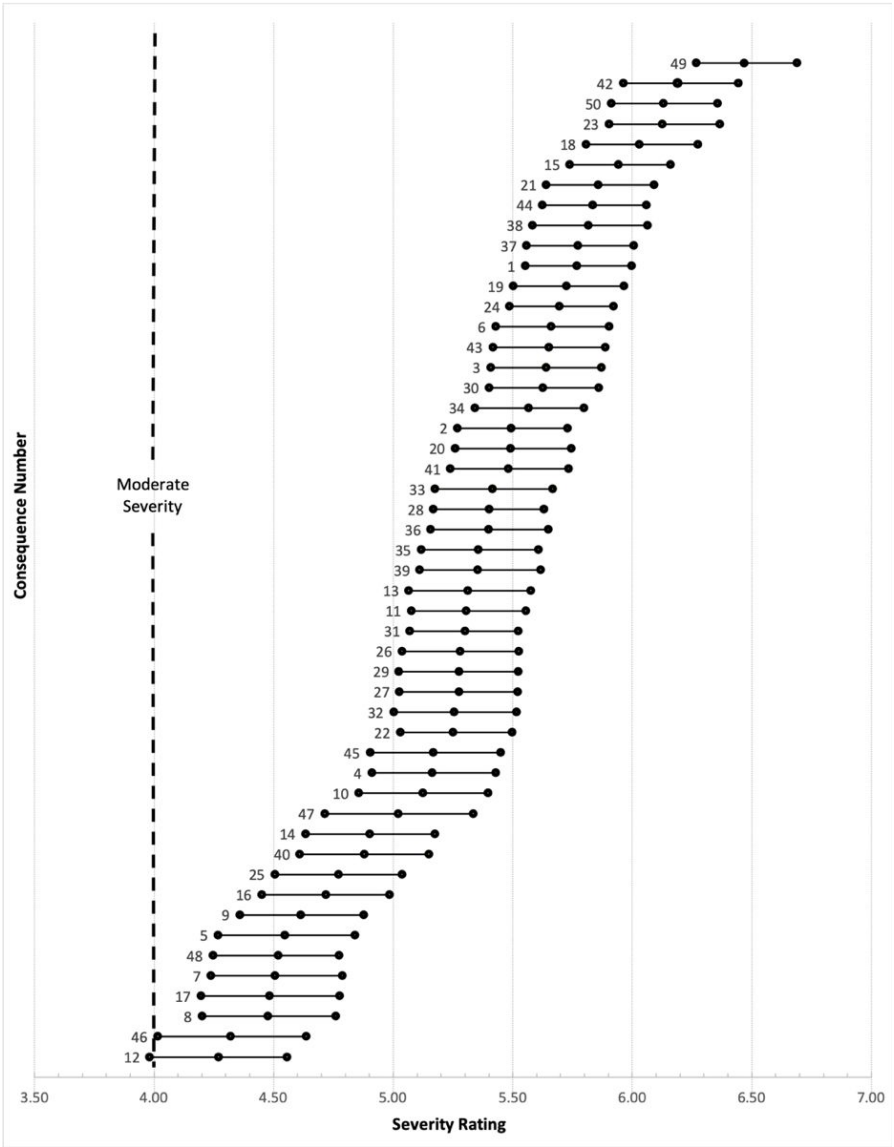

Figure 1: The bootstrap means and 95% confidence intervals of the perceived severity ratings for individual consequences. The severity rating scale ranged from not severe (1) to severe (7). Descriptions of each consequence can be found in Appendix A.

would take with that data. Participants rated an attacker accessing information as significantly more severe than deleting that information ($V = 1284$, $p < .001$). Thus, the various things an attacker could do to or with accessed information were perceived as more severe than them deleting that information.

The third test compared perceived severity ratings for an attacker logging into the user's computer (Consequence 41) to those for the attacker logging into the user's Internet account (Consequence 40). These specific consequences were chosen because they differed from each other mainly in terms

Table 1. Consequences that fell into each STRIDE category.

