

Usability Testing

«UCD: User-Centered Software Development»

Prof. Dr. Cl. Müller-Birn, Institute for Computer Science, HCC.lab

June 22, 2015

Usability Testing Introduction

Context for today

Last week

- The Human Memory
- Attention
- Visual Processing
- Color Usage
- Accessibility

This week

Usability testing

- Defining goals
- Conducting test
- Comparing test data

How can you test your high fidelity prototype?

Outline

1. Preparing a usability test
2. Integrating usability testing and heuristic evaluation
3. Conducting a Think-Aloud test
4. Collection data after a test

Learning goals for today

- » You can describe the process of planning and preparing a usability test.
- » You can carry out a usability test based on the thinking-aloud method and you are aware of existing pitfalls.
- » You know how to use and to evaluate results from the system usability scale.

Usability Testing

“Process of systematically collecting data that informs us about what it is like for a particular user or group of users to use a product for a particular task in a certain type of environment” (Preece et al. 2002)

Usability testing

Preparing a usability test

General questions for preparing a test

1

What to test—based on where the product is in development

2

Where to conduct the test—based on choices for lab testing, field testing, testing remotely, or some combination

3

How to test—based on resources, timing, and your goals

What to test - product

Formative evaluation

- » Discover usability problems as part of an iterative design process.
- » It can be used to test very early paper prototypes or partially developed products, the information architecture of the product, or a particular feature by itself.
- » Goal is to uncover as many problems as possible.

Summative evaluation

- » Assess the usability of a prototype, or compare alternatives.
- » It's useful and valuable to do summative testing when you want to confirm that requirements have been met for the product.
- » Goal is a reliable, statistically valid comparison.

Where to test - location

- » In a lab—by testing in your own lab or one that you will rent
- » In a conference room—by reserving a room for the test
- » In the field—by going to the users in their environment
- » At a distance—by testing remotely

Labor testing - Brighton University Usability Lab



Field testing in a UCD software project



Approaches for remote usability testing

Method	Description	Examples
Synchronous (moderated)	Test users and evaluators are separated in space	<ul style="list-style-type: none">• Observing user behavior via webcam or video conference• Audio-Feedback
Asynchronous (automated)	Test users and evaluators are separated both in space and time	<ul style="list-style-type: none">• Auto logging• User-reported Critical Incidents (UCI)• Forum-based online reporting and Discussion• Diary-based longitudinal user reporting

How to test - general test setting

“Typical” test of the product

- » You present users with a number of tasks within scenarios, which gives you similar feedback on their experience with your product.
- » Usually formative for products in development.

Benchmarking

- » You test your product with users to establish metrics, or benchmarks, for the product, as well as requirements for new product development
- » Usually summative for completed products.

How to test - general test setting (*cont.*)

Comparison of designs

- » You present users with two or more designs so that you can see whether a preference emerges.

