Planning Usability Tests For Maximum Impact

Scott McDaniel, Laura Snyder

Usability tests make products better. Those of us who have seen their results understand their value, but we sometimes have difficulty convincing managers or developers of their worth.

The pay-off of a usability test depends on its context. Both the type of test you perform and the timing of that test influence what you can do with the results. At best, you can guide the design of a product or clearly demonstrate its usability. At worst, you can ruin your credibility and have no effect on the product whatsoever.

This paper explores what you get from a usability test at each stage of the user interface design process. It also provides advice on selling testings to management.

INTRODUCTION

Nobody starts out saying, "I want to create a product that's hard to use!" Yet many managers and developers ignore a basic tool to enhance their product's usability – the usability test. Either they haven't heard of them, or they don't know how to use them properly.

To execute a proper usability test, you must know what kinds of tests there are as well as when to do them. Usability testing too late in the process, for example, can ruin your credibility – decreasing your odds of being able to regularly use usability methods. Similarly, measuring task completion time when you are still working out the overall navigation and user interface concept makes little sense.

First, we discuss the types of usability tests available, spelling out the pros and cons of each. Then, we use a common framework for user centered design to discuss what you do and don't gain by testing at each stage of the design cycle.

USABILITY EVALUATION

At its most basic level, a usability test is the act of asking your user to perform a task using a product. Unlike a demo, you do not provide an introduction to the user interface or any help along the way. The goal is to measure the product's inherent usability or to guide the design of the product.

This paper does not provide details on how to conduct a usability test; instead, it concentrates on what you can get out of them. For a thorough grounding in how to conduct usability tests, see Dumas & Redish's A Practical Guide To Usability Testing and Jeff Rubin's Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests.

There are two types of usability test: summative and formative.

Summative Usability Testing

Summative usability tests are concerned with metrics. They measure a product's usability and are useful for demonstrating that you have, in fact, made the new version of a product more usable than a previous version. Similarly, you can use them to compare the usability of competing products.

Typical measurements for a summative usability test are the timeit takes to complete a task, the number of errors users make, or the number of clicks required to complete a task.

You cannot usually address the problems that summative tests reveal right away. Rather, they serve as input to future releases, not immediate projects.

Summative tests often work best in a formal, lab setting. Metrics call for added rigor, and it is often useful to have video tape of the sessions to aid in counting errors, timing tasks, and presenting results.

Formative Usability Testing

Formative usability tests are concerned with user interface design. When you need to find out if your screen navigation is usable, a formative usability test will provide you with the information you need to make decisions. You can target formative tests to reveal high-level issues, like whether users understand navigation and metaphors, or you can target them to detailed tasks, like whether users have trouble with specific screens.

Formative tests typically use a "think-aloud" strategy, in which facilitators invite users to keep talking throughout the tasks they perform. The facilitator probes for more information frequently. All this discussion is useful for making design decisions, but would get in the way of summative metrics like time to complete a task.

Formative tests work well as either informal or formal tests. It is quite possible to get good, actionable results from a simple paper prototype, with just a facilitator, observer, and a couple of hours of planning. Formal

tests provide greater depth in your findings and can help justify design choices to skeptical third parties.

WHEN TO EVALUATE USABILITY

The timing of a usability test is very important. A common mistake among new usability professionals is to insist upon a usability test right away, even if the product is just a couple of weeks away from release.

By understanding the software development life-cycle, usability advocates can recommend only the tests that will have a maximum impact on the final product. It may take foregoing a usability test on one release so that you can plan it properly into the next release's activities.

The LUCID Framework™ (Logical User-Centered Interaction Design) can help you recognize where your project is and recommend the appropriate usability test. Developed primarily by Dr. Charles Kreitzberg and Whitney Quesenbery, LUCID is one of the most widely-used frameworks for designing interactive products and user experience
It consists of six stages, and usability tests performed during each stage yield different results:

- Envision
- Discovery
- Design Foundation
- Design Detail
- Build
- Release

Envision

The central event of the Envision stage is an Envision Session, in which the entire team and all stakeholders come together to establish the vision and direction for a project. This is the time to build high-level project plans, identify roles and responsibilities, and identify the user segments you want to concentrate on. This is also the time to establish usability goals.

Obviously, it is too early to usability test an interface for this new project. What may make sense, however, is to conduct summative tests on an earlier version of the product or of a competing product. A summative test here will establish a usability baseline against which you can measure later efforts.

A (usually) lower-cost alternative to a usability test at this stage is a heuristic review of a previous or competing product. This expert's review of an interface's usability can provide a good starting point for an Envision session's discussion of usability goals.

