
Following a Fast-Moving Target: Recording User Behavior in Web Usability Testing

by Laurie Kantner

Web usability testing presents special challenges for taking notes about user behavior. Capturing user behavior accurately and completely for immediate reporting is difficult to accomplish in “real time,” with fast-clicking and complicated user interface elements to track. However, it is critical for the rapid usability feedback Web site developers demand. This article describes the challenges of capturing user Web-site behavior and the recording methods that usability specialists at Tec-Ed have developed.

Context for Web Usability Testing

The development cycle for many Web sites is more incremental than for other types of software. A new version of a Web site does not require announcing or shipping—users simply encounter the new version when they next visit the site.

Web-site owners are concerned when server logs show that pages they want users to see are receiving very few hits. Knowing how users navigate the site and where they get sidetracked is vital to paving a smoother path to the key pages.

However, the ability of an organization to make incremental improvements to Web sites makes developers unwilling to wait long for usability feedback. At many organizations, the usability specialists perform “off the cuff” expert evaluation to meet the demands for instant feedback.

Web sites designed to generate revenue generally have more structured development processes. Usability improvements to such sites have more visible impact on the bottom line—for example, increasing the number of software downloads from a site. Thus, revenue-generating sites are likelier candidates for laboratory usability testing than information-only sites.

Web Site Use is Difficult to Record

In most Web sites, as in many hypertext systems, users have enormous freedom of action. Clicks can happen quickly, and we cannot always tell at the moment if the user’s action is germane to the issues of concern. Thus, we must err on the side of recording too much, not too little. In essence, we must record every click so that we can retrace the user’s steps.

The difficulties of recording this behavior are:

- **Users of Web sites can take numerous paths** to reach their goal. It is difficult to develop a short-hand for identifying so many paths.
- **Links and buttons that have similar names but different destinations** are rampant on the Web. We must record these user choices accurately for later analysis, yet, a short-hand for identifying links and buttons is difficult with so many similar-sounding names.
- **Users can traverse many individual web pages** to reach their destinations, and recording these locations is important for determining where a problem exists. However, recording web page titles is difficult with pages that lack titles or have wordy or awkward titles.
- **Web-page users often cycle through pages repeatedly**, trying to get to their desired destinations. Recording return visits is not only important for identifying where users are getting lost or confused, but also difficult because users tend to speed up when they repeat steps they have already taken.
- **Recording detailed behavior on dynamically generated pages** is a challenge in real time, especially for unanticipated pages created “on the fly.” An accurate recording of events leading to display of that page is crucial to replicating the user’s behavior, should the observer’s notes about that page lack sufficient detail.

Whether a usability specialist takes notes on the spot or from viewing videotapes or Lotus Screencam(R) recordings later, it is physically difficult to label every choice the user makes. Tec-Ed usability specialists prefer to take detailed notes in real time to have ready data for quick analysis and reporting. Taking notes at the user’s pace, however, further aggravates the problem of describing the user’s choices precisely—distinguishing identically named buttons and links while the user has moved to another page is a problem we encounter all too frequently.

Automated Data Collection Methods Can Delay Results Reporting

Automated methods for recording user behavior are popular with today’s usability professionals. These methods include videotaping, Lotus Screencam recording, data-logging software, and server log files. All these methods reduce how much note-taking the usability specialist must perform during the session. However, they do not meet all the criteria for collecting accurate data that will be available for immediate analysis and reporting:

- **Videotaping.** The advantage of videotaping, for both Web-based and non-Web-based software, is it records fast-paced user activity that handwritten notes might miss. It also captures user commentary and cursor-pointing behavior. Thus it promotes accuracy and completeness. However, a single fixed video camera may not capture all user behavior. Most important, this method delays results reporting, because the usability specialist must review tapes to formulate a complete set of data to analyze. Depending on whether you take notes from the videotape or actually transcribe it, compiling data from videotapes can take from one to eight hours per hour of tape.

