
**Information
Architecture
for the Web**

Module 7:

Conduct Usability Testing



Info.Design
1725 Q Street NW, #201
Washington, DC
20009-2498

www.infodotdesign.com

Module 7: Conduct Usability Testing

Successful design comes from a basis in direct observations of, not assumptions about, the users or potential users of your product.

*JoAnn Hackos and Ginny Redish, **User and Task Analysis for Interface Design**, 1998*

Objectives:

- 7.1** → **Explore reasons for user testing**
- 7.2** → **Identify techniques for user testing**
- 7.3** → **Conduct user testing throughout**
- 7.4** → **Identify other ways to gather information from users**
- 7.5** → **Analyze results of user testing and revise text**


Introduction

We'll explore the importance of testing and discover ways we can test.

As developers of Web content, we sometimes feel we don't have enough time for synthesizing content. Certainly we often feel we don't have time for testing. Yet by testing our site architecture, labels, text, and links, we have the opportunity to learn about how our users think, and can then build documents that help them do so.

In this module, we will explore the importance of testing and discover ways we can test categories, labels, text, and site structure. We'll explore methods of determining when we should test and how we should analyze results.

Testing does not have to be expensive, nor does it need to be complicated. It does, however, need to take place. By testing your site you will discover questions you've never considered. You'll find nuances that can help your users. You'll find a route to user-focused structure.

7.1 

Explore reasons for user testing

Usability rules the web...[It] has assumed a much greater importance in the Internet economy than it has in the past.

**Jakob Nielsen,
Designing Web Usability,
2000**

Users face a great deal of frustration. They still feel guilty when they make mistakes and often feel that they should somehow be able to figure out what to do. The fact that IDG's *For Dummies* series generated \$121 million in revenue in 2000 suggests something about how users view themselves, as well as their hunger to master technology.¹

Some scholars define usability as the degree to which a given piece of software assists the person sitting at the keyboard to accomplish a task, as opposed to becoming an additional impediment to such accomplishment.²

Usability testing routinely reveals important problems that document designers, even expert ones, may fail to detect. Involving the audience as active participants in document design provides substantial benefits by improving the documents that people use and by broadening the scope of problems that document designers can attend to during revision.³

As information architects, we must attend to how our users relate to categories, labels, and text.



Question

What role does user testing play in your organization?

Does your organization consider good usability to be a business goal?

What do we know about user testing from our exploration of users and their tasks?

¹Charles B. Kreitzberg and Ben Shneiderman. "Making Computer and Internet Usability a Priority." *Essays on Usability*. 2001.

²Michael D. Levi and Frederick G. Conrad, *Usability Testing of World Wide Web Sites*, 1998.

³Karen A. Schriver, *Dynamics of Document Design*, 1994.

Test categories and labels with users

Your goal is to present information the user expects—and wants to use. You have many opportunities to test your text so it works for your users. You can use one of these four methods to test categories and labels with users:

- Ask users to create labels.
- Ask users to respond to labels.
- Ask users to create categories and labels.
- Ask users to perform tasks with categories and labels.

Method 1: Ask users to create labels

Your challenge when testing is to gather information from your users' heads. Use the following procedure to test labels with real users:

1. Show the user groups of content. Do not give or suggest labels for the groups.
2. Ask the user to describe what is common among all items in the group.
3. Ask the user to think aloud when trying to figure out the grouping.
4. Record comments.
5. Ask the user to give each group a name (label).
6. Record labels.

Method 2: Ask users to respond to labels

Similarly, you want to gauge users' responses when they see labels. Use the following procedure to test labels with real users:

1. List all possible labels for the categories you want to test.
2. Show each possible label—one-by-one—to users.
3. Ask users to describe the kind of information he or she would expect to find behind each label.

Try this strategy: Collect results in a spreadsheet. Score user interest on a scale of 0 to 3 (not interested to very interested). Use similar labels to identify those that hold the most resonance for your audiences.