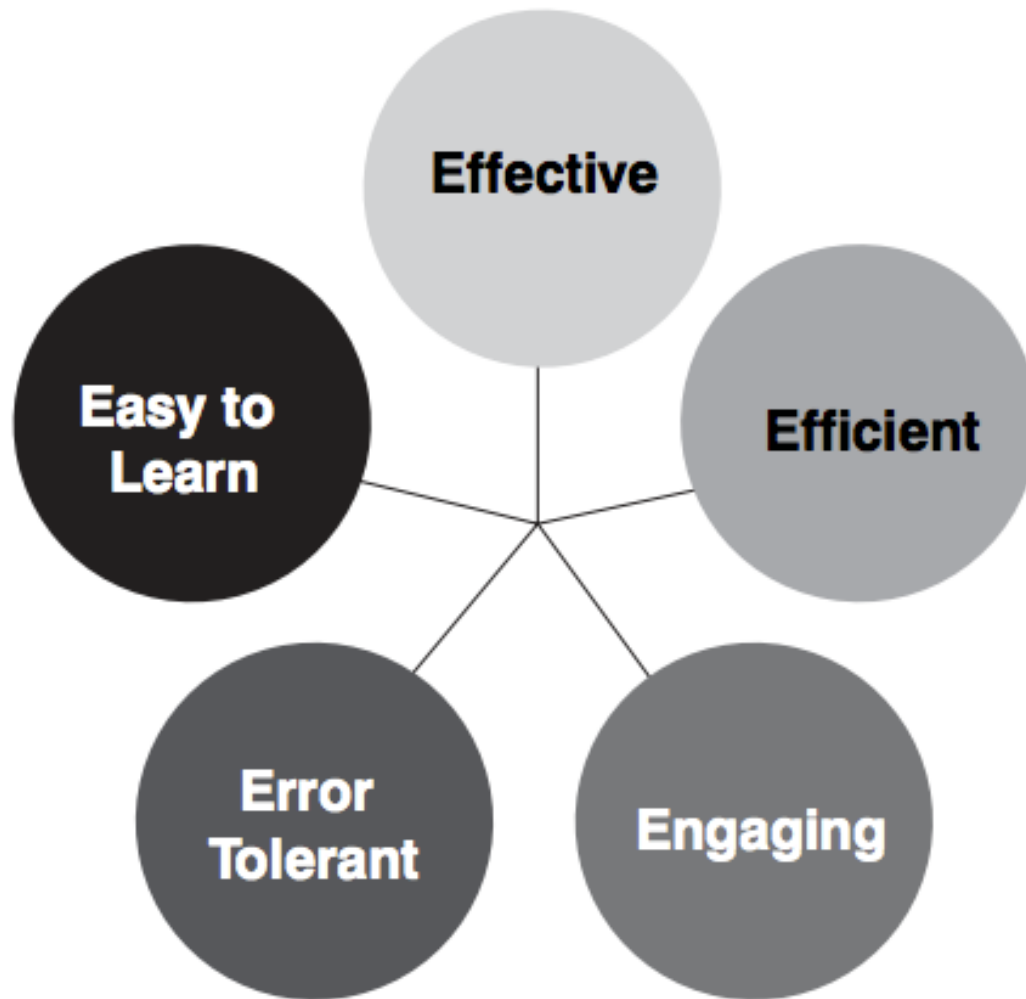
Competitive evaluation

- » You present users with tasks to complete in your product and one or more competitor products to learn their preferences or to measure your product against the competition.

How to test - define your test goals

Test goals focus on what you want to learn about your users' experience with the product at the point in development where you will be testing.

How to test - goals: Whitney Quesenbery's 5Es



Effective: 20%
Efficient: 20%
Engaging: 20%
Error Tolerant: 20%
Easy to Learn: 20%

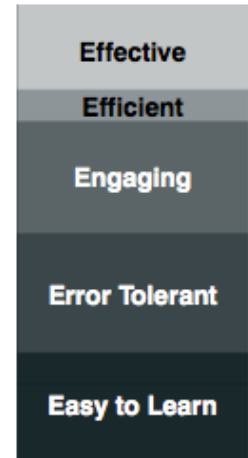
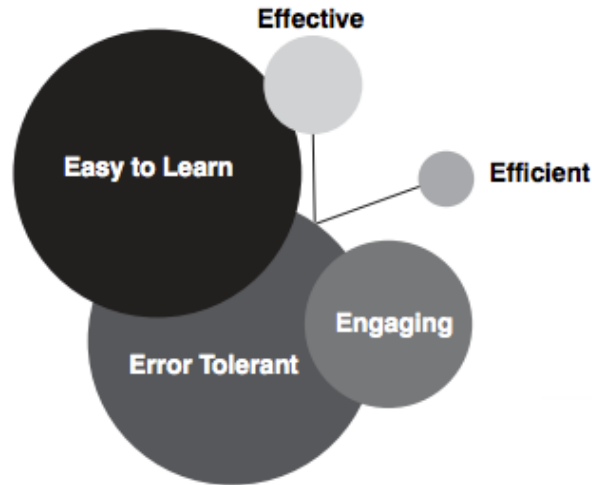
Whitney Quesenbery's 5Es

- » **Efficient:** Can users find the information they need to complete tasks without assistance? Can users perform a process within a predetermined timeframe?
- » **Effective:** Can users successfully place an order or sign up for a service?
- » **Engaging:** Do users rate their experience as satisfying or enjoyable? Do their comments (and body language) suggest that they are having a positive experience?
- » **Error tolerant:** Do users experience errors? If so, how many? And when they experience errors, do they recover successfully? If they receive error messages, do they understand them?
- » **Easy to learn:** Can users get started right away? Does their ability to do tasks improve as they become familiar with the system? Does the system architecture match their mental model for the way they expect the system to work?

