Poster: Factors Associated with Online Privacy Knowledge

Masooda Bashir
Library and Information Science
University of Illinois Urbana-Champaign
mnb@illinois.edu

Kevin Hoff
Information Trust Institute
University of Illinois Urbana-Champaign
hoff1@illinois.edu

Gahyun Jeon
Department of Psychology
University of Illinois Urbana-Champaign
jeon29@illinois.edu

1. INTRODUCTION

The recent widespread transition to cloud computing-based technologies has fundamentally altered the dynamics of digital privacy. An enormous amount of personal information is now stored and managed in the cloud, apart from users' computers. Although cloud service providers must generally have users consent to privacy notices before collecting any personal information, past research has shown that the vast majority of consumers do not read privacy notices before accepting them [1]. Thus, serious questions remain regarding what consumers actually understand about present-day privacy issues, such as how cloud computing works and what service providers are able to do with customer information stored in the cloud.

In order to investigate the above questions and more, we conducted an extensive examination of consumer privacy knowledge through an online survey. We also collected demographic information in order to assess the personal characteristics most associated with knowledge. In this poster abstract, we report the significant demographic differences found in our survey related to age, gender, and education. The findings can be applied to guide ongoing and future efforts aimed at promoting informed consumer decision-making through usability improvements to the current "notice and choice" approach to digital privacy in the United States [2].

2. METHOLODOGY

The survey link was distributed through email and social media at a large, Midwestern university in the United States. Survey completion was promoted through a prize drawing for a $10 Amazon gift card. All together, we collected 455 complete responses.

We thoughtfully constructed the knowledge-based questions with an eye towards minimizing bias and subjectivity. In order to better facilitate objectivity, most of the questions were written as third-person scenarios.

The survey contained 42 knowledge-based questions organized into five sections based on subject matter. The first section, Cloud Computing, contained ten questions that assessed respondents' basic knowledge of cloud computing and how it works. The next section, Online Security, also contained ten questions, but they focused on basic issues related to cyber security. There were seven questions in the third section, Internet Economics, which revolved around how companies make money off the internet. The fourth section, Educational Records, was comprised of six questions that asked about provisions of the Family Educational Records and Privacy Act (FERPA), a federal law that governs the privacy of educational records in the United States. The last section, Legal Aspects, contained nine questions that focused on the significance of digital privacy laws in the United States.

3. RESULTS

3.1 Demographic Overview

Of the 455 respondents, there were slightly more females (55%) than males (45%). In terms of age, 37% of respondents were between 18-21 years old, 28% were between 22-29, 16% were between 30-39, and 19% were age 40 or older. For undergraduate field of study, 20% of respondents reported "computer science" or "engineering" while the remaining 80% reported other fields. The sample was highly educated: 60% of respondents had completed an undergraduate-level degree and almost half of this figure (29% overall) had also completed a graduate-level degree.

3.2 Overall Knowledge Scores

Overall knowledge scores were computed by assigning equal weight to all 42 knowledge-based questions. The mean overall percentage score was 60%.

For education, we examined differences in overall scores based on undergraduate field of study and highest level of education. We found that the overall score of respondents with computer science or engineering majors was significantly different from all other respondents, F (1, 396) = 15.33, $p < .05$, such that computer science and engineering majors (M = 65.27, 95% CI = [62.71, 67.83]) scored about seven percentage points higher than everyone else (M = 58.31, 95% CI [56.59, 60.04]). For highest level of education, we grouped respondents into one of three ordinal categories (non-college graduate, college graduate, and advanced degree). Overall knowledge scores differed significantly across the three groups, F (2, 450) = 21.97, $p < .05$. Tukey post-hoc comparisons indicated that non-college graduates (M = 54.21, 95% CI [51.70, 56.73]) had significantly lower scores than both college graduates and those with advanced degrees, but college graduates did not differ from those with advanced degrees.

Overall knowledge scores also significantly differed across four age groups, F (3, 449) = 15.87, $p < .05$. As shown in Figure 1, the 18-21 year old group had the lowest knowledge scores. Tukey post-hoc comparisons confirmed that that this age group (M = 53.56, 95% CI [50.99, 56.12]) had significantly lower scores than the other groups, but the three remaining groups did not significantly differ from one another.

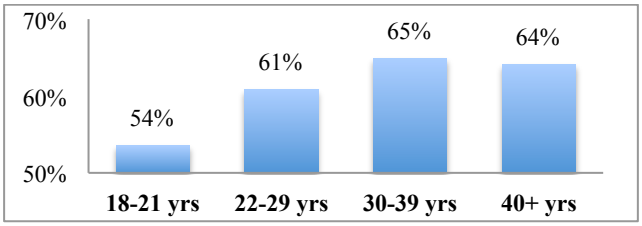

Figure 1. Mean overall knowledge scores by age group

Additionally, we investigated whether age and education level each predicted privacy knowledge controlling for each other. Results of multiple regression analyses indicated that age (b = 2.29, p < .05) and education level (b = 3.08, p < .05) both predicted overall scores. Overall scores also significantly differed across gender (F (1, 450) = 22.18, p < .05), such that men were more likely to have higher scores.

