RepliPRI: Challenges in Replicating Studies of Online Privacy

Sameer Patil
Helsinki Institute for
Information Technology HIIT
Aalto University
Aalto 00076, FInland
sameer.patil@hiit.fi

Abstract
Replication of prior results has recently attracted attention and interest from the CHI community. This paper focuses on the challenges and issues faced in carrying out meaningful and valid replications of HCI studies. I attribute these challenges to two main underlying factors: (i) a domain of inquiry that simultaneously covers people, social systems, and technology; and (ii) deficiencies in result reporting and data archiving. Using examples from investigations of online privacy, I outline how these challenges manifest themselves in HCI studies. Longitudinal approaches, international collaboration, and sharing of study instruments could help address these challenges.

Author Keywords
Replication, Privacy, Cultural differences

ACM Classification Keywords
H.1.2 [User/Machine Systems]: Human factors.

General Terms
Human Factors, Security

Introduction
Replication of prior results has recently attracted attention and interest from the CHI community. The

resulting discussions tackle replication from two important perspectives: higher level epistemological debate on the place and merits of replication in the scientific (publishing) enterprise and the lower-level practical considerations for replicating previous studies from the literature. Growing interest in RepliCHI suggests increasing recognition for the value of replicating prior studies. I hope and anticipate that this trend will foster continued community discussion on how to justify, appreciate, and reward replication as a valuable scientific pursuit. Therefore, in this paper I focus on the latter aspect, viz., challenges and issues faced in carrying out meaningful and valid replications of HCI studies.

I attribute these challenges to two main factors:

1. Domain of inquiry: A large proportion of HCI studies tackle research problems where results typically exhibit simultaneous and interacting influence of individuals, social systems, and technology. Each of these three factors changes at drastically different rates and magnitudes. For instance, technology used in a study may become obsolete within months or a couple of years, while physical and cognitive capabilities of adults change at much slower rates (and the magnitude of the change is often comparatively small and predictable). These differences in the evolution trajectories of humans, cultures, and technology make it difficult to replicate studies at a later time and to determine and attribute causes behind differences in results, if any.

2. Insufficient and/or incomplete reporting: Typically the only resource available for replicating a study is the publication describing the results of the study. Unfortunately, due to page limits and other editorial reasons, publications often do not include all information — about methods and/or data — necessary for carrying out the study the way it was originally conducted. For instance, instead of including the entire questionnaire instrument, the publication may include only those questionnaire items that led to statistically significant results. Similarly, results may be presented in the aggregate or as percentages, making it difficult to replicate analyses that require details of individual data points.

In the following section, I outline how I have found these challenges to manifest themselves in investigation of user preferences and practices regarding online privacy. I conclude with some thoughts on addressing the challenges.

Replicating Studies of Online Privacy

When thinking about and carrying out replications of research related to privacy, I have encountered several practical challenges:

Privacy is a nuanced and complex issue affected by individual characteristics, context of operation, and the technology under consideration. For instance, individuals have been classified into different groups based on their inherent level of privacy concern [7], and privacy concerns have been shown to exhibit cultural variation [3]. People's mental models and understanding of the underlying technology also affects their preferences and practices regarding privacy [4]. This implies that even when considering the *same* technology, replication conducted at a later time ought

to take into account the impact of learning effects on privacy issues. Replications may also encounter the selection-maturation threat to validity owing to major external events that occur after the original study, such as news coverage of privacy breaches. Such events affect the population's overall understanding and awareness of privacy issues, thereby potentially affecting the results of replications of studies that were originally conducted prior to these event(s).

The majority of attention in replication has been devoted to replication at a different (later) *time*. In the case of privacy, however, it is equally important to consider replication across different *cultures*. For example, we administered a questionnaire simultaneously in the US and India, enabling us to draw interesting and surprising observations from comparison across cultures [5]. Our results confirmed earlier findings regarding low levels of consumer privacy concerns in India. Surprisingly, by examining interpersonal privacy separately from consumer privacy, we found that interpersonal privacy concerns in India were not only higher than consumer privacy concerns but also higher than interpersonal privacy concerns in the US. Our study considered culture at the broad level of national cultures. However, it should be noted that for replication purposes "culture" could be construed to connote any large groups with shared characteristics and/or values, such as students, engineers, mothers, liberals, etc. Moreover, if replication across cultures is conducted at a time later than the original study, then learning effects and maturation threats need to be taken into account (as discussed above).

In theory, replication with a different cultural sample is a simple case of re-running the study with subjects drawn from a different culture, with translation of instruments and study materials, if necessary. In practice, however, cultural differences pose several hurdles. For instance, the same word or term may be interpreted differently leading to the same question being answered differently. For example, we found that the term "cubicle" was understood differently in the US and India owing to differences in office layouts and density. This difference was one of the factors crucial for understanding the differences in results between the US and India [5]. In other studies, I discovered that the demographic question about ethnicity, which is commonly asked in the US (and even mandated for NSF-sponsored studies), was considered potentially offensive and confusing in Europe. Differences in lifestyle and beliefs can also affect whether questions and tasks from one study can yield valid results, or even make sense, when replicated in a different cultural context. For instance, some privacy studies have asked Western respondents about premarital sex, sexual practices, extramarital affairs, and number of sexual partners (e.g., [1]). Such questions are unlikely to produce meaningful results in cultures where such practices are uncommon and/or forbidden. Resolving this issue can be complicated when such culturally-specific questions comprise parts of standard scales; using the scale without modifications will not yield meaningful results and dropping and/or modifying items in the scale risks affecting the validity of comparison across studies. Finally, it is also necessary to consider whether results across cultures are affected by differences in sampling techniques and sample characteristics. For instance, although our comparison of the US and India was limited to software professionals, the mean and median ages of the Indian

participants were lower than those of the US participants.

