

Addendum (Team E)

1. Evaluation Criteria

For me, a usability problem is any aspect of the design, functionality, or content of the thing that makes it difficult for people to use it for its intended purpose. This usually amounts to things that confuse the user, don't meet their expectations or needs, or simply don't work.

2. Resources used

Date	Activity	Duration (minutes)
4/12	Read instructions	30
4/21	Reread instructions	30
4/21	Watch videos, take notes	70
4/25	Watch videos, take notes	120
5/9	Ponder report contents	30
5/20	Draft report	65
5/31	Reread draft, ponder contents	75
6/2	Review videos, recheck times	55
6/2	Finish writing report	160

3. Comments on the evaluation

How realistic is this evaluation, compared to the ones I normally carry out?

I rarely conduct usability tests myself anymore, because I spend much of my time trying to show other people how to do them themselves.

I tried to approach observing these sessions in the way that I try to get other people to do it:

- **Test a small number of users.** I only observed three of the unmoderated sessions, since I believe that three tests are enough to uncover more serious issues than people have the resources to fix in a reasonable time period (at least in products that haven't already gone through a lot of testing and fixing).
- **Focus on the most serious problems first.** I reported only five problems, even though I observed many more. (One of these--E5--was what I refer to as profoundly low hanging fruit: something that while not a serious problem can be fixed by one person in less than an hour without getting approval from anyone else.)

I feel that reporting anything more than the most serious problems only leads to people fixing low hanging fruit and avoiding the problems that are harder to fix. A huge "laundry list" of problems also has a dispiriting effect ("How can we ever fix all of them?").

This is why my report only contains a small number of problems.

If I was doing this for a client, I would only report these most significant problems to them, and tell them that I would be happy to give them more problems to work on, but **ONLY** after they had fixed the most serious ones.

- **Create a very short report.** When doing client work, I actually do not generate a written report at all: I report in a conference call with screen sharing, and encourage the client to get as many people as possible to take part.

The report I created for purposes of this study is the kind of report I recommend teams create as the output of their debriefing sessions: A one-page bullet list of the most serious problems that the team has committed to fixing in a short period of time, along with a brief suggestion of the way they think they will fix each one—preferably by making the smallest change possible that will mitigate (though not necessarily eliminate) the problem for most (but not necessarily all) users.

4. Other

One thing that concerns me about this whole process is that I think that to be a good observer of the tests one has to know the thing that is being tested and its context.

For instance:

- You have to know how the entire site works. It can be difficult to interpret what you see in tests without knowing how the thing is supposed to work.
- You have to know what the entire user experience is like. (In this case, how does the online experience interact with the experience when the user goes to pick up the rental?) Things that seem like problems may not be when the whole process is considered, and things that don't seem like problems may turn out to be problems.
- You have to know what the organizational context is. For instance, what resources are available to fix usability problems? What is the internal level of commitment to fixing them? What are the technology limitations? What tradeoffs were made that resulted in the existing problems and why were they made? Without this kind of knowledge, it can be very difficult to assess just how usable the thing can actually be made to be, which in turn affects what decisions about which problems are worth reporting.

Additional Questions

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

I have probably rented vans from them three or four times in my life (to move locally, to a new apartment), but the last time was probably 25 years ago.

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

I had never seen their Web site.

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

Probably six times.

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

Probably four times.

5. How much time did you spend analyzing U-Haul.com before you watched the first video?

I *may have* tried doing the tasks before watching the videos, but I can't remember. I know I debated whether to do it or not (since I feel like you have to be familiar with the thing being tested before you can be a good observer of it), but I don't have a record of it in the log I kept. I suspect I *thought* about doing it, but didn't. If I did, I didn't spend very long on it; about 15 minutes doing the tasks.

Part of the reason that my memory is vague is that I *did* use the site after watching some of the videos to see why some things had happened. For instance, I tried adding a truck rental to the cart to see if the storage space rates would all show up as \$0.

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

When I watched the first video, I paused it very many times (perhaps 30 or more) since I was trying to capture all the data (including the times when problems occurred, and well-worded descriptions of the problems) needed for the spreadsheet.

When I was two-thirds of the way through the first video, I realized that this was incredibly time-consuming and inefficient, and—more importantly—that it was nothing like my normal process. So I started over and watched all three videos, with almost no pauses—while taking notes on paper.

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?

While I was putting the report together, I watched all the videos a second time to double-check the timestamps, to think about the issues, and to observe some things in detail that I hadn't been looking for the first time. (For instance, I was checking to see at what point each participant encountered the pre-selected quantities for some items, and whether it caused them trouble or not.)

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?

No. I actually didn't miss the ability to probe in these cases—although I have often missed it in unmoderated tests. What I did miss at some points was the ability to keep them on-task.