- **Screencam recording.** Like videotaping, this method records fast-paced user activity accurately and completely. While it captures cursor-pointing behavior, it does not capture user commentary. Thus it is slightly less complete than videotaping. Reviewing recordings also delays reporting of results.
- **Data-logging software.** The advantages of using data-logging software are that the note-taker can record textual notes more quickly than writing by hand, and can code observations into categories such as “Error” or “Observer Comment”. These advantages mean the data is more complete and already somewhat sortable into categories for more immediate analysis. This method still does not solve the problem of labeling user choices at the user’s pace.
- **Server log files.** This automated method is unique to Web software, and its advantage is it records a lot of detail—so much so that one might think every keystroke is captured. However, log files in fact miss important information: client-side events such as pages displayed from cache (return visits), cursor-pointing behavior, and JavaScript activity. Equally disadvantageous are the large amounts of time required to synthesize the individual records into episodes.

Until automated tools are developed that can convert videotaped sessions or electronically captured user keystroke sequences into organized tabulations of task episodes—complete with timing information and user commentary—additional note-taking will be required, and time will be needed to convert the recorded material into a database of observations for analysis. For now, we seek the simplest method that meets our requirements.

Writing by hand is not exciting in the world of technology. Yet Tec-Ed specialists have found time-savings from taking detailed notes during the actual sessions, and have developed session note-taking methods to meet the special challenges of Web usability tests.

A Note-Taking Method That Works

For lab test observation, Tec-Ed has migrated to a hybrid method of note-taking that meets our requirements for accurate, complete information, as well as ready data for immediate analysis and reporting. We have adapted this method to meet the special challenges of Web usability testing.

In usability tests where the user’s task is linear or contains few branches, we can record user behavior using a checkoff data-collection form. This type of data collection device requires us to consider the range of outcomes that may occur, based on our use of the software and the development team’s concerns, and pre-organize these outcomes into multiple-choice lists on which the test administrator or observer can check off items as they occur. These data-collection forms also provide white space for jotting down user comments or “custom” observations.

The success of this data collection method depends on five factors:

- The usability test is of a defined set of web pages; that is, tasks do not include learning the site for unknown destinations.
- A usability colleague reviews the data-collection form to identify additional choices or improvement in presentation of choices for fast recording.
- The dry-run of the usability test includes a dry-run of the data-collection form.
- The usability specialist practices using the data-collection form so that finding and checking off items does not mean missing user actions. The usability specialist must become skilled at handling these forms unobtrusively. We use this method both in the room with the user and in a separate observation room. Fortunately, our experience is that the user's attention remains focused on the screen, not on our note-taking.
- The testing or observation room has adequate surface area for handling the paper.
- The usability specialist is skilled in writing quickly, because ultimately the data-collection form cannot anticipate every possible behavior the user's chosen chronology.

Of course, the data-collection form also include prompts for Start Time and End Time, and the usability specialist is practiced in filling these in diligently.

Examples of Data-collection forms for Software Testing

Here is an example of part of a filled-in user data-collection form for software in which the user's task is extremely linear:

Start Time: <u>10:43</u>	
Correct Procedure	Procedure Followed by Participant
1. Select document for which to set access rights.	<u>Opts SysAdmin V/G</u>
2. Select View/Access List.	<u>USER ACRTS</u>
3. Choose Users tab.	<u>Unit</u>
4. Set access level for user.	<u>"I want to see the status"</u>

In the above example, the usability specialist would either place a checkmark next to the steps completed or write additional steps on the right side, drawing an arrow to the left side where they occurred within the sequence of correct steps.

Here is an example of part of a data-collection form for software in which the user can make more branching choices:

How does the participant display the format?	
<input type="checkbox"/>	Select the document icon, then the display choice
<input type="checkbox"/>	Double-click on the document icon
<input checked="" type="checkbox"/>	Drag-and-drop the document icon to the display choice
<input type="checkbox"/>	Right-button menu, Display Now command
<input type="checkbox"/>	Press Enter
<input type="checkbox"/>	Window menu Display command
<input type="checkbox"/>	Icon at bottom of another format display
<input type="checkbox"/>	Administrator prompts participant to try another method to display the format.