Different users, different usability needs

User group A needs

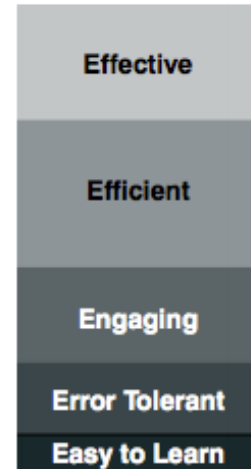
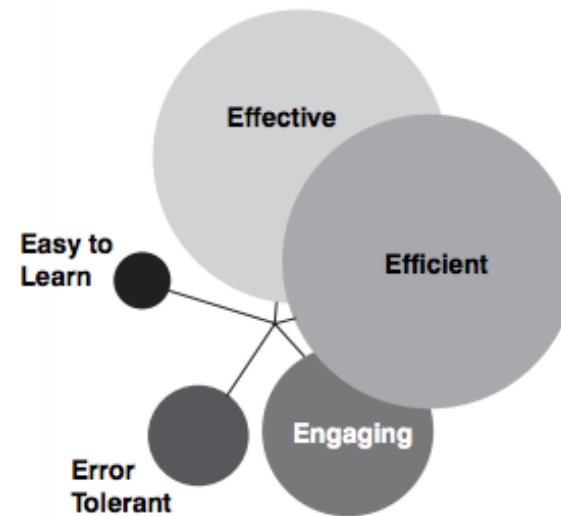
- Good instructions to replace the personal interview (ease of learning)
- Confirmation of not only their data updates, but any impact these changes will have on their benefits (error tolerant)
- Reassurance throughout the process and confidence that they were accurate in their entry (engaging/effective)



Different users, different usability needs

User group B needs:

- To be able to complete routine functions quickly (efficient)
- A good overview of the employee they are working with on a single screen, so they can focus on the conversation rather than the interface (engaging/efficient)
- A way to confirm all changes with the employee before they are made (effective)



Using the 5s for planing your testing

Effective: Watch for the results of each task, and see how often they are done accurately and completely.

Efficient: Time users as they work to see how long each task takes to complete.

Engaging: Watch for signs that the screens are confusing, or difficult to read. Look for places where the interface fails to draw the users into their tasks. Ask questions after the test to see how well they liked the product.

Error Tolerant: Create a test in which mistakes are likely to happen, and see how well users can recover from problems and how helpful the product is.

Easy to Learn: Control how much instruction is given to the test participants, or ask experienced users to try especially difficult, complex or rarely-used tasks.

Usability testing
Integrating
usability testing and heuristic evaluation

Hands-on: Do you remember homework 7-2?

Aggregate and evaluate the results of your heuristic evaluation (submit as a team)

Meet as a group and discuss the various problems found in the team. Create an aggregated list that contains all problems and the violated heuristic. Discuss the severity of each problem found and document your final decision in a document. Arrange all problems in the order of rating (starting with the highest severity).

Submit the final document as a team to blackboard. This list should contain the following information: problem, heuristic, severity.

Hands-on

Work in your team (20 minutes)

Review the results of your homework.

Answer the following question:

- What kind of problems have been found by using heuristic evaluation and how do you addressed them in your high-fidelity prototype? (also include insights gained from the last testing session)

Present the results to the auditorium (3 minutes per group).

Putting both HE and UT together (1)

Do a heuristic evaluation (or expert review) first.

The results from this review can be used as follows:

- » You can clean up the interface to get rid of those issues that can interfere with the user's experience.
 - If there's time, this can be an extensive revision.
 - If there isn't a lot of time, you can go for the "low-hanging fruit."
- » If you don't have time to make any changes to the product before testing, you can use the review to identify your goals for usability testing in order to identify possible problems.

Putting both HE and UT together (2)

Do a usability study.

After your heuristic evaluation, you have either fixed the problems or identified a list of problems that you want to explore in a usability study:

- » If you have fixed the problems before testing, you can learn what additional problems emerge for users.
- » If you haven't fixed the problems before testing, you can learn whether the issues you identified in the review are, in fact, problems for users.
 - When they are, you have increased the validity of your findings by combining your processes.
 - When problems you've identified don't actually bother users, you can learn what doesn't need to be fixed.

Putting both HE and UT together (3)

Use both methods at the same time.

When there's money in the budget to use both of these tools but insufficient time to do one and then the other, divide and conquer.

- » Conduct an expert review and a usability study in the same timeframe, reporting the results separately, then combining the findings for your recommendations.