We found that understanding privacy-related cultural nuance often requires insights derived from qualitative methods (such as interviews, focus groups, field visits, etc.) and/or insider knowledge of the culture and its practices [6]. Currently the CHI community is focused mostly on replication of studies that employ quantitative methods, such as experiments, questionnaires, or usability evaluations. Complementing quantitative replications with qualitative insights has potential to broaden the scope of these replication endeavors. Toward this end, it may also be fruitful to tackle whether and how qualitative studies could be effectively replicated.

Discussion and Conclusion

The previous section utilized examples from investigations of online privacy attitudes and behaviors to illustrate some of the challenges and issues in replicating HCI studies. Online privacy cuts across the individual, the social, and the technical, in much the same way as many studies in HCI do. Therefore, I believe that many, if not all, of these concerns are also likely to arise in HCI investigations of other topics.

The RepliCHI workshop is an important milestone toward developing a comprehensive compilation and understanding of various challenges involved in the replication of HCI studies. Moving forward, it is necessary to apply this knowledge and insight for constructing best practices to follow and pitfalls to avoid. Toward this end, I offer suggestions that address the two important considerations outlined in the Introduction, viz., (i) domain of inquiry that simultaneously covers individuals, social systems, and technology; and (ii) result reporting and data archiving.

The second of these, in particular, could be easily addressed by requiring inclusion of full instruments and study protocols as appendices[1]. Similarly, authors of accepted papers could be asked, or even required, to upload the raw data after taking steps necessary to protect participant anonymity. In this regard, ACM, IEEE, NSF, and other prominent HCI funding and sponsoring organizations can follow the lead of the NIH, which mandates raw data availability. In a similar vein, an open source inspired approach could encourage authors to release the source code of systems and scripts used for conducting studies and carrying out analyses. An open question regarding data and code sharing is how to deal with commercialization and intellectual property issues (especially when corporate entities are involved in conducting the study)[2].

One approach for addressing the issue of intersection of people and technology is to encourage longitudinal investigations carried out at regular intervals over several years. Depending on the details and logistics of the study, a longitudinal investigation could utilize the *same* participants or different participants with the same sampling method and sample characteristics. The former approach can help examine the impact of changes in individual characteristics, evolution in lifestyles, and effects of learning. The latter approach can help illuminate the impact of changes in

[1] This also provides the additional benefit of addressing one of the most common comments raised in peer reviews — lack of methodological detail.

[2] Data used by studies conducted by corporations was a hotly debated topic at the WWW 2012 conference [2].

technology. For replications across cultures, however, it is perhaps best to target simultaneous study deployment. Fostering international collaborations and/or leveraging international students to gain cultural knowledge and access could help in this regard.

Requiring a replication component in Bachelor's and Master's theses could provide a starting point for repeating studies from the literature, simultaneously serving a valuable pedagogical purpose by training the next generation. Further, conferences and journals could explicitly solicit replications of specific studies. Special conference sessions or journal sections could be devoted solely to replication studies. Discussions and follow-up activities from the RepliCHI workshop could lead the way toward legitimizing and promoting replication as a valuable scientific pursuit within HCI.

Acknowledgments

I thank Mihir Mahajan and John McCurley for editorial comments.

References

[1] Grossklags, J., and Acquisti, A. When 25 cents is too much: An experiment on willingness-to-sell and willingness-to-protect personal information. In *Workshop on the Economics of Information Security (WEIS)* (2007).

[2] Markoff, J. Troves of personal data, forbidden to researchers. http://www.nytimes.com/2012/05/22/science/big-data-troves-stay-forbidden-to-social-scientists.html, May 2012.

[3] Milberg, S., Burke, S., Smith, H., and Kallman, E. Values, personal information privacy, and regulatory approaches. *Communications of the ACM 38*, 12 (1995), 65–74.

[4] Patil, S., and Kobsa, A. Uncovering privacy attitudes and practices in Instant Messaging. In *Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work*, GROUP '05, ACM (New York, NY, USA, 2005), 109–112.

[5] Patil, S., Kobsa, A., John, A., and Seligmann, D. Comparing privacy attitudes of knowledge workers in the U.S. and India. In *Proceedings of the 3rd International Conference on Intercultural Collaboration*, ICIC '10, ACM (New York, NY, USA, 2010), 141–150.

[6] Patil, S., Kobsa, A., John, A., and Seligmann, D. Methodological reflections on a field study of a globally distributed software project. *Information and Software Technology 53*, 9 (2011), 969–980.

[7] Taylor, H. Most people are "privacy pragmatists" who, while concerned about privacy, will sometimes trade it off for other benefits. *The Harris Poll 17* (2003), 19.