Category	Consequence #s
Spoofing	10, 11, 16, 20, 26, 27, 28, 31, 44
Tampering	3, 5, 6, 9, 12, 14, 21, 24, 25, 30, 31, 34, 35, 36, 38, 48
Repudiation	18, 27, 32, 33, 34, 43, 49
Information Disclosure	1, 6, 19, 22, 29, 31, 37, 45, 46, 47
Denial of Service	4, 5, 6, 7, 8, 13, 17, 23, 31, 39, 40, 41, 42, 50
Elevation of Privilege	1, 2, 3, 6, 7, 12, 15, 30, 31, 37, 38

of whether a device or Internet account was being accessed. Ratings for the attacker logging into a computer were significantly higher than those for an attacker logging into an Internet account ($V = 3632.5$, $p < .001$). Thus, an attacker gaining access to a device was perceived as more severe than them gaining access to an online account.

Comparing Perceived Severity of STRIDE Categories

To group the 50 consequences into the six STRIDE categories, we placed each consequence in a STRIDE category if a cyber-attack in that category would lead to that consequence. For example, the consequence "the cyber-attacker intercepted Internet traffic as it passes between your computer and the Internet" (Consequence 22) was placed in the information disclosure category because it is a consequence of a man-in-the-middle attack. Eleven consequences were grouped into multiple STRIDE categories because they could occur from attacks in multiple categories. A list of consequences in each STRIDE category is presented in Table 1.

The raw (not bootstrapped) mean of the perceived severity ratings for each consequence were averaged across consequences within a given category to create an overall mean of the perceived severity rating for that category. Those averages were subjected to a one-way analysis of variance (ANOVA) with STRIDE category as the independent variable and mean severity rating as the dependent variable. The main effect of STRIDE category was significant, $F(5, 830) = 34.30$, $p < .001$, $\eta_p^2 = .17$.

A Tukey HSD test revealed repudiation had the highest mean rating (5.67), which was significantly different from mean ratings for all other STRIDE categories. Therefore, consequences associated with repudiation were perceived to be more severe than consequences associated with all other STRIDE categories. However, the mean severity rating for repudiation was high partly because one consequence in this category concerned financial consequences ("The cyber-attacker took control over one of your financial accounts." 6.47). When this consequence was removed from the analysis, the mean severity rating for repudiation (5.53) was no longer significantly different from that for elevation of privilege (5.44), $t(332) = 0.72$, $p = .474$. As such, it appears that forms of repudiation that do not concern one's finances were not perceived as more severe than elevation of privilege. The Tukey HSD test also revealed that mean ratings for elevation of privilege (5.44) were significantly different from those for spoofing (5.31), denial of service (5.26), and

tampering (5.22). Elevation of privilege differing from spoofing may reflect that non-experts are not as aware of the risks involved in impersonation attacks compared to experts (Bartsch & Volkamer, 2013). Elevation of privilege differing from denial of service and tampering suggests that perhaps participants perceived consequences related to denial of service and tampering to largely be temporary nuisances.

DISCUSSION

Implications for Warning Message Design

The current study is the first to provide specifics regarding non-experts' perceptions regarding a wide range of cyber-attack consequences. The current study's results suggest not all consequences are perceived to be equally severe. Thus, although warning messages should describe personal consequences, designers need to consider non-experts' perceived severity of the consequence. Matching the wording of the severity of the consequence to the users' perceived severity could prevent distrust of warning messages and improve compliance (Bartsch et al., 2013).

Furthermore, descriptions of certain consequences may need to be more elaborate than others to provide the information non-experts need to appropriately assess consequence severity. Participants perceived consequences that, at face value, seemed quite similar to one another to be different from one another in terms of severity. For example, participants perceived an attacker logging into their computer as significantly more severe than an attacker logging into one of their Internet accounts. Accordingly, the information in warning messages should not be general (e.g., "the attacker could log into your system"). Rather, it should contain sufficient detail so that users understand the specific nature of the attack consequence (e.g., "the attacker could log into your computer").

Implications for Threat Modeling

The results of the consequences grouped into the STRIDE categories provide useful information for threat modelers. Perceived severity is a moderator for user security behavior (Ng et al., 2009). Therefore, threat modelers can use perceived severity ratings to make better predictions about cyber-attacks non-experts are more likely to protect against (Dodel & Mesch, 2017; Ng et al., 2009). The current study's results suggest users are more likely to protect against cyber-attacks that affect their money or property and repudiation attacks. However, users are less likely to protect against spoofing attacks or attacks that affect certain functions of their device because users do not think consequences of those attacks are as severe. Armed with this information, threat modelers can make better predictions of when users will exhibit more secure behaviors and create threat models that better predict user behavior.

