DOI:10.1145/1787234.1787236

CS Expertise for Institutional Review Boards

IRBS NEED COMPUTER scientists, a point highlighted by the Viewpoint “Institutional Review Boards and Your Research” by Simson L. Garfinkel and Lorrie Faith Cranor (June 2010). Not just over the nature of certain CS-related research but because social scientists (and others) administer online surveys, observe behavior in discussion forums and virtual worlds, and perform Facebook-related research. In this regard, the column was timely but also somewhat misleading.

First, the authors created a dichotomy of computer scientists and IRBs, saying IRB “chairs from many institutions have told us informally that they are looking to computer scientists to come up with a workable solution to the difficulty of applying the Common Rule to computer science. It is also quite clear that if we do not come up with a solution, they will be forced to do so.”

However, any institution conducting a significant amount of human-subjects research involving computing and IT ought to include a computer scientist on its IRB, per U.S. federal regulations (45 CFR 46.107(a)): “Each IRB shall have at least five members, with varying backgrounds to promote complete and adequate review of research activities commonly conducted by the institution. The IRB shall be sufficiently qualified through the experience and expertise of its members...”

Though CS IRB members do not have all the answers in evaluating human-subject research involving computing and IT, they likely know where to look. It would also mitigate another problem explored in the column, that “many computer scientists are unfamiliar with the IRB process” and “may be reluctant to engage with their IRB.” Indeed, if an IRB member is just down the hall, computer scientists would likely find it easier to approach their IRB.

Second, the authors assumed the length of the IRB review process represents a problem with the process itself though offered only anecdotal evidence to support this assumption.

Two such anecdotes involved research on phishing, an intrinsically deceptive phenomenon. Deception research, long used in social sciences, typically takes longer to review because it runs counter to the ethical principle of “respect for persons” and its regulatory counterpart “voluntary informed consent.” Before developing a technical solution to perceived IRB delays, the typical causes of delay must be established. Possibilities include inefficient IRBs and uninformed and/or unresponsive researchers. Moreover, as with any deception research, some proposals may just be more ethically complex, requiring more deliberation.

Michael R. Scheessele, South Bend, IN

Authors' Response:

Scheessele is correct in saying an increasing number of social scientists use computers in their research and is yet another reason IRBs should strive to include a computer scientist as a member. Sadly, our experience is that most IRBs in the U.S. are understaffed, lack sufficient representation of members with CS knowledge, and lack visibility among CS researchers in their organizations.

Simson L. Garfinkel, Monterey, CA

Lorrie Faith Cranor, Pittsburgh, PA

How Many Participants Needed to Test Usability?

No usability conference is complete without at least one heated debate on participant-group size for usability testing. Though Wonil Hwang's and Gavriel Salvendy's article “Number of People Required for Usability Evaluation: The 10±2 Rule” (Virtual Extension, May 2010) was timely, it did not address several important issues concerning numbers of study participants:

Most important, the size of a participant group depends on the purpose of the test. For example, two or three participants should be included if the main goal is political, aiming to, say, demonstrate to skeptical stakeholders that their product has serious usability problems and usability testing can

find some of them. Four to eight participants should be included if the aim is to drive a useful iterative cycle: Find serious problems, correct them, find more serious problems.

Never expect a usability test to find all problems. CUE-studies¹ show it is impossible or infeasible to find all problems in a Web site or product; the number is huge, likely in the thousands. This limitation has important implications on the size of a test group. So go for a small number of participants, using them to drive a useful iterative cycle where the low-hanging fruit is picked/fixed in each cycle.

Finally, the number and quality of usability test *moderators* affects results more than the number of test participants.

In addition, from a recent discussion with the authors, I now understand that the published research in the article was carried out in 2004 or earlier and the article was submitted for publication in 2006 and accepted in 2008. All references in the article are from 2004 or earlier. The authors directed my questions to the first author's Ph.D. dissertation, which was not, however, included in the article's references and is apparently not available.

Rolf Molich, Stenløse, Denmark

Reference

1. Molich, R. and Dumas, J. Comparative usability evaluation (CUE-4). *Behaviour & Information Technology* 27, 3 (May 2008), 263–281.

Correction

“CS and Technology Leaders Honored” (June 2010) mistakenly identified the American Academy of Arts and Sciences as the American Association for the Advancement of Science. Also, it should have listed Jon Michael Dunn, Indiana University, as one of the computer scientists newly elected as an American Academy 2010 Fellow. We apologize for these errors.

Communications welcomes your opinion. To submit a Letter to the Editor, please limit your comments to 500 words or less and send to letters@cacm.acm.org.

© 2010 ACM 0001-0782/10/0800 \$10.00