App-Based Diary Studies as an Alternative to Traditional Usability Testing for Mobile Applications

Ashley M. Schaad, Arlen Rosenthal, James Dayton | ICF
Daniel Baker | MFour Mobile Research

Dream big.
Then call ICF.

Contact
Ashley Schaad
ashley.schaad@icf.com

Abstract

Even though usability testing for mobile applications (apps) is an important step for ensuring adoption by the target audience, it is often skipped – particularly for small-scale apps and those with a low-incidence target audience. This may be due, in part, to the time and cost involved in conducting traditional laboratory-based, moderated usability testing. Unmoderated testing options, on the other hand, do not require as much researcher time and do not have the costs associated with laboratory testing. However, unmoderated testing generally involves screen and voice/video recording that requires specialized software and results in data that can be time-consuming to analyze. Here, we describe three mobile app usability tests conducted in a diary study format using an app-based survey panel with national representation. These unmoderated tests consisted of a series of daily testing scenarios over a one- or two-week testing period. We did not use screen-recording software; instead, screen shots were used to confirm that respondents had successfully completed the scenarios. Most questions in the scenarios were closed-ended, requiring little time to analyze. For the open-ended questions, respondents were randomly assigned to enter either text-based responses or to record video responses which were then converted to text for analysis. This allowed us to further examine the benefit – or lack thereof – of video-based responses as opposed to text-only methods.

Background

App Usability Testing: Usability testing evaluates a product or service by testing it with representative users. Typically, during such a test, a participant will complete typical tasks while an observer watches, listens, and takes notes. The goals are to identify any usability problems, collect qualitative and quantitative data, and determine the participant's satisfaction with the product.[1]

Traditional Moderated Usability Testing

- Involves direct observation of, and interaction with, the participant as s/he completes the tasks in real time.
- Oftentimes takes place in a laboratory environment, but can be done virtually.
- **Pros:** Allows for interaction between the researcher and the participant as tasks are completed.
- **Cons:** Costly; requires much logistical planning; can have limited geographic reach unless done remotely.

Traditional Unmoderated Usability Testing

- Participant completes tasks without direct observation by the researcher.
- Data typically include screen and video/voice recordings, and sometimes behavioral data (such as tracking hand movement on the screen).
- **Pros:** Can easily be done remotely; can be less costly.
- **Cons:** No direct participant-researcher interaction; requires specialized recording software; data can be difficult and time-consuming to analyze.

Common Usability Testing Outcomes and Issues

- In general, participants rate apps highly when asked to provide an overall rating.[2,3]
- Common usability complaints include issues with functionality, requested features, and app crashes.[4]
- Complaints most likely to affect overall app ratings include privacy/ethical concerns and hidden costs.[4]

Unmoderated App-based Diary Study Methodology

Using an unmoderated diary study methodology, we conducted usability testing studies to evaluate three mobile apps. Participants were recruited using the MFour Mobile Research panel, and all studies were conducted through MFour's Surveys on the Go (SOTG) mobile app. Each study consisted of a one- to two-week testing period during which participants received one to two scenario-based tasks to complete each day, followed by a short series of questions about their experiences completing each task. Follow-up questions included a mix of closed- and open-ended items. In addition, we administered first and final impression surveys on the first and final days of each study to gather over-arching feedback. For two of the studies, we also asked some of the diary study participants to take part in a telephone focus group after the diary study was completed to gather additional feedback and follow-up on diary study feedback. Specific information about each study is listed below, followed by an example scenario with follow-up questions.

Research Questions

- Can an unmoderated diary study methodology result in comparable feedback to that received using more traditional usability testing methods?
- How do video-based responses to open-ended questions compare to text-based responses in terms of length and data quality?

	Study 1	Study 2	Study 3
Topic of App	Opioid Overdose	Sexual & Reproductive Health	Smoking Cessation
Number of Participants	69	88	25
Study Length	1 week	1 week	2 weeks
Number of Diary Entries	7	7	5
Follow-up Focus Group?	Yes	No	Yes
Other Notes	Conducted in English and Spanish		Included Pre- and Post-test Surveys for Behavior and Attitude Changes

Scenario: Imagine your partner regularly uses heroin. In case your partner overdoses, you want to be prepared to recognize the signs. Use the app to find the signs of an opioid overdose. When you find the answer, take a screen shot of the page(s).

