

Addendum to Team R Report

Description of Evaluation Criteria

I reviewed each videotaped session with an eye toward positive and negative findings. I logged findings as I encountered them. Because the website was unfamiliar to me, I had to go back often in the review of the first user's session to understand what the user had done and what the result was. In the "real world," I would already be familiar with the interface and would have led the planning for the study, which would give me familiarity with the tasks as well as the results.

Normally, we would log the findings using Morae and review the logs in a findings meeting. For this review, I took notes on a legal pad, and then reviewed the notes for the compilation of top positive and negative findings.

For me, a usability problem is something that frustrates a user, causes a user to express a concern or a wish for a different way of doing things, or, in the worst case, prevents successful completion of a task. In this study, all users "successfully" completed the tasks, but they failed to order the right sized truck or the right sized storage unit, guessing in both cases as to the size they needed. This is a show-stopper problem that will result in very unhappy customers if the issue is not addressed.

Resources for CUE-9 in Person Hours

This project has been far more time consuming than I anticipated and, in fact, I have not been able to complete the full report I would normally deliver.

- Preparation time—40 minutes—to download videos, check audio and video quality; review instructions (completed on receipt of instructions in April). Note: user 1 audio was very poor quality and made it hard to hear everything.
- Review instructions again to begin the analysis of sessions (May12)—20 minutes
- Review the videotapes for each session (May 12)—4 ½ hours
- Write report (May 13)—5 hours
- Write addendum (May 13)—1 hour
- Complete spreadsheet (May 20)—3 hours
- Edit/proofread the report(s) (May 23)—1 hour

Total hours: 15.5

Comments on the Evaluation

This evaluation did not in any way match the way I do user testing. I had to do the review and write the report with one arm tied behind my back because I did not have information about the following, all of which are part of my process:

- Would like to see the screener and understand the criteria for the screener. There was almost no information about the users, including their level of proficiency with the web and their knowledge of U-Haul's website or other competitor's websites. Because I believe that effective criteria for screening and effective participant recruiting are the most critical foundation for a successful small study, I felt very much in the dark about who these users were and what their motivation was for completing these tasks.
- Would like to have heard the recorded comments at the end of the study. As there were no post-task or post-test assessment tools/techniques used, I have no way to triangulate the data from my observations with any other methods. I would have really liked to know if the tasks took more time or less time than expected (from the user's perspective) and whether they were satisfied, etc. etc.
- In addition, I had no contact with the users, so I was strictly an observer after the fact. I have no idea how these unmoderated sessions compared with the moderated sessions, but my process is to always conduct moderated sessions, whether in our lab or remotely.
- I had no contact with the client and no knowledge of the client's goals for the study, which is a critical component of our planning process and resulting test protocol. Even without knowing the study goals or the recruitment criteria for the study, I felt that the scenarios/tasks had some flaws in them that did not match real user experience (not knowing, for instance, at the start that the friends would want moving pads and a dolly, etc.), and ending with a somewhat biased last task that suggested to users that the experience had left a good impression. When users tried to add items to their shopping cart, they had problems, but their experience might have been different if they were selecting them from the beginning. I could, however, envision a task in which they changed their mind about an item and needed to remove it from the shopping cart later.
- I never work alone. I always have one other person from my organization and at least one person from the client organization. We plan the study together, conduct it together, analyze the findings together and set the top issues list together. I strongly believe in this process and won't take work that would require me to conduct a study without this client involvement. I also would never analyze a study alone, as I believe that more than one perspective strengthens the validity and scope of the findings (as I'm sure we will see in the workshop).
- I also don't rank top findings by severity ratings, which I have come to see as too prescriptive for the qualitative, small studies I primarily conduct. I found it problematic to use the rating codes provided because, for instance, there were no examples of problems that "caused frequent catastrophes" as "when the test participant cannot solve a reasonable problem," nor were there any instances in which a test participant was delayed "for some minutes, but eventually allowed to continue." Still there were some major findings that I addressed in the report, as well as others that the developers can and should consider changing. Organizing issues by priority, is, to me, a better way to present the findings.

Addendum to the Addendum

1. How familiar were you with the company U-Haul before we announced that we would use it for CUE-9?

Very familiar with the company.

2. How familiar were you with U-Haul's website before we announced that we would use it for CUE-9?
- Never used the website prior to this study.

3. Approximately how many times have you rented a car?

Several times.

4. Approximately how many times have you rented a car on the web?
- Never.

5. How much time did you spend analyzing U-Haul.com before you watched the first video?
- I did not analyze the U-Haul website before beginning the study. I assumed we were not supposed to do this. I take it, from the nature of this question, that we could have done this in advance, but that was not my impression from reading the instructions.

6. Approximately how many times did you pause the videos to deliberate or catch up with your notes?
- Quite a lot for the first session for two reasons: (1) I didn't know the website (see my answer to Q. 5) and (2) I had a hard time hearing the participant. After the first participant, I paused to confirm where the user had clicked or what the user had entered in the search box or in the checkout process.

7. Did you watch all or parts of the videos several times? If yes, approximately how many times did you watch each part and how long were the parts of the videos that you watched several times?

As stated in my response to Q. 6, I watched parts of the video more than once.

U1 session 30 minutes—my time 1 hour

U2 session 20 minutes—my time 25 minutes

U3 session 32 minutes—my time 1 hour

U4 session 41 minutes—my time 55 minutes

U5 session 37 minutes—my time 45 minutes

8. Were there any burning questions that you would have asked the test participants during or after the video recorded sessions if you had been moderating the sessions? If yes, what were they?

I would have asked post-task questions and post-test questions, and I would have conducted a post-test interview to get at top issues from the user's perspective as well as feedback on my observations.

I also would have conducted a pre-test interview to get additional information about users' experience renting trucks in general and using U-Haul in particular.