3.3 Section-by-Section Knowledge Scores

The survey was divided into five sections based on subject matter: 1. Cloud Computing (CC); 2. Online Security (OS); 3. Internet Economics (IE); 4. Educational Records (ER); and 5. Legal Aspects (LA). Overall, the mean scores were highest in the IE section (M = 77%) and lowest in the ER (M = 45%) and LA (M = 49%) sections. The average scores in the CC (M = 67%) and OS (M = 63%) sections were intermediary.

When investigating the effects of education level, one-way ANOVA results indicated that non-college graduates scored significantly lower than both college graduates and respondents with advanced degrees in the CC section, F (2, 450) = 22.83, p < .05, the OS section, F (2, 450) = 17.79, p < .05, the IE section, F (2, 450) = 19.08, p < .05, and the LA section, F (2, 450) = 7.57, p < .05. (See Figure 2 below.)

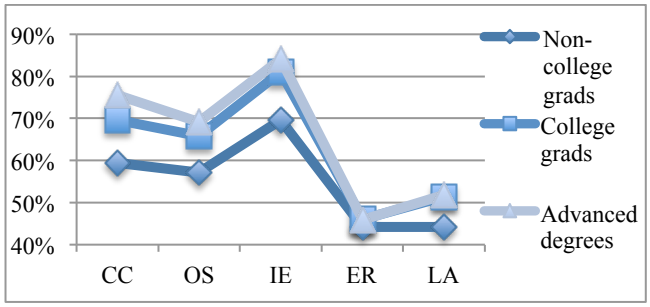

Figure 2. Mean section-by-section knowledge scores by education level

We also examined the effects of age and education level in predicting knowledge scores in each section. Based on the results of multiple regression analyses, we found that both variables simultaneously predicted CC scores (for age: b = 2.75, p < .05; for education level: b = 5.47, p < .05) and OS scores (for age: b = 2.93, p < .05; for education level: b = 3.28, p < .05), but only age was a significant predictor for IE scores (b = 5.09, p < .05).

For undergraduate field of study, we found that the division between engineering or computer science majors and all other respondents was a significant distinction in the CC, OS, and LA sections. One way ANOVA results indicated that computer science and engineering majors scored significantly higher in the CC section, F (1, 397) = 21.57, p < .05, the OS section, F (1, 397) = 22.92, p < .05, and the LA section, F (1, 397) = 9.13, p < .05.

As shown in Figure 3, knowledge scores were significantly different between men and women in every section except ER. One-way ANOVA results indicated that men scored significantly higher than women in the CC section, F (1, 450) = 33.57, p < .05, the OS section, F (1, 450) = 10.55, p < .05, the IE section F (1, 450) = 5.02, p < .05, and the LA section, F (1, 450) = 10.90, p < .05.

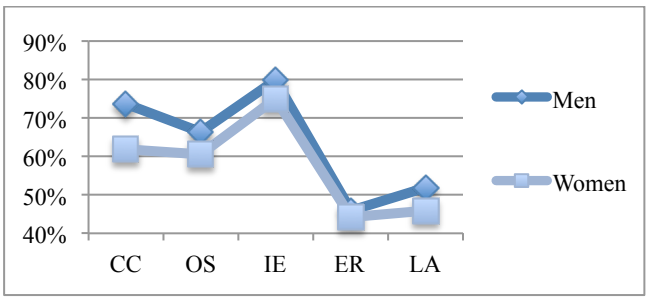

Figure 3. Mean section-by-section knowledge scores by gender

4. DISCUSSION

The demographic differences identified in our survey suggest that online privacy knowledge varies across age, gender, and education. The most significant finding in relation to age was that respondents between the ages of 18-21 years old (the approximate age of undergraduate students in the U.S.) performed substantially worse than all older respondents. These knowledge deficiencies may help explain the findings from past research that suggest that younger generations care less about their online privacy than older generations [3]. For education, we found that respondents with a college or graduate-level degree performed significantly better than non-college graduates, and knowledge scores were higher for respondents with a background in computer science or engineering. The survey results also suggest a potential gender discrepancy, with men have greater privacy knowledge than women. Still, men and women had similar patterns of knowledge scores across the survey sections so it seems that the same general knowledge gaps are present for both genders.

5. CONCLUSION

Based on the results of our survey, we think knowledge deficiencies are an important consideration which should be factored into the design of privacy policies. If policies better accommodated the preexisting knowledge levels of diverse individuals (e.g., across age, gender, and education), readership rates might increase. Service providers may also wish to consider knowledge levels in order to better address consumers' security concerns [4]. In addition, the results of our survey suggest that college students are particularly uninformed about privacy issues so future educational efforts should target this demographic group and younger generations.

6. REFERENCES

[1] Bakos, Y., Marotta-Wurgler, F., and Trossen, D. R. "Does anyone read the fine print? Testing a law and economics approach to standard form contracts", In CELS Conference on Empirical Legal Studies Paper, 9-40, 2009.

[2] Cranor, L.F. "Necessary but not sufficient: Standardized mechanisms for privacy notice and choice", *Journal of Telecommunications and High Technology Law*, 10(2), 2012.

[3] dos Santos Brito, K, Durao, F. A., Garcia, V. C., and Romero de Lemos Meira, S. "How people care about their personal data released on social media," in *Privacy, Security and Trust (PST)*, 111-118, 2013

[4] Nishioka, D., Saito, Y., & Murayama, Y. "Poster: The influence of the knowledge level about information security on Anshin factors." In *Symposium On Usable Privacy and Security (SOUPS), 2013.*