In this example, the usability specialist marks the appropriate checkbox. For lists where users might perform multiple items, the usability specialist numbers the items in chronological order. The space at the right is available for recording additional information such as user comments.

The advantage of the checkoff data-collection form is that it enables the usability specialist to capture predictable actions quickly and fairly accurately. Most important, after the session, the results are easy to tally for analysis and reporting. The disadvantage of this form is that users can make surprising choices, and then the usability specialist has to annotate the form. However, those instances themselves are extremely informative in telling us where everyone's expectations were incorrect.

As mentioned earlier, we add the space to check-off forms for additional observations or capturing user comments, with labels identifying the intended information. For example:

<input checked="" type="checkbox"/>	Administrator assistance needed to add a field to a profile card:
<i>pointed to Fields choice list</i>	
Participant comments about terminology or process:	
<i>"Fields" ok for doc characteristics</i>	

Another disadvantage of this form style is that these forms can occupy lots of pages; skill and practice are required in handling these with ease.

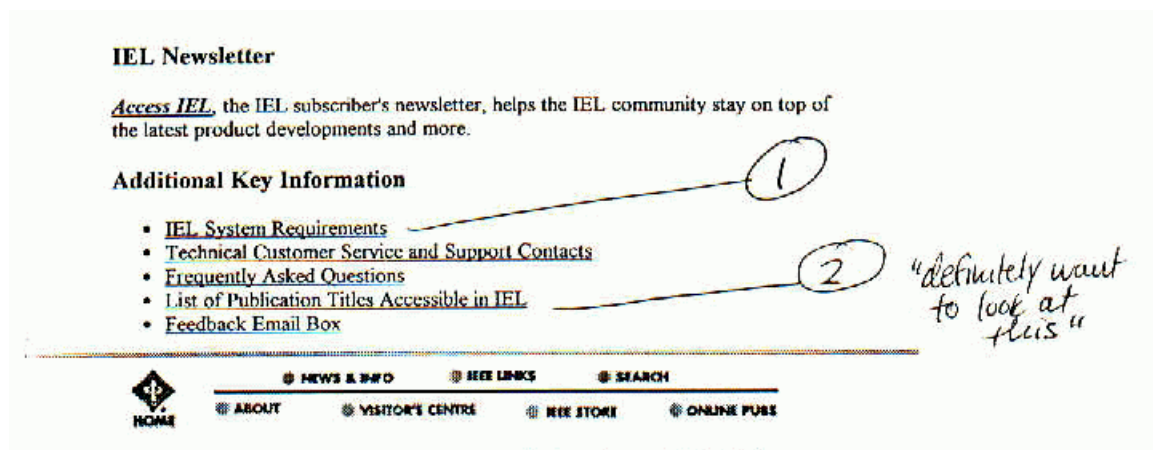
Some usability specialists prefer to use blank paper and take "stream of consciousness" notes, or to use a computer and simply type into a blank document. These methods can generate complete notes, but usually offer no pre-organization of information from which to tally quickly.

Adapting the Data-collection form for Web Studies

Despite our efforts to create a usable checkoff data-collection form, it has proven insufficient for note-taking of participant behavior during Web usability tests. Why?

- To provide a list containing every possible choice a user can make on a Web page would require absurdly long lists of items—and the observer couldn't possibly find the desired items quickly enough. Our data-collection forms already occupy 10 to 15 pages for a 2-hour session; we believe that is enough pages to master.
- Web pages continue changing right up to the test day. However carefully we label the Web-page choices we want to track, some of those choices will change or even go away on testing day.

Therefore, we replaced the detailed lists of checkbox items with printouts of the Web pages on which we record user actions. These printouts can be generated at the last minute to reflect the latest updates. We photo-reduce the web-page printouts to create a column of blank space at the right. For the first visit, we place a "1" next to the first link chosen, and then put the number "2" on the page this link goes to. For a repeat visit, we write user actions in the blank space, pointing to links for clarity. We also use the blank space to record user comments. See the example below.



To facilitate page-flipping, the usability specialist is wise to use sticky notes on the edges of key pages. All possible pages of interest must be included in the packet.

