Lessons for Labeling from Risk communication

August 17, 2021

Abstract

Labeling for software and Internet of Things devices (including the software embedded within them) will better meet the goals of the Executive Order on Improving the Nation's Cybersecurity [17] if the design and implementation integrates previous results on risk communication and warnings. Here we provide a partial summary of previous research in online risk communication that could inform the design of effective, usable, and helpful labels. We also identify a handful of classic works on warnings to provide a potentially useful foundation for exploring the wealth of findings which might be helpful in designing labels for software and the devices in which is it embedded.

Introduction

Decision-making in the face of privacy and security risks can be considered holistic risk communication [16]. The harms resulting from exposure to the risks are stochastic, and vary based on the vulnerability of the exposed. The results of the same security flaw or privacy lapse may range from simply annoyance to severe loss. Risk communication offers guidance for the creation of labels where the vulnerability to exposure and ability to mitigate can vary widely. Labels that are grounded in risk communication offer the potential to impinge the risk perception of those who interact with them.

There is the potential of a virtuous circle in security and privacy if effective labels support changes in risk perception and decision-making. Such changes in decision-making would reflect changes in habits of use and expectation. Since habits of use, perceived risk, and dimensions of the underlying risk predict cautionary behavior, a change in use that results from clear risk communication has the potential to impinge future choices and behaviors [24].

Background and Recommendations

Loss of privacy and loss of security are risks. As a result, the development of labels should be informed by risk communication and warning science. In this submission we identify previous work leveraging risk communication in security and privacy to support this core recommendation. Based on this foundation we make three central observations, and close with recommendations.

Timing matters. Warnings are provided before risk exposure, to enable risk avoidance and harm minimization. In the case of IoT and software, this means that labeling information should be provided during purchase, before installation or use. The importance of providing labels before exposure has been reified in studies on privacy permissions manifests on mobile devices. Prior research has found that the iOS vetting and run-time warnings were less effective than Android's community ratings and permissions manifest mechanism [10, 20]. Previous work with apps found that when permissions were included in the app description page instead of being presented after people chose to install an app, people chose apps that had fewer permissions [18, 15]; later work found stronger evidence in IoT selections [14]. There is also evidence that people are desensitized toward repeated run-time dialogs that are widely used on mobile, stationary, and embedded devices [1, 25]. Multiple research investigations with a range of methods has found that people may view these warning dialogs as yet another interruption rather than meaningful alerts, and simply click through them [23, 29, 9, 5, 8].

Multi-layer warning systems can empower a wide range of stakeholders. The same software and devices may be used in homes with only non-expert consumers, or in workplaces with dedicated technical staff. Labels designed for the average consumer will exclude the most vulnerable half of the population by definition. Here warning science offers a solution: when there is a highly variable audience, effective warnings are those designed for the low-end extreme to include the entire population [28]. As a result, multi-layer labels can be designed as warning systems to serve both the expert and the non-expert.

Simple interaction are needed to inform those with the least expertise. These may be augmented with audio, first of course to ensure accessibility. Secondly, a combination of visual indicators to support informed decision-making with an audio to increase risk awareness can create an effective labeling and warning system [26, 28, 21].

Yet simple icons and audio indicators can fail to serve the needs of experts. These also may not perfectly align in customizable systems. For example, there may be products which offer far greater functionality yet require expertise to enable these functions. Such products could be more safe in the default configuration yet more risky when used fully enabled. A multi-layer warning system could communicate information about both situations. Concerns about non-expert consumers may be mitigated with secure configurations, with the hope that the less expert may either not change the default or seek information. In terms of the first, the power of defaults in security is substantial. In terms of the second, when a product is perceived to be complex then the less expert users have long been found to be more likely to read instructions [19].

Align with the mental models of less expert users. Repeated experiments and evaluations on the mental models of computer security risks illustrates that expertise has an impact. More expertise corresponds to perceptions of risks as ubiquitous, with mental models that correspond to consistent awareness and severe results (e.g., health, grievous bodily harm). Those with less expertise more frequently choose mental models that correspond to less severe risks (i.e., nuisance crimes) or risks when the victim has no ability to mitigate (i.e., war) [2, 6, 7].

Even among those with less expertise, different mental models can correspond with different behaviors in terms of risk avoidance and effective mitigation [27]. Thus selecting a model that is both salient and aligned with safe behavior holds the most promise. Such an approach has been evaluated in the context of of web browsing, where simple icons based on mental models has been found to improve online safety by changing risk behavior [4, 7].