When new consequences of cyber-attacks occur, threat modelers can use the STRIDE category perceived severity ratings to make predictions about non-experts' perceived severity of the new consequence. The new consequence can be placed into the correct STRIDE group, and the perceived severity

rating for that group can be applied to the consequence. Therefore, threat modelers can have a way to estimate non-experts' perceived severity and likelihood of exhibiting safe security behaviors without collecting data about the perceived severity of the new consequence.

ACKNOWLEDGMENT

This research was supported in part by the U.S. National Science Foundation (Award #: 1564293). Opinions, findings, and conclusions are those of the authors and do not necessarily reflect the views of the NSF.

REFERENCES

Arief, B. and Besnard, D., 2003. Technical and human issues in computer-based systems security. *School of Computing Science Technical Report Series*.

Bartsch, S. and Volkamer, M., 2013. Effectively communicate risks for diverse users: A mental-models approach for individualized security interventions. *INFORMATIK 2013–Informatik angepasst an Mensch, Organisation und Umwelt*.

Bartsch, S., Volkamer, M., Theuerling, H. and Karayumak, F., 2013, June. Contextualized web warnings, and how they cause distrust. In *International conference on trust and trustworthy computing* (pp. 205–222). Springer, Berlin, Heidelberg.

Bauer, L., Bravo-Lillo, C., Cranor, L. and Fragkaki, E., 2013. Warning design guidelines. *'^'eds.'): Book Warning Design Guidelines, Carnegie Mellon University, Pittsburgh, PA*.

Blythe, J. and Camp, L.J., 2012, May. Implementing mental models. In *2012 IEEE symposium on Security and privacy workshops* (pp. 86-90). IEEE.

Dodel, M. and Mesch, G., 2017. Cyber-victimization preventive behavior: A health belief model approach. *Computers in Human behavior*, *68*, pp. 359–367.

Foster, E. K., Jones, K. S., Armstrong, M. E., & Namin, A. S. (2021, July). User Perceptions of Phishing Consequence Severity and Likelihood, and Implications for Warning Message Design. In *International Conference on Applied Human Factors and Ergonomics* (pp. 265–273). Springer, Cham.

Hardee, J.B., West, R. and Mayhorn, C.B., 2006. To download or not to download: an examination of computer security decision making. *interactions*, *13*(3), 32–37.

Ibrahim, T., Furnell, S.M., Papadaki, M. and Clarke, N.L., 2010, August. Assessing the usability of end-user security software. In *International conference on trust, privacy and security in digital business* (pp. 177–189). Springer, Berlin, Heidelberg.

Johnson, J. A. (2005). Ascertaining the validity of individual protocols from web-based personality inventories. *Journal of research in personality*, *39*(1), 103–129.

Ng, B.Y., Kankanhalli, A. and Xu, Y.C., 2009. Studying users' computer security behavior: A health belief perspective. *Decision Support Systems*, *46*(4), 815–825.

Xiong, W., & Lagerström, R. (2019). Threat modeling–A systematic literature review. *Computers & security*, *84*, 53–69.

Wash, R., 2010, July. Folk models of home computer security. In *Proceedings of the Sixth Symposium on Usable Privacy and Security* (pp. 1–16).

APPENDIX A

Cyber-attack consequences and non-bootstrapped perceived severity ratings. All consequences began with "The cyber-attacker …".

