FAQ: How to conduct Usability Testing
FAQ: How to conduct Usability Testing

User Experience Direct (UX Direct) is an Oracle Applications User Experience (UX) program that provides user experience expertise to Oracle customers and partners for their implementations, customizations, and usage of Oracle enterprise applications. The goal of this program is to enhance end user experiences during and after customer implementations and improve user adoption of Oracle’s enterprise applications.

Overview

Usability testing is a research method for answering the question: Is this product easy to learn and use? The “product” may be a whole application, part of one, or part of an application that has been customized. In addition to the product being tested, you need typical users of the product and typical tasks for them to do.

The foundation of the method is to ask users to perform the tasks while you watch and take notes about what you see. The problems that users have during the course of the usability test become very obvious with this method. They hesitate, go down incorrect paths, fail to complete tasks and express frustration at the experience. Finding a way to fix the cause of those problems is an additional challenge.

The remaining sections of this document contain a series of questions and answers about how to conduct a usability test.

1. Is this a new method?

No. Usability testing has been used to evaluate the usability of products for about 30 years. It is a standard technique that is used primarily by user experience (UX) professionals. All of the large software development companies use this method.

2. What are the advantages of usability testing?

Usability testing uncovers the most serious problems that provide barriers to users who are performing tasks, and it uncovers the problems quickly. A user experience professional can usually define the problem after watching no more than three to five users.

Usability testing also has the advantage of built-in credibility, that is, when skeptics see users failing at tasks, they do not have doubts that there is a problem to be addressed. And the problem can’t be dismissed as “just users griping.” For example, when a company customizes a software application and some users complain that they can’t or won’t use it, their concerns can be ignored if those concerns are not supported by data from a usability test.

3. What are the disadvantages?

Usability testing requires some planning on the testers’ part. Tasks need to be identified that probe the user interface of the product, and typical users need to be recruited and agree to attempt the tasks. It takes a few days to plan and conduct the test.

4. How many users does it take for a valid test?

One of the values of this method is that a small number of users can find the most serious usability problems quickly. If you have the luxury of having many users to choose from, about five users is enough; many tests have been conducted with three users. There is ample research showing that these small samples do uncover the usability problems.

5. Who are these “users?”

The users who participate must be people who will use the product when it is finished or released. For example, you wouldn’t expect an IT professional or an administrative assistant to understand the tasks that a compensation specialist does. People who are not compensation specialists will miss problems that a compensation specialist would find right away.

6. How do I find the users?

One way to find and recruit the right people is to have a user profile of who the users are. Use it to create questions for candidates to see if they qualify. For example, you might go to the manager of human resources (HR) and ask if there are any compensation specialists. If some are identified, you might ask them if they have at least one year of experience in their job and whether they have used the company’s current software application for at least six months.

---


2 See the UX Direct: User Profile document.
7. How do I choose the tasks?

This is where a user profile would also come in handy, because it lists the most important tasks users say they do. If you don’t have a profile, there are other things you can do. You can go the manager of the users and ask what tasks the users perform frequently. If you have a sample of users that is large enough, you can ask one of the people who qualifies to work with you to create realistic tasks. You want tasks that are typical rather than exceptions. As a rule of thumb, these tasks should be the kinds of tasks that users perform 80 percent of the time. When you are testing a configuration that has been tailored by the organization, you also want to include tasks that probe those functions or pages. For example, if the design has options in it that are intended to make it easy for a compensation specialist to compute workers compensation rates, you will want a task that includes having to use the screens that provide access to those calculations.

8. How do I present the tasks to the users?

It takes a bit of care to make sure that the wording of the tasks you have chosen is clear to the users and does not suggest how to perform it. You will present the task in the form of a scenario. The task scenario describes the objective of the task and any information the user needs to accomplish it. For example, a task scenario for a compensation specialist might read like this:

“The new VP for product development wants to know the salary of one of his staff, Ms. Rebecca Ellison.”

Notice that there is no hint about how to perform the task, such as “On the Compensation tab…” It does contain a specific name to look for, which is more realistic than saying something like, “Find the salary of everyone on the new VP of product development’s staff.”

It is a good idea to do a quick rehearsal of the wording of your scenarios by asking someone to read them and then perform the tasks.

It’s best to write each scenario on a card or piece of paper and ask the user to read it aloud before attempting it. If you read it for them, the user may not hear all of it or may later want you to spell the name of the person whose salary is of concern. By having the user read the scenario, you know that he or she has seen every word. Ask the user if the task is clear after he or she has read it. If there is more than one task, make sure the user can’t see the second one while working on the first one.

An important aid to finding problems is asking users to talk out loud as they work. Most people find this easy to do after a few minutes. The talking provides the tester with information about the strategies users try.