Expertise also influences security perceptions, so that the less expert may engage in more risk-seeking behaviors. For example, Gallagher et al. showed that expertise level is a strong determinant of usage patterns in Tor [11]. Stanton et al. discussed the correlation between users' technological expertise and their security awareness [**?**]. Both found that expertise corresponded with more risk avoidance. Less expert consumers have the greatest need for guidance.

Flawed risk communication can result in perverse results. Security is not immune to the phenomena of risk compensation, where increases in safety technology do not result in decreased harm due to increases in risk-seeking behaviors. This phenomena was apparent in a study on privacy behaviors in social networks, where increasing knowledge about privacy without risk communication lowered risk perceptions [13]. This has also been seen in the mobile context, with a comparison of risk communication in Android. Interactions that focused on improving the permissions interface through simplified text, explanations with more text, and icons using eyes found that those in the group with just the icon accepted fewer high-privacy and high-security apps, while rejecting more relatively safe ones [3]. Evaluating the effect of labels on long-term behaviors requires evaluation of how these influence individual risk thermostats [22].

There are many explanations for the lack of adoption of technologies to enhance security and privacy: lack of concern, lack of usability of tools, or lack of knowledge about risk. An empirical evaluation of the three explanations found that lack of knowledge or risks was the driving factors for over-exposure [12]. Labels that are designed to embed effective risk communication may better enable consumers to engage in informed self-protection.

Recommendations

Security and privacy labels are risk communication. Thus the design of labels should be informed by design of warnings, and include multiple levels of communication. Labels should be available for evaluation and comparison before any risk is encountered, ideally at the time of purchase or selection, to be effective. Different label designs embed distinct conceptual models, and selection of these should be informed by mental models. Labels that simply list information may not be adequate, and may even create perverse effects. It is worth evaluating a modular approach which enables evaluating an entire product, platform, or device or a only a specific desired service. Any label must be evaluated with an awareness of the potential influences of biases in human decision-making in the face of uncertainty.

In summary we argue for multi-level as well as single-level warning systems that are presented at the time a purchase or download decision is made. User-centered design that embeds mental models can serve this goal. We strongly encourage that any design be subject to evaluation based on its impact on risk perception. Finally, we argue for public campaigns grounded in risk communication with a focus on severity of consequences of loss of digital privacy, security, and safety concurrent with the development and diffusion of labels.

References

[1] Bonnie Brinton Anderson, Jeffrey L. Jenkins, Anthony Vance, C. Brock Kirwan, and David Eargle. Your memory is working against you: How eye tracking and memory explain habituation to security warnings. *Decision Support Systems*, 92:3 – 13, 2016.

[2] Farzaneh Asgharpour, Debin Liu, and L Jean Camp. Mental models of security risks. In *International conference on financial cryptography and data security*, pages 367–377. Springer, 2007.

[3] K. Benton, L. J. Camp, and V. Garg. Studying the effectiveness of android application permissions requests. In *2013 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, pages 291–296, March 2013.

[4] Jim Blythe and L Jean Camp. Implementing mental models. In *2012 IEEE symposium on Security and privacy workshops*, pages 86–90. IEEE, 2012.

[5] José Carlos Brustoloni and Ricardo Villamarín-Salomón. Improving security decisions with polymorphic and audited dialogs. In *Proceedings of the 3rd Symposium on Usable Privacy and Security*, SOUPS '07, pages 76–85, New York, NY, USA, 2007. ACM.

[6] Jean Camp, Farzaneh Asgharpour, Debin Liu, and IN Bloomington. Experimental evaluations of expert and non-expert computer users' mental models of security risks. *Workshop on Economics of Information Security*, 2007.

[7] Sanchari Das, Jacob Abbott, Shakthidhar Gopavaram, Jim Blythe, and L Jean Camp. User-centered risk communication for safer browsing. In *International Conference on Financial Cryptography and Data Security*, pages 18–35. Springer, 2020.

[8] Steve Dodier-Lazaro, Ruba Abu-Salma, Ingolf Becker, and M Angela Sasse. From paternalistic to user-centred security: Putting users first with value-sensitive design. In *CHI 2017 Workshop on Values in Computing*. Values In Computing., 2017.