Follow-up Questions:

1. Were you able to find what you were looking for? (Yes/No)
2. Please upload the screenshot you took at the end of the scenario.
3. Was there enough information for you to recognize an overdose? (Yes/No)

<u>If no,</u>

1. What additional information would you like to see included on this topic?
2. Approximately how long did it take you to find the information? (Less than 1 minute, 1-2 minutes, 3-5 minutes, Longer than 5 minutes)
3. How difficult or easy was it to find the information? (Very difficult, Somewhat difficult, Neither difficult nor easy, Somewhat easy, Very easy)

<u>If Very difficult, Somewhat difficult, or Neither difficult nor easy,</u>

1. What could make the process of finding this information easier?

App-Based Diary Studies as an Alternative to Traditional Usability Testing for Mobile Applications

Ashley M. Schaad, Arlen Rosenthal, James Dayton | ICF
Daniel Baker | MFour Mobile Research

Dream big.
Then call ICF.

Contact
Ashley Schaad
ashley.schaad
@icf.com

1 Effectiveness of Unmoderated Diary Study Approach

The diary study methodology was successful in testing participants' interactions with various features of the app. The scenarios were short enough to yield high participation rates, and nearly all participants were able to successfully complete each of the scenarios. Overarching findings from the three studies include:

- All three apps were rated highly overall;
- No major usability issues were identified; and
- Each study revealed at least one feature of the app that should be modified, added, or eliminated.

In addition to the diary study components, the focus group component provided additional feedback and recommendations for improvements that went beyond the data gathered during the diary study.

2 Video-versus Text-based Responses

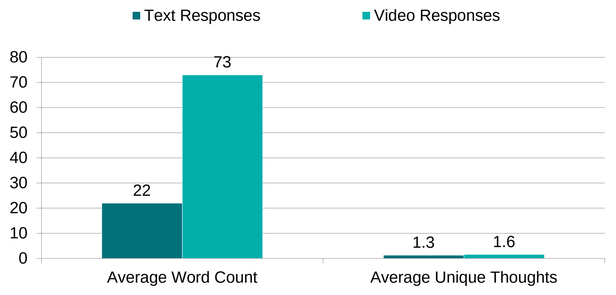
Two of the studies utilized both text-based responses (responses entered using the phone's keyboard) and video-based responses (responses captured using the phone's camera to record a video of the participant providing responses verbally). Participants were randomly assigned to one of the two conditions and maintained that condition throughout the full diary study. Only a portion of the open-ended questions were available for video responses, with an average of one to two video-based questions per scenario for those assigned to the video response condition.

We did not track drop-off rates for each condition, but the total number of participants in each condition who completed the study suggests that participants preferred the text response condition. Upon entering the studies, half of the participants were assigned to each condition. In total, 99 participants completed the studies in the text response condition, and 55 completed the studies in the video response condition.

Conclusions

- Diary-based usability testing can yield insights similar to those found with more traditional testing methods, particularly when combined with follow-up focus group discussions, and may be a less costly option.
- Given the extra time required to analyze video-based responses, the limited additional return in response quality, and an apparent dislike among participants, we suggest not assigning participants to complete the open-ended questions via video for this type of study. For studies where the data being collected via video will be used to validate a human participant or collect image-based data, however, video responses may still be useful.

Video response participants used 3 times the number of words but provided similar quality responses.



References

1. "Usability Testing." May 2, 2017. https://www.usability.gov/how-to-and-tools/methods/usability-testing.html.
2. Ickin, S., Wac, K., Fiedler, M., Janowski, L., Hong, J.-H., & Dey, A. (2012). Factors influencing quality of experience of commonly used mobile applications. *IEEE Communications Magazine*, 50(4), 48–56. https://doi.org/10.1109/MCOM.2012.6178833.
3. Kortum, P., & Sorber, M. (2015). Measuring the Usability of Mobile Applications for Phones and Tablets. *International Journal of Human-Computer Interaction*, 31(8), 518–529. https://doi.org/10.1080/10447318.2015.1064658
4. Khalid, H., Shihab, E., Nagappan, M., & Hassan, A. E. (2015). What Do Mobile App Users Complain About? *IEEE Softwa*re, 32(3), 70–77. https://doi.org/10.1109/MS.2014.50