TASKS	Dog Owners	Cat Owners	Gift Givers	Fish Owners	Bird Owners	Pet Seekers	TOTAL
Buy gift certificate	1	0	3	0	1	0	0.6
Order pet food	3	3	1	3	3	0	2.1
Cancel an order	2	3	2	3	3	0	2.1
See order status	3	2	1	2	3	0	1.8
Return product	2	3	0	2	1	0	1.3
Compare food prices	3	3	1	3	3	2	2.5
Compare food brands	3	3	2	1	3	2	2.3
Find the right pet	0	0	1	0	0	3	0.6

Example of spreadsheet showing scale from 0-3. Locate a template on the class website, <http://www.infodotdesign.com/class/ia>.



Workworld Consideration

A story: We tested labels for a health care site. The marketers had identified a label they preferred: Follow Healing Path. But when we tested this label, no one wanted to click on it. They didn't know what it meant. On the other hand, they were very interested in Following an Individualized Health Care Plan.

Method 3: Ask users to create categories and labels

Your goal when testing categories and labels is to ensure that your users create the same meaning as you intended. One way you can discover users' reactions to information is to discover their relationships.

Use the following procedure to test categories and labels with real users:

1. Write each content item from the content inventory and content wish lists on a note card, piece of paper, or sticky note.
2. Ask users to place content items together into groups that make sense.
3. Ask the user to think aloud while doing this.
4. Make notes about which content items are most difficult to categorize.
5. When all content items are grouped, ask the user to name the group and write a brief description of the kind of content that the group contains.
6. Record the groupings and labels.


Labeling is a form of representation. Just as we use spoken words to represent thoughts, we use labels to represent larger chunks of information in our Web sites...Labeling requires planning to succeed.

Louis Rosenfeld and Peter Morville, *Information Architecture for the World Wide Web*, 1998

Method 4: Ask users to perform tasks with categories and labels

Use the following procedure to test labels with real users; it requires two observers:

1. List the labels that users would see on the entry page to your site on a sheet of paper. Repeat this for each top level page of your site. Decide how far down you want to test.
2. Ask a user why he or she would come to your site.
3. List all the goals and tasks he or she mentions.
4. Explain that you'll display words on pieces of paper to mimic the words that would be seen on a screen at your site.
5. Ask him or her to accomplish a goal or task selected from the list in Step 3 (above) by following the links that he or she thinks will provide the information needed. You'll be the website by changing pages according to those choices.
6. Show him or her the paper with the front page labels. You will be the website by changing pages according to those choices. Following the picks, show him or her the next screen.
7. Ask the user to think aloud while doing this. The second observer records choices, frustrations (verbal and nonverbal), where he or she gets lost, finds answers, gives up, and other reactions.
8. Repeat Steps 5 through 7 (above) for each task or goal on the list.

7.3 

Conduct user testing throughout

Testing reminds you that not everyone thinks the way you do, knows what you know, uses the Web the way you do.

Steve Krug, *Don't Make Me Think*, 2000

Identify reasons for user testing throughout

The goal of usability testing is simple—find out how people use a website or application and make it better. We test users because it improves the sites people use.

Research (such as that conducted by Karen A. Schriver and described in *Dynamics in Document Design*) shows that a reader-focused evaluation can produce a substantial improvement in document quality. Testing often offers results that are different than what any expert would predict.

Web-specific research shows a strong connection between user success and business success. Forrester Research shows that redesigns can boost revenue, lower support costs, reduce development waste, and improve customer satisfaction.

Testing can be done at any phase in the development life cycle. But as a general rule, the earlier you test, the greater your cost savings. As you test, you can learn about a user's ability to get a job done and the user's satisfaction with the process.

Jakob Nielsen (www.useit.com) recommends two techniques for user testing:

- Guerilla user testing
- One-hour user testing

These techniques are described in the following sections.