			95% CI	
#	Consequence	Mean (SD)	Lower	Upper
1	accessed your computer files.	5.77 (1.46)	5.55	6.00
2	accessed your computer programs.	5.50 (1.50)	5.27	5.73
3	accessed your information stored in an Internet site.	5.64 (1.54)	5.40	5.88
4	caused a program on your computer to crash.	5.17 (1.67)	4.91	5.42
5	caused your computer program to run very slowly.	4.55 (1.85)	4.27	4.83
6	caused your computer to crash.	5.66 (1.51)	5.43	5.89
7	caused your computer to run very slowly.	4.51 (1.76)	4.24	4.78
8	caused your Internet connection to run very slowly.	4.47 (1.79)	4.21	4.75
9	caused your request for a certain Internet page to actually take you to a different Internet page.	4.62 (1.69)	4.36	4.87
10	changed a device's serial number.	5.13 (1.78)	4.85	5.40
11	changed how an Internet service functions to benefit the attacker.	5.31 (1.58)	5.07	5.55
12	changed the appearance of an Internet site.	4.27 (1.83)	3.99	4.55
13	completely filled your computer's storage space.	5.32 (1.63)	5.07	5.57
14	deleted your information stored in an Internet site.	4.90 (1.81)	4.63	5.18
15	determined your password.	5.95 (1.35)	5.74	6.15
16	disrupted the availability of an Internet service.	4.72 (1.71)	4.46	4.98
17	floods your inbox with a very large number of emails.	4.49 (1.85)	4.20	4.77
18	forged a digital signature on an electronic document.	6.04 (1.51)	5.80	6.27
19	gained information about existing hidden pathways by which they can enter your system from the Internet.	5.73 (1.51)	5.50	5.96

(Continued)

#	Consequence	Mean (SD)	95% CI	
			Lower	Upper
20	gained information about the device that you use to create your home network.	5.50 (1.56)	5.26	5.74
21	gains your username and password for a given Internet site.	5.86 (1.47)	5.64	6.09
22	intercepted Internet traffic as it passes between your computer and the Internet.	5.26 (1.53)	5.02	5.49
23	irreparably damaged your computer's hardware.	6.13 (1.51)	5.90	6.36
24	issued commands to your computer's operating system.	5.70 (1.42)	5.48	5.92
25	made an Internet page that you use act differently than intended.	4.77 (1.74)	4.51	5.04
26	made you think an Internet site that the attacker created was a legitimate Internet site.	5.28 (1.57)	5.04	5.52
27	made you think that an email that you received from the attacker came from someone else.	5.28 (1.66)	5.02	5.53
28	made you think that information sent to your Internet browser came from a trusted source.	5.40 (1.52)	5.17	5.63
29	made you think that you had a secure connection to an Internet site when it was not secure.	5.28 (1.64)	5.03	5.53
30	made your computer perform tasks that benefit the attacker.	5.63 (1.49)	5.40	5.86
31	made your computer run software that your computer did not intend to run.	5.30 (1.53)	5.06	5.53
32	made your computer think that your password was entered when it was not.	5.26 (1.66)	5.00	5.51
33	modified the content of a digital message without your awareness.	5.42 (1.60)	5.17	5.66
34	modified your computer files to hide their activities.	5.57 (1.50)	5.34	5.80
35	modified your information stored in an Internet site.	5.36 (1.59)	5.12	5.60
36	modified your information within an Internet database.	5.40 (1.60)	5.16	5.65
37	opened new hidden pathways by which they can enter your system from the Internet.	5.78 (1.45)	5.56	6.00
38	performed actions on an Internet site as if they were you.	5.82 (1.57)	5.58	6.06

(Continued)

#	Consequence	Mean (SD)	95% CI Lower	95% CI Upper
39	prevented you from accessing your home network.	5.36 (1.62)	5.11	5.61
40	prevented you from logging into an Internet site.	4.88 (1.71)	4.62	5.14
41	prevented you from logging into your computer.	5.49 (1.61)	5.24	5.73
42	prevented you from using your computer until you pay a ransom.	6.20 (1.55)	5.96	6.43
43	removed your computer files to hide their activities.	5.65 (1.55)	5.42	5.89
44	rerouted your Internet requests to a device that they control.	5.84 (1.44)	5.62	6.06
45	saw what was presented on your computer screen.	5.17 (1.84)	4.89	5.45
46	sent you an email that asks you to click on a given Internet link.	4.32 (2.03)	4.01	4.63
47	sent you an email that asks you to respond with certain personal information.	5.02 (1.97)	4.72	5.33
48	shut down an Internet site that you were using.	4.51 (1.78)	4.24	4.79
49	took control over one of your financial accounts.	6.47 (1.37)	6.26	6.68
50	used your computer to store and distribute stolen software.	6.13 (1.47)	5.91	6.36