[9] Serge Egelman, Lorrie Faith Cranor, and Jason Hong. You've been warned: An empirical study of the effectiveness of web browser phishing warnings. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 1065–1074, New York, NY, USA, 2008. ACM.

[10] Zheran Fang, Weili Han, and Yingjiu Li. Permission based android security: Issues and countermeasures. *computers & security*, 43:205–218, 2014.

[11] Kevin Gallagher, Sameer Patil, and Nasir Memon. New me: Understanding expert and non-expert perceptions and usage of the tor anonymity network. In *Thirteenth Symposium on Usable Privacy and Security ({SOUPS} 2017)*, pages 385–398. USENIX Association, 2017.

[12] Vaibhav Garg, Kevin Benton, and L Jean Camp. The privacy paradox: a facebook case study. In *Telecommunications Policy Research Conference*, 2014.

[13] Vaibhav Garg and L Jean Camp. Cars, condoms, and facebook. In *Information security*, pages 280–289. Springer, 2015.

[14] Shakthidhar Gopavaram, Jayati Dev, Sanchari Das, and L Jean Camp. IoT Marketplace: Willingness-To-Pay vs. Willingness-To-Accept. In *Proceedings of the 20th Annual Workshop on the Economics of Information Security (WEIS 2021)*, 2021.

[15] Shakthidhar Reddy Gopavaram, Omkar Bhide, and L. Jean Camp. Can You Hear Me Now? Audio and Visual Interactions That Change App Choices. *Frontiers in Psychology*, 11:2227, 2020.

[16] D Henshel, MG Cains, B Hoffman, and T Kelley. Trust as a human factor in holistic cyber security risk assessment. *Procedia Manufacturing*, 3:1117–1124, 2015.

[17] The White House. Executive Order 14028 on Improving the Nation's Cybersecurity, 2021.

[18] Patrick Gage Kelley, Lorrie Faith Cranor, and Norman Sadeh. Privacy as part of the app decision-making process. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3393–3402. ACM, 2013.

[19] M Letho and JM Miller. *Warnings Volume 1: Fundamentals, Design, and Evaluation Methodologies.* Fuller Technical, 1986.

[20] Jialiu Lin, Shahriyar Amini, Jason I Hong, Norman Sadeh, Janne Lindqvist, and Joy Zhang. Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 501–510. ACM, 2012.

[21] D.S. Mileti and J.H. Sorensen. Communication of emergency public warnings: A social science perspective and state-of-the-art assessment. *Oak Ridge National Laboratories Technical Report*, (ON: DE91004981), 8 1990.

[22] Barry Pless. Risk compensation: Revisited and rebutted. *Safety*, 2(3):16, 2016.

[23] Joshua Sunshine, Serge Egelman, Hazim Almuhimedi, Neha Atri, and Lorrie Faith Cranor. Crying wolf: an empirical study of SSL warning effectiveness. In *Proceedings of the 18th Conference on USENIX Security Symposium*, SSYM'09, pages 399–416, Berkeley, CA, USA, 2009. USENIX Association.

[24] Paul Van Schaik, Jurjen Jansen, Joseph Onibokun, Jean Camp, and Petko Kusev. Security and privacy in online social networking: Risk perceptions and precautionary behaviour. *Computers in Human Behavior*, 78:283–297, 2018.

[25] Anthony Vance, Brock Kirwan, Daniel Bjornn, Jeffrey Jenkins, and Bonnie Brinton Anderson. What do we really know about how habituation to warnings occurs over time? A longitudinal FMRI study of habituation and polymorphic warnings. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, page 2215–2227, New York, NY, USA, 2017. Association for Computing Machinery.

[26] W Kip Viscusi and Richard J Zeckhauser. Hazard communication: Warnings and risk. *The Annals of the American Academy of Political and Social Science*, 545(1):106–115, 1996.

[27] Rick Wash. Folk models of home computer security. In *Proceedings of the Sixth Symposium on Usable Privacy and Security*, pages 1–16, 2010.

[28] Michael S Wogalter, Dave DeJoy, and Kenneth R Laughery. *Warnings and risk communication.* CRC Press, 2005.

[29] Haidong Xia and José Carlos Brustoloni. Hardening web browsers against man-in-the-middle and eavesdropping attacks. In *Proceedings of the 14th International Conference on World Wide Web*, WWW '05, pages 489–498, New York, NY, USA, 2005. ACM.