Follow a four-step process for Guerilla user testing

1. **Select truly representative users**
2. **Prepare the users for testing**
3. **Listen to the users' responses**
4. **Evaluate the users' findings**

1 Select truly representative users

- You don't need many testers
- One or two people from each user group is sufficient for guerilla testing usability
- Testers will certainly point out problems you will need to notice
- Beware the disease of familiarity
 - Don't select colleagues who are overly familiar with internal thinking rather than external thinking.
 - By this point, you are tainted, too. Even if you are an outsider to begin with, you now know too much.
- Avoid your boss' or client representative's opinions
 - This may be your big political fight. Warn your client and your boss ahead of time that you need to gather user feedback...and it might just vary from their ideas.
 - Explain to them how they are not typical of your representative user.

Familiarity breeds confusion. Those suffering from the disease of familiarity are the experts in the world who, so bogged down by their own knowledge, regularly miss key points as they try to explain what they know.

**Richard Saul Wurman,
Information Anxiety 2,
2000**

2 Prepare users for testing

- Complete the logistics
 - Be sure your users know what they will be doing, when, where, how.
 - Decide if you want to record users as they evaluate the site. Get permission from the user if you choose to do so.
- Plan your schedule
 - It takes 39 hours to test a website for usability the first time you do it.
 - The time estimate includes planning the test, defining test tasks, recruiting test users, conducting a test with five users, analyzing the results, and writing the report. With experience, Web user tests can be completed in two work days.

Jakob Nielsen, **Alertbox**, May 3, 1998

- Prepare the materials (tasks for testing)
 - List sample goals and tasks based on previous contacts with users (information you gathered during the analysis phase).
 - As you create tasks, remember to use strong verbs. For example:
 - Find out the membership dues for this organization.
 - Buy the latest report by J. Smith on global warming.
 - Find out if your company will allow you to load software from home on your computer.
 - Find out how your congressperson voted on yesterday's bill.
- Prepare the tester (before testing and on test-day)
 - Treat users as partners.
 - Work with users one-on-one.
 - Offer a framework for the information you are looking for:
 - Tell users you will have them complete **tasks**.
 - Tell users you are seeking a *self-focused diagnosis*. Ask the reviewer to walk through each task, constantly referring to "I" and talking out loud throughout the process.
 - Invite users to comment on preference if they want, but remind them *your* focus is on performance.

3 Listen to users' responses

Be quiet. Observe. Listen. Record.

- Be quiet
 - **WARNING:** If you had anything to do with developing the site, you'll be tempted to try to help users, by asking leading questions, answering their questions, or guiding them. **DON'T.** This defeats your purpose.
 - Prompt users only if necessary. If the user is not talking, ask: "What are you thinking?"
- Observe
 - Watch users' eyes, face, posture, and gestures.
- Listen
 - Listen carefully for their **EXACT** words. You may want to use a tape recorder or video camera.
- Record
 - Try to capture exact quotes.
 - Identify snippets of frustration, confusion, delight, etc.

4 Evaluate users' findings

- Ask yourself these questions:
 - Which major user goals are being impeded by the site?
 - Where did the users get lost most often?
- Identify the problems:
 - **Problems of commission**—created by what is actually on the page (how text is written or buried inside chunks).
 - **Problems of omission**—caused by what is **NOT** on the page. These problems are caused by missing content, underdeveloped ideas, gaps in logic, incomplete procedures, nonexistent transactions, or missing graphics.⁴

The important things that you learn from usability testing usually just make sense. They tend to be obvious to anyone who watches the sessions.

Steve Krug, *Don't Make Me Think*, 2000

⁴ Karen A. Schriver, *Dynamics of Document Design*, 1994.

- **Problems at a global level**—users face organizational problems. They didn't understand the hierarchy, it doesn't work for them, they can't find their way around.
 - **Problems at a local level**—users question some of the details, such as word choices or misleading images.
- Use results to make the case for change:
- Ask yourself, “What have we learned?” about these users that can guide us when planning and revising this document and other documents.
 - Use your user testing information, not as an end in itself, but to show the importance of including users in site building.

