

Addendum Team L

1. Evaluation Criteria

I evaluated the videos by watching them, and then stopping the video when I've found an issue, idea, or positive finding. I then recorded it in a spreadsheet, noting the time of the video. When I found an issue I had already entered in the spreadsheet, I'd go back to the spreadsheet and add the new instance video time to the list.

I define a problem when the participant either does something wrong or makes a wrong assumption (knowingly or unknowingly), and/or mentions that they are confused by something. If the participant struggled with something, and then made a recommendation based on that struggle, I would usually try to record those as separate entries (the problem, and then an idea). When someone expressed liking something, or it working well, I'd define it as a positive finding. I used some discretion overall; I would not record recommendations that don't seem very plausible, or their report of problems that weren't really problems (these were rare occurrences).

After filling out the spreadsheet I went back through the findings and combined them where necessary, and categorized them to enable me to group similar topics together in the report.

2. Resources used

Preparing for evaluation – 1 hour over 1 day

- Accessing and downloading videos, reading instructions, looking at website, setting up spreadsheet

Watching videos and taking notes – 5 hours over 3 days

- Watching videos, with pausing to enter findings and video times

Analysis – 2 hours over 1 day

- Cleaning up spreadsheet, combining and categorizing findings

Writing the Report – 5 hours over 3 days

- Taking spreadsheet entries and rewording them in a powerpoint deck, adding executive summary, method, and appendices. Report did not include screenshots.

Timesheet for above activities:

Day 1: 1 hour prep, 1 hour video

Day 2: 2 hours videos

Day 3: 2 hours videos

Day 4: 2 hours analysis, 1 hour report

Day 5: 2 hours report

Day 6: 2 hours report

3. Comment on the Evaluation

I have never written a report off of unmoderated videos before. In a typical (moderated) usability test I would have a note-taker taking very detailed notes and verbatims, while I watch or conduct the test and take more general notes on my own. Then later I go back to the notes and scour through the entire document line by line until I've made sure I've covered everything in my list of findings (usually compiling an outline during that process). Finally I take that outline and write the report from it. If I need to pull video clips (which I rarely do), I go back to the notes and try to find the spots in the videos that match verbatims/notes I have (sometimes my notetaker notes times).

This evaluation process was not very similar to the way I typically carry out analysis. The need to record video times to support findings caused me to have to stop the videos and record findings and times as I was watching them. I felt that this was the easiest way to carry out the evaluation

for this CUE exercise, as I didn't want to have to go back to videos and find times to support my findings. I also needed to take note of the sequence of findings, which further incited me to record findings in this way. This is not something I typically do. Approaching analysis in this way forced me to be a bit more attentive than usual during the videos. I am usually very thorough with reporting findings, but my findings are only as thorough as my notetaker's and my notes. In the case of this CUE exercise, I had the raw footage so unless I daydreamed for a moment, I didn't miss anything.

Therefore I don't think these reports would necessarily represent the typical level of detail/thoroughness in a usability test. However, I suspect such specific instructions were given to see the consistency of findings in even a very controlled environment where everyone was forced to pay very close attention to the videos. I'm interested to see the outcome.

The other thing I will note is that I typically include screenshots in my report, which I did not this time for the sake of saving time. That would usually add another hour or two in pulling screenshots, cropping, and formatting my slides. Speaking of formatting, I'm not very happy with the visual appeal of the powerpoint deck, and in reality would have spent more time (perhaps an additional hour) making the text sizes more consistent, more space on each slide, perhaps animation to progressively disclose each bullet point (when giving the presentation), etc.

I also rarely include severity ratings, although I find them to be very helpful. In the face-paced work environment I'm in, it just adds another layer of thinking (and thus time) that I would have to wade through. In especially important tests, or those where I have more reporting time, I will include severity ratings. I typically do something similar—a High, Medium, Low or Positive finding.

4. Other

I was very surprised that the UserTesting.com participants were very articulate participants with great observations and ideas. This is not at all what I would expect of a tool like this, and I'm curious to understand more about whether these videos were chosen as the "cream of the crop" of a larger set, or that UserTesting.com itself has a very good panel of participants (or both). These five sessions were more fruitful than any random five sessions I would moderate in a lab.

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. I've rented a truck from them one or two times, and been there while friends have rented trucks as well. I also have helped someone move their things into a U-Haul storage facility.

2. How familiar were you with U-Haul's website before we announced that we would use it for CUE-9?

Not familiar at all. I don't think I'd ever been there.

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

About four.

4. Approximately how many times have you rented a car on the web?

Twice.

5. How much time did you spend analyzing [U-Haul.com](https://www.uhaul.com) before you watched the first video?

None.

6. Approximately how many times did you pause the videos to deliberate or catch up with your notes?

Too many to count. I did this constantly—this is how I recorded my findings and all of the instances of video clips. I would guestimate perhaps about 15-20 times per video, on average.

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?

No, the only time I saw a part more than once was if I was trying to find my place from where I had last left off, or was trying to replay a section to get an accurate time to record. (Or, if I zoned out for a second and wasn't paying close enough attention).

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?

Only in some cases, some follow-ups. For instance, if someone laughed or grunted or said, "huh" I would have asked them to elaborate on what they were thinking. But these moments were rare, as the participants in these videos were fairly good at articulating their thoughts.