Pause 15''

Usability Testing

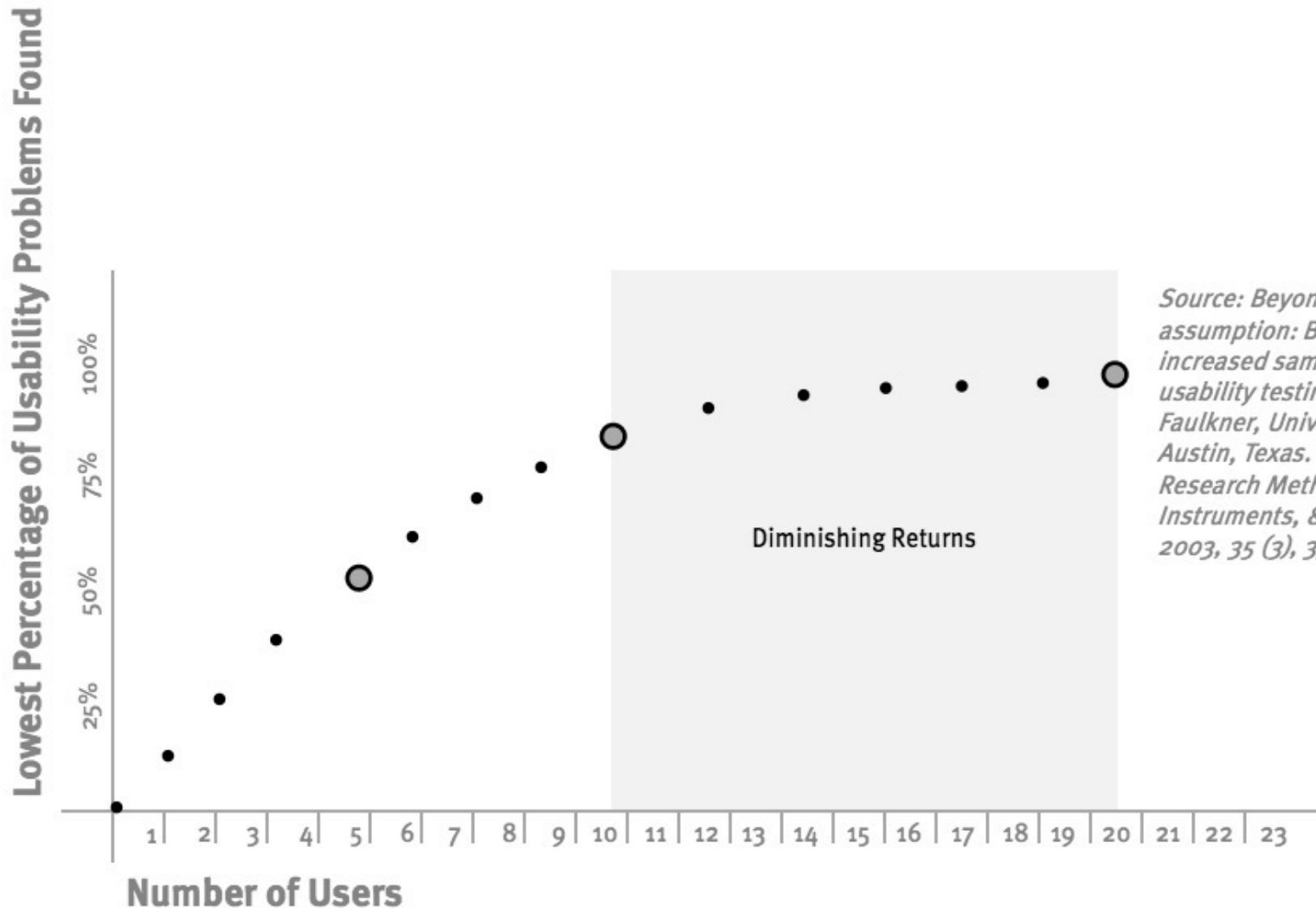
Conducting a Think-Aloud test

Think-Aloud user testing

Evaluating the usability of your work by encouraging a user to think out loud as they use your software or service.

- Is considered a “quick and dirty” method, as compared to a formal experiment

Think-Aloud: how many users?

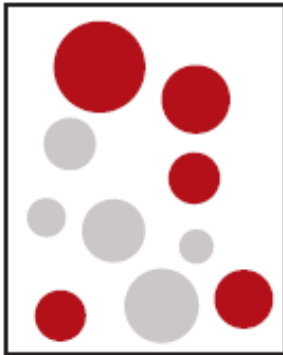


Source: Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. By Laura Faulkner, University of Texas, Austin, Texas. In Behavior Research Methods, Instruments, & Computers. 2003, 35 (3), 379-383

Does the number of testers matter?

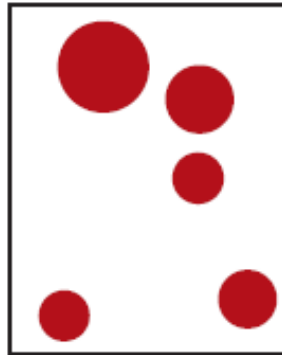
ONE TEST WITH 8 USERS

8 users



Eight users may find more problems in a single test.
But the worst problems will usually keep them from getting far enough to encounter some others.