7.4 

Identify other ways to gather information from users

Conduct one-hour testing

One hour is better than nothing

Jakob Nielsen writes: “It’s a misconception that it has to be expensive or difficult to conduct a usability test of a website. In fact, you can collect a small amount of user data in as little as an hour.”

Jakob’s one-hour usability test

- Catch a user or two on their way to the cafeteria (for an intranet), at a tradeshow, or any other place where a reasonable representation of users gather
- Spend 10 to 15 minutes with each user, giving them your laptop and asking them to perform a single task on your site
- Write a 50-line email message listing the usability catastrophes you observed
- Mail this mini-report to your entire Web team from your hotel room that evening

Gather information via online surveys

Many organizations like the flexibility of online surveys. A Web survey can provide information on patterns of use, opinions, needs, and desires. Certainly, surveys offer value because they solicit information from users—and any information is better than no information.

Surveys are not, however, as simple as putting a few questions and a 1-5 scale online. You face challenges with results, challenges with tabulation, and challenges with length and content.

Challenges with survey results

You want your results to be accurate. Researchers look for two measures: reliability and validity. Reliability is a measure of consistency. Researchers wonder how well the same scale will give the same rating in the same conditions. Unfortunately, humans produce varying output in response to the same repeated input. So in a “reliable” survey, questions are asked in different ways while remaining the same question. Validity is a measure of accuracy. A survey must be reliable to be valid.

Challenges with tabulation

If you want to build a survey that actually tabulates in real time and provides varying ways to analyze data and sample respondents, you're looking at a programming challenge. Some survey building tools are available online, as are other tools for gathering and responding to customer feedback. Typically, these survey tools are directed at large companies who might have money to spend for the service.

Challenges with survey length and requirements

Web users will not complete lengthy surveys. The typical pop-up survey violates many customer requirements in terms of soliciting feedback. For example, many pop-up surveys ask for personal information. Research suggests that users want to provide feedback anonymously. Pop-up surveys also ask what the website researcher wants to ask, which doesn't necessarily equate to information the user wishes to provide. These surveys often ask questions about preference, rather than performance.



Tips

Maintain caution with surveys

Follow these guidelines when gathering data with online surveys:

- ✓ Ask what topic the user was looking for and whether or not they were able to find the desired information.
- ✓ Include a text box that allows users to fill in whatever questions or comments they have. It allows users to describe any problems they encountered.
- ✓ Keep the survey short.
- ✓ Offer your users something in return for filling out the survey form.
- ✓ Make the survey in several stages, each stage offering some “goodie” in return.
- ✓ Include a small survey question in a side bar.
- ✓ Demonstrate to the users that their input matters by showing how it actually provokes change.
- ✓ Quantify this system by measuring the number of different users who provide feedback and the number of actual changes that are made to your site as a result.

7.5 

Analyze results of user testing and revise text

Designing and redesigning is a never-ending process. There's always room for improvement—even in utopia.

Tucker Viemeister (Executive Vice President for R&D, Razorfish), *Wired*, 2001

Analyze results

Analyze results of user testing to decide which categories and labels work most effectively with representative users.

At the end of [a] session a full group discussion is held to gauge the participants' subjective reactions and solicit suggestions for improvement. Our experience is that the conversations between participants during such debriefings are frequently more instructive than the direct participant-tester dialogue.

Michael D. Levi and Frederick G. Conrad, *Usability Testing of World Wide Web Sites*, 1998

Revise based on results

Based on the feedback from users, revise your categories and labels to meet user expectations.

The cleverest system in the world does no good if users avoid it because they find it annoying.

Michael D. Levi and Frederick G. Conrad, *Usability Testing of World Wide Web Sites*, 1998

