

Measuring Usability of a Depression Screening App: A 12-Week Follow-Up Study with Undergraduate Students

Kyung-mi Chung^{1,2}, Jin Young Park^{1,2}, Kyungun Jung³

¹Department of Psychiatry, Gangnam Severance Hospital, Yonsei University Health System
Seoul, South Korea

²Department of Psychiatry and Institute of Behavioral Science in Medicine, Yonsei University College of Medicine
Seoul, South Korea

chungkyungmi@yuhs.ac; empathy@yuhs.ac

³Department of Psychiatry, International St. Mary's Hospital, Catholic Kwandong University
Incheon, South Korea
kyungun12@gmail.com

Extended Abstract

The current study aims to measure usability improvement of a newly developed depression-screening app that was redesigned based on the usability problems encountered from usability testing (UT). It is well-known that ecological momentary assessment [1] contributes to reduced retrospective recall and associated biases, such as memory biases and self-concept biases, increasing ecological validity when employing self-report measures. To overcome the limitations of the most pencil-and-paper-based depression screening tests that retrospectively ask how often the respondents felt the depressive symptoms during the last two weeks (e.g., the Patient Health Questionnaire, PHQ-9 [2], the Center for Epidemiologic Studies Depression Scale-Revised, CESD-R [3], and the Korean version of CESD-R, K-CESD-R [4]), we implemented a mobile-optimized K-CESD-R app in which users could report their experienced symptoms in the last 24 hours, particularly within 'yes/no' binary choice options. After conducting task scenario-based UTs (a 1st 1-hour UT – a 2-week use of the app – a 2nd 30-min UT) with 5 young and 8 elderly adults, we redesigned the K-CESD-R Mobile app to universally correspond with both the young and the elderly's mental models.

To evaluate the improved usability of the app, 9 male and 5 female undergraduate students ($M = 24.00$, $SD = 1.41$; aged 22 to 28 years) were enrolled in a 12-week follow-up study. All participants were guided to install an earlier version of the K-CESD-R Mobile on their own smartphones and repeat a 2-week long test 5 times for a 10-week period. At the final day of the 10-week period, participants completed the pre-USE questionnaire [5] modified to comply with the purpose of this app. The USE questionnaire was composed of 4 factors: usefulness (USE), ease of use (EOU), ease of learning (EOL), and satisfaction (SAT). After removing the app in use and installing an updated version of the app, all continued to take the same once-a-day retrospective test for 2 weeks and filled out the same post-USE questionnaire on the last day.

A paired-samples *t*-test was performed to compare the mean scores of USE, EOU, EOL, SAT, and overall usability for the pre- and post-USE questionnaires rated on 5-point Likert scales. Compared to the responses to USE ($M = 3.699$, $SE = .226$) and EOU ($M = 4.012$, $SE = .178$) of the initial evaluation, those of the final evaluation were significantly improved: USE [$M = 4.074$, $SE = .121$; $t(13) = -2.758$, $p < .05$] and EOU [$M = 4.377$, $SE = .100$; $t(13) = -2.258$, $p < .05$]. The final overall usability ($M = 4.351$, $SE = .103$; $t(13) = -2.106$, $p < .055$) and the SAT ($M = 4.184$, $SE = .172$; $t(13) = -2.113$, $p = .055$) scores showed a trend in improvement compared to the initial overall usability ($M = 4.048$, $SE = .197$) and SAT ($M = 3.804$, $SE = .275$) scores. EOL showed no significant difference between the initial ($M = 4.679$, $SE = .145$) and final ($M = 4.768$, $SE = .096$) evaluation, $t(13) = -.616$, *NS*. Unlike other improved 4 factors, its learnability was already highly achieved before redesigning the app.

Acknowledgement

This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number : HM16C1991).

References

- [1] A. A. Stone and S. Shiffman, "Ecological momentary assessment (EMA) in behavioral medicine," *Annals of Behavioral Medicine*, 1994.
- [2] K. Kroenke, R. L. Spitzer, and J. B. Williams, "The Phq-9," *Journal of general internal medicine*, vol. 16, pp. 606-613, 2001.
- [3] W. W. Eaton, C. Smith, M. Ybarra, C. Muntaner, and A. Tien, "Center for Epidemiologic Studies Depression Scale: review and revision (CESD and CESD-R)," in *The Use of Psychological Testing for Treatment Planning and Outcomes Assessment*, M. E. Maruish, Ed., 3rd ed. : Instruments for Adults ed Lawrence Erlbaum, Mahwah, NJ, 2004, pp. 363–377.
- [4] S. Lee, S. T. Oh, S. Y. Ryu, J. Y. Jun, K. Lee, E. Lee, *et al.*, "Validation of the Korean version of Center for Epidemiologic Studies Depression Scale-Revised (K-CESD-R)," *Korean Journal of Psychosomatic Medicine*, vol. 24, pp. 83-93, 2016.
- [5] A. M. Lund, "Measuring usability with the use questionnaire12," *Usability interface*, vol. 8, pp. 3-6, 2001.