

USER INTERFACE DESIGN PROCESS AND EVALUATION

AN AFTER-SCENARIO QUESTIONNAIRE FOR USABILITY STUDIES: PSYCHOMETRIC EVALUATION OVER THREE TRIALS

James R. Lewis
Design Center/Human Factors
IBM Boca Raton, Florida

INTRODUCTION

The purpose of this report is to (1) describe a psychometric evaluation of a printer scenario questionnaire (PSQ) that was used to assess user satisfaction during scenario-based usability studies of printers and (2) compare the psychometric properties of the PSQ with those of the After-Scenario Questionnaire (ASQ) (Lewis, 1991b). The PSQ is similar to the ASQ, but is composed of three five-point scales rather than three seven-point scales.

METHOD

A total of 70 users participated in the studies, ten users per printer. Seven printers were evaluated, and the studies had four scenarios in common. Each scenario was presented three times during the study. After participants finished a scenario, they completed the PSQ.

RESULTS

Based on previous research (Lewis, 1991b), items were hypothesized to cluster by scenario. Factor analyses confirmed this hypothesis. Coefficient alpha for the scale for each scenario ranged from .64 to .93, and averaged .80. These results indicated acceptable scale reliability, but less reliability than that reported for the After-Scenario Questionnaire (ASQ) (see Lewis, 1991b).

A three-way (Printer-by-Scenario-by-Trial) analysis of variance was conducted using the scale. Although the main effect of Printer was not significant ($F(6,42)=0.6, p=.71$), the Printer-by-Scenario interaction was ($F(18,108)=2.1, p=.01$). This result indicated that the questionnaire was sensitive to important independent variables.

Over all tasks and trials, the scale was significantly correlated with successful scenario completion ($r(69)=-.38, p=.0006$), the number of requests for assistance ($r(69)=0.42, p=.0001$) and the number of problems observed ($r(69)=-.31, p=.004$). These results indicated that the scale had concurrent validity with behavioral measures.

DISCUSSION

The results of this study of the PSQ are similar to those reported for the ASQ (Lewis, 1991b). The factor analyses showed the same pattern of association with scenarios, but extended through three trials. The analysis of variance had the same pattern as well, with no main effect of printer but a significant Printer-by-Scenario interaction. The correlation between the summed PSQ ratings and successful scenario completion for the first trial was -.38, very close to that of the ASQ. In the only discrepant result from the ASQ study, coefficient alpha for the PSQ ranged from .64 to .93. Coefficient alphas reported for the ASQ all exceeded .90. Because the internal consistency of a questionnaire is partially determined by the number of scale steps per item (Nunnally, 1978), it is likely that the smaller coefficient alphas found for the PSQ were primarily due to the use of five-point scales rather than seven-point scales.

CONCLUSION

This research showed that the usefulness of the ASQ is probably not dependent upon the order, format, or the exact wording of the items, and the ASQ should remain useful across multiple trials. Therefore, the ASQ (see Lewis, 1991b) should be seriously considered when choosing a measurement of participant satisfaction for scenario-based usability studies.

REFERENCES

- Lewis, J.R. (1991a). An after-scenario questionnaire for usability studies: Psychometric evaluation over three trials (IBM Tech. Report 54.584). Boca Raton, FL: International Business Machines Corp.
- Lewis, J.R. (1991b). A psychometric evaluation of an after-scenario questionnaire for computer usability studies: The ASQ. SIGCHI Bulletin, 23, 78-81.
- Nunnally, J.C. (1978). Psychometric theory. New York, NY: McGraw-Hill.

CONTACT INFORMATION

For more information about this study (Lewis, 1991a), contact James R. Lewis, IBM Design Center/Human Factors, P.O. Box 1328, Boca Raton, FL 33432, (407) 443-1066.