Often, rough or high-level design sketches emerge from an Envision Session. It is not necessarily too early to do a quick, informal formative test to see how users react to these early efforts. These sketches and any test results, however, are preliminary. Information gathered during the Discovery stage typically changes the design direction.

Discovery

During the Discovery stage, you will perform the user research and analysis that you need to be certain that your product's design will help users accomplish their tasks. Typical products of the Discovery stage are personas (a type of user model), scenarios (a type of task model), and requirements.

Usability tests are not typically performed as part of the Discovery stage, but they can be quite illuminating if you have a previous version to analyze.

At this point, it is not still too late to do a summative test, as described in the last stage. You could also do a formative test of a previous version, in which you worry less about task times and instead encourage users to talk during the test. The goal of such a test now is to illuminate how users think, discover the words they use, and find their problem areas.

There are, of course, other ways to conduct user research, such as Contextual Inquiry (described in Hackos and Redish's *User and Task Analysis* or in Beyer and Holtzblatt's *Contextual Design*). Interviews, surveys, and focus groups are also common methods.

Design Foundation

The Design Foundation stage is the iterative part of user-centered design, and it offers the greatest opportunity for usability testing to affect the finished product.

This it the point at which the project team takes the user information and requirements gathered during the Discovery stage and turns them into a high-level, conceptual design for a user interface. It's a bit like creating a detailed outline of a term paper before you begin writing it.

The first design iteration during Design Foundation is often preliminary, but it should still be suitable for usability testing. Such a test is often informal, and with only enough users to see if the overall navigation and metaphors make sense. Paper prototypes are common at this point. In fact, a heuristic review of the first iteration design often makes sense, to get the obvious flaws out of the way. Cognitive walkthroughs – a less task-based method of evaluation – also make sense.

As iterations proceed, the team continually refines the design prototype. As the prototype gains definition and

depth, the formative usability tests can get more detailed and more task-based.

Depending on the time and money available, it often makes sense to do two or three rounds of informal tests and then a bigger, formal test before a final revision and "lock down" of the design.

Design Detail

The conceptual design is done, and the team has (at least in principle) locked down the design. The Design Detail stage, then, is the time to produce the complete specifications for the user interface. This is the time to worry about field labels, tab order, button names, and other detailed matters.

Usability tests at this stage, then, have a different focus. They are still formative tests, but they should concentrate on specific screens or critical task flows rather than overall navigation or underlying concepts in the user interface. In fact, if your test suggests changes to these fundamental aspects of the UI, you risk frustrating stakeholders and losing credibility.

Build

During the Build stage, the design team typically reacts to development problems as needed. The design work is done, so this is a time for troubleshooting. If developers discover that a particular screen won't work well as specified, they may come back and ask for an alternative design. Sometimes such design changes call for a quick, informal formative usability test to be sure that users will be able to understand the new screen.

Additionally, sometimes it may not be possible to fully test a screen until it is finished, with a working database behind it. If they are critical screens, it may be worth usability testing them.

The same warnings that applied to usability testing during the Design Detail stage also apply to the Build stage. Do not usability test anything that cannot be changed.

The Build stage is also the first opportunity to conduct a summative test that measures and documents the product's usability, as compared with earlier versions or the competition.

Release

After the software has been built and tested, the Release stage begins. The team puts any finishing touches on the installation and packaging materials, ensuring that the users first experience with the product is a good one. Or, perhaps, the team works out the final details of the web site registration process. After product rollout, the team may need to do follow-up surveys or otherwise measure user satisfaction.

Depending on your organization's need to demonstrate usability improvements, the team may perform a summative usability test during the Release stage. The results of this test, at least in terms of design changes, must be fed forward to future versions – it's too late to change the product at this point.

One useful type of summative test is called the "out-of-the-box" test. Whether your product has an actual box or not, this type of test asks users to begin using the product from scratch. The results can help the team refine packaging materials, smooth web site registration, and tweak documentation. They are, of course, most useful before the actual product rollout.

HOW TO MANAGE USABILITY EVALUATION

Convincing a team that a usability test is a good idea is only half the battle. It doesn't matter if a manager is sold on the concept if she has no idea how to plan the test, or work it into her schedule and budget.

Schedule and Budget Estimates

The key to figuring out how long it will take to conduct a usability test is a Work Breakdown Structure – a detailed listing of the activities required to conduct the test. Each test is different, but there are enough similarities to arrive at a template activity list:

Activity
Create Evaluation Plan
Determine product availability
Review and learn product
Establish goals for the usability evaluation
Coordinate test platform
Design evaluation tasks
Create evaluation plan
Recruit Participants
Identify participants to recruit
Create Recruitment Screeners
Obtain contact info for participants
Recruit participants
Prepare Evaluation Materials
Set up product to be tested (screens, wireframes, etc.)
Create/assemble scripts, forms, materials

Activity
Facilities logistics (find lab, etc.)
Prepare lab equipment
Conduct dry run
Conduct Evaluation
Sessions: x participants in y days
Analyze evaluation data
Present Findings
Prepare draft report/presentation
Review draft report/presentation
Prepare final report/presentation
Present final report/presentation
Facilitate decision-making

Typically, tests require at least a facilitator and an observer. For each person involved in the test, estimate the number of hours they will spend on each activity. Multiply these figures by their hourly rates to arrive at a labor estimate for the usability test. Then add in the fixed costs (lab rental, etc.), if any, to arrive at the final cost estimate.