In addition to web-page printouts, we pre-organize some of the expected outcomes into higher-level data-collection forms with spaces for tabulations, to record more summary-level observations. After a user completes a task or subtask, we use the data-collection forms to summarize what happened, for quick post-test reporting.

Here is an example of pre-organizing outcomes on a data-collection form filled out on completion of a session:

Did the participant recognize "[Product Name]" as what s/he is looking for?	<input checked="" type="checkbox"/> Yes	<input type="checkbox"/> No
Did the participant want to read [Product] background information on-line?	<input checked="" type="checkbox"/> Yes	<input type="checkbox"/> No
Did the participant read the URLs?	<input type="checkbox"/> Yes	<input checked="" type="checkbox"/> No
Did the participant ever see the [Product] home page?	<input type="checkbox"/> Yes	<input checked="" type="checkbox"/> No
Recognize it?	<input type="checkbox"/> Yes	<input type="checkbox"/> No

In the above case, the data-collection form was for a study in which the user was to find and download a type of software without knowing its exact name. The user could potentially visit up to 70 web pages to complete the test tasks. Our goal was to record all user choices, while also making sure we captured the answers to the above questions which addressed key concerns.

As test administrator in the same room with users, I kept one stack of web-page printouts and one set of checkoff data-collection forms, sitting at a working surface that enabled me to alternate between the two sets as needed. The user was focused on the screen and rarely looked at me or my recording. If I got behind, I caught up during the time between tasks, when the participant was filling in a questionnaire.

Analyzing the Notes after the Session

At the end of the first session, the usability specialist can begin designing intermediate tabulation forms for collecting all of the data in one place for analysis. The check-off forms simplify choice and timing tabulations. The most time-consuming part is going through the annotated web-page printouts and recording events considered significant. However, using annotated printouts to accomplish this task is far less time-consuming than reviewing videotapes.

Here is an example of an intermediate tabulation of notes taken for a web page study, using the method described above. (Note that we occasionally also spot-viewed videotapes to confirm or fill in our notes; this spot-checking took less time than reviewing the tapes to create notes.)

Participant	First Link Chosen from Home Page	Number of Pages Visited	Recognizes "[Product Name]"?
Pilot	LinkA	10	Already knew its name
1	LinkB	8	Figured it out
2	LinkA	8 (hint)	Figured it out (hint)
3	LinkA	12 (hint)	Figured it out (2 hints)
4	LinkA	11 (gave up)	Had to be told directly
5	LinkC	10 (hint)	Figured it out (hint)

Participant	First Link Chosen from Home Page	Number of Pages Visited	Recognizes “[Product Name]”?
6	LinkD	3	Figured it out
7	Search	9 (hint)	Had to be told directly
8	LinkA	13+	Used search
9	LinkE	10 (hint)	Figured it out (3 hints)
10	LinkB	9 (hint)	Figured it out (hint)
11	LinkA	7	Figured it out
12	LinkA	8	Already knew its name

In this example, we see that LinkA was a popular but less productive choice than LinkD, and that many users required hints to succeed. The data for the First Link Chosen from Home Page column and the Number of Pages Visited column came from analysis of web-page-printout annotations, and the data from the Recognizes [Product Name] column came from the checkoff data-collection forms.

Conclusion

Web-site usability testing requires fast, accurate data collection in an accessible format if the usability team wants to deliver immediate results—which Web site developers demand. Using a combination of Web page printouts for detailed data collection and checkoff data-collection forms for pre-organizing higher-level outcomes gives the usability specialist the range of instruments needed to give results fast.

The Web page printouts enable the usability specialist to track detailed Web usage data in “real time” without the added difficulty of labeling the user’s choices precisely. The data-collection forms answer higher-level concerns, summarizing the bottom line of how the user worked with the site.

This method of collecting data about Web usage works for all sizes of Web sites, although extremely large sites increase the difficulty of making sure all pages the user might visit are represented in the data-collection forms. This method also accommodates those last-minute changes to Web sites we have learned to expect. We still ask the Web site developers to keep the site unchanged during the actual days we conduct the usability testing. And sometimes it happens!