9. Do I need a laboratory to conduct a test?

No. A usability test can be conducted in a cubicle or in a typical conference room. All that is needed is some way to show the user interface to users. If the user interface is an interactive application, a computer is the only equipment necessary.

If you want to record the session, set up a camcorder that points at the screen, which will then show what the user sees and record what is said for playback.

10. What if the users work in different locations?

Remote users are not a barrier to conducting usability test sessions. Most companies have phone conferencing systems and Web conferencing capabilities. The Web conferencing system allows testers to share their desktop with users and to give them control of the cursor, and the phone conferencing allows the two people to talk.

11. What does the moderator do during the task?

Moderating a test is a skill that can be learned easily. If there is an experienced UX professional available, he or she can help. But in many cases, there is no UX professional. The best way to handle moderating when you are new at it is simply to watch and listen. Say as little as possible while the user is working on the task. Even if the user is struggling and you know how to do the task, let the user struggle for awhile, just like they might if you were not there. Avoid approaching the session as a demonstration. If you find yourself saying, “I can show you how to do that,” you are intervening in a way that makes it harder to document what problem the user is having. Keep in mind that the objective is to find as many problems as you can in the short time available.

12. How do I identify usability problems?

One reason usability testing is so popular is that severe usability problems, the “show stoppers,” reveal themselves quickly and are quite obvious. Basically, they present

---

4 See Dumas & Loring, op. cit.
themselves as a barrier or hindrance to users completing a task correctly and efficiently, and/or show that users are not satisfied with the experience.

As you watch users perform tasks, look for indicators of problems: hesitation, going down the wrong path toward a solution, failure to complete a task, asking for assistance, not understanding what words in the interface mean, and expressions of frustration. If you can videotape the session, you can go back and record when and how often these events occur. But often, it’s more efficient to just take notes as the user works. It helps to have a form with the task scenario at the top and some checkable boxes to record errors and failures. There should also be space to write down what you see as the problem.

13. Do I treat all usability problems the same way?

No. Some problems are more important than others. There are two ways in which problems differ. First is the impact of the problem on users. For example, does it cause users to be unable to complete a task, or does the task take much longer to complete than expected? On the other end of the spectrum, the problem may be a minor misunderstanding that only shows up the first time the product is used. Also, some problems affect every user, while others affect only a small subset.

Second is the scope of the problem. Does it affect only one screen or does it have a wider impact? For example, structural issues, such as the organization of tabs or menus that do not match the users’ work flow, may affect more than one task. A single misunderstood term, however, is likely to affect only a step within a task.

A simple but effective way to categorize the severity of problems is to use a three-level scheme: high, medium, and low severity. A person who has analyzed the data can make this judgment, and it will not be far off from those of a UX expert.

14. How do I organize the test data?

What you do with the data depends on whether you need to justify the existence of the problems in order to get them fixed. For example, you might have to present the list of problems to an implementation committee or to a senior manager to get approval to fix them. In that case, you need to create a list of the problems, preferably in their order of severity, and build a case for each with your data. For example, you might say: “None of the users were able to edit their expense report because they could not find it once it had been created. In addition, they spent wasted minutes exploring the interface to try to find where the report was, and several of them expressed their frustration at this experience. The problem is that once the report is completed and saved, it becomes a ‘memorized’ report, which is a term that users did not associate with a completed report.”

Notice that the problem is justified with performance measures so that it cannot be dismissed as just griping. You can put these statements about the problems into a PowerPoint presentation or into an e-mail. If you have any videotape showing the problem, it will highlight the problem and make it easier to capture the attention of decision-makers.

If you do not have to justify the problems to anyone else but yourself, then you need do no more that make a list and start with the most serious problems.

15. How do the problems get fixed?

The solutions to usability problems need to be carefully designed. A simple solution arrived at quickly may not solve the problem or may make it worse. Typically, the team of people who are creating the product or customizing it get together to work on a solution. Sometimes the severe problems require a major design change; other times, a fix to one page or an added dialog box or tool tip can solve the problem.

To be sure that the problem is fixed, someone needs to go back to the users and try the new design with them. Ideally, the new user will be someone who did not participate in the original test, but in the case of a customization used by a small number of users, that may not be possible.

16. Is testing the only way to evaluate a product?

No. A method called heuristic evaluation can also be used. Heuristic evaluation is the process of systematic inspection of the user interface to uncover usability issues. The issues are grouped into 10 categories, which are the heuristics. The end result is a list of usability issues to be addressed.

The method is most effective when a UX expert does the inspection. Research has shown that UX experts consistently find more issues than people who are not trained in UX. But if no UX expert is available, someone on the development or implementation team can do the review. It is even better if several people do reviews independently of each other and then combine their results.

---

5 See the UX Direct: Heuristic Evaluation How-To document.