To estimate the schedule, go back to the Work Breakdown Structure. Using the activities, and assuming 6 useful hours per workday, determine the calendar schedule for the usability test. (Note: 6 hours per day is not to suggest that we're all goofing off, but rather to acknowledge the fact that we get called into staff meetings, attend training, answer e-mail, and perform other useful – but not project related – activities.) Remember to account for any dependencies. For example, the team can't have one person begin recruiting participants before someone else finishes the recruitment screener. On the other hand, work on the screener and on the test script can go on concurrently.

In general, it is the manager's job to make the call as to what takes too much time or costs too much money. We have found, though, that it is much more difficult to get time and money out of an established project plan. Feed this information to a manager early, and you stand a much better chance of getting at least some of the usability tests you ask for.

Problems With Usability Tests

Another major part of a manager's job is assessing and dealing with risk – unexpected events that make a project's schedule or budget deviate from original estimates. (Risk can even be positive.)

A major argument for usability tests is that they reduce the risk of rework. Discovering a user interface's problems during the design stage can prevent huge costs later in patches, customer relations, and emergency releases.

Usability tests are not without some risks themselves, however. One of the most common problems is actually finding users to test. If the design team is not allowed to contact users directly, work with marketing or sales representatives who can . Such gatekeepers usually have legitimate concerns about employees making promises or otherwise embarrassing statements to their customers. Perhaps the solution is to identify a short list of customers that the design team may contact without approval. Whatever the solution, try to make it a repeatable one – one that you can use for later projects as well.

Another common problem is making the prototype ready to test. If, for example, a clickable prototype is stored on an internal network, it will be difficult to do a remote usability test with a user outside the company. Try to reserve space on an external server that you can use. You may need to negotiate terms, such as removing all files after each test or round of testing. Again, aim for a repeatable solution.

CONCLUSION

Usability testing is a powerful tool — One that can do much good in the product design universe. Used inappropriately, however, you can hurt your credibility by pointing out serious problems that no one has the ability to fix. Understand what you can achieve with a usability test at each stage of a project, and you can recommend only those tests that will have the maximum impact on the final product. You can also improve your odds of getting the recommended tests done by estimating the test's cost and schedule ahead of time, as well as pointing out the pitfalls the test will help you avoid.

REFERENCES

(1) http://www.lucidcommunity.com/
(2) Beyer, H. & Holtzblatt, K. (199?). Contextual Design: <The rest of the title>, etc.
(3) Dumas, J. & Redish, J. (1999). Practical Guide to Usability Testing. Intellect.
(4) Hackos, J. (1994). Managing Your Documentation Projects. John Wiley & Sons.
(5) <Hackos & Redish. User & Task Analysis>
(6) Project Management Institute. (2000). A Guide To the Project Management Body of Knowledge. Project Management Institute.
(7) Rubin, J. (1994). Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests. John Wiley & Sons.

Scott McDaniel
Interaction Designer
Cognetics Corporation
1320 Fenwick Ln., Suite 209
Silver Spring, MD 20910 USA
(301) 587-7549

Scott McDaniel is a user-centered designer with more than 8 years of experience in designing and documenting software, networks, and telecommunications systems. As a designer, he helps establish vision and direction for products, researches prospective users, determines requirements, and provides detailed designs and specifications for both software and web products. He is a member of the Usability Professionals Association (UPA), The Society for Technical Communication (STC) and founded and managed the STC DC Usability Special Interest Group.

Laura Snyder
Interaction Designer
Cognetics Corporation
51 Everett Drive
Suite B-103, PO Box 386
Princeton Junction, NJ
(609) 799-5005

Laura Snyder is a user-centered designer and manager of the Princeton-based consulting staff at Cognetics Corporation. As a designer, she helps establish vision and direction for products, researches users, conducts testing, and provides detailed designs and specifications for a wide range of products. As a project manager, her focus is on determining the “right” amount of process needed to have a successful outcome for both the customer and the users. She is a member of the Usability Professionals Association (UPA) and STC. She is currently tprogram director for UsabilityNJ, a grassroots group for usability advocates in the New Jersey area.