**TOTAL PROBLEMS
FOUND: 5**



TWO TESTS WITH 3 USERS

First test: 3 users



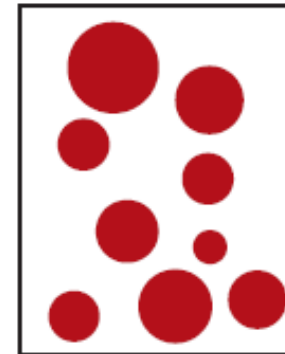
Three users may not find as many problems in a single test.

Second test: 3 users



But in the second test, with the first set of problems fixed, they'll find problems they couldn't have seen in the first test.

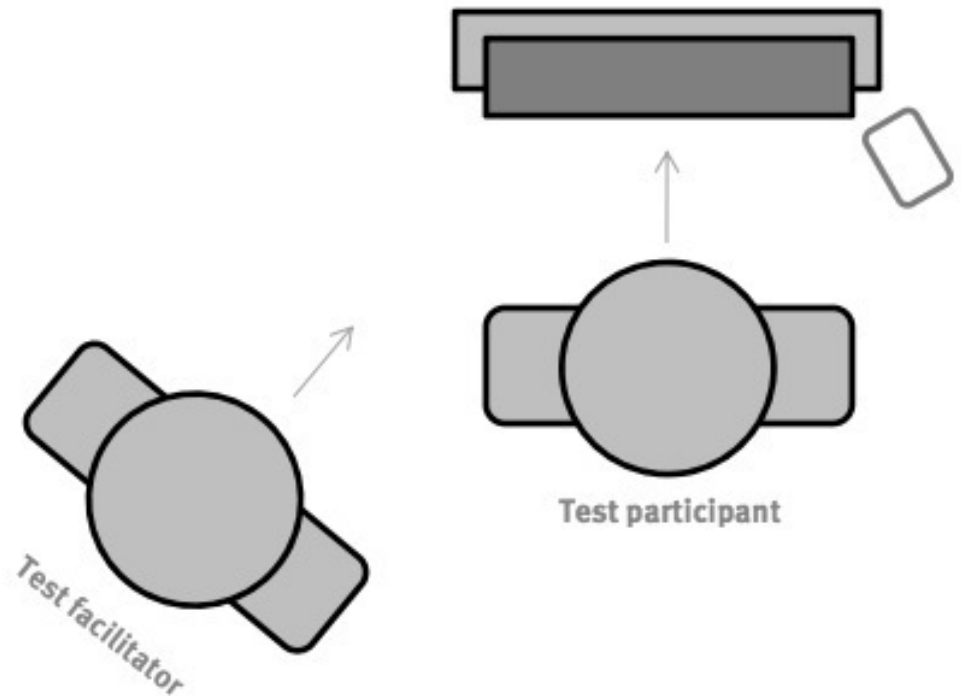
**TOTAL PROBLEMS
FOUND: 9**



Running a Think-Aloud user testing session

Get set up.

- » Prepare the physical arrangement.
- » Get written consent to tape.
- » Don't forget to start taping!



Running a Think-Aloud User Testing Session

Explain to the user:

- » Who you are & what you are doing
- » That you are testing your interface, and not testing them
- » That they can quit at any time
- » That you won't be able to help them
- » That you require them to continue talking, and you will remind them to "please keep talking" if they fall silent

Verify that the user understands the tasks (have them read the tasks aloud too, and ask if there are any questions)

Running a Think-Aloud user testing session

If the user falls silent for more than three seconds, prompt them “please keep talking”

Do not help the user complete a task (if the user asks for help, explain that you cannot help, and prompt them to try what they think is correct)

Don't defend your designs! This is not a critique of your design skills; don't even mention that they are your designs.

While the session is running, do not say things like:

- » “Please explain what you are doing”
- » “Note any design problems you see”
- » “Tell us if you have any suggestions”
- » “Why are you doing what you are doing”

Usability testing

Collection data after a test

Creating a post-test questionnaire

Rather than creating your own post-test questionnaire, you can use standard usability post-test questionnaires, such as ***System Usability Scale***

System Usability Scale (SUS)

- » Developed by John Brooke at Digital Equipment Corporation in 1986
- » It uses 10 Likert-type statements with responses based on a 5-point scale
- » The 10-item questionnaire is short, and free, as long as you acknowledge the source
- » A SUS score is best presented as a single number - ranging from 0 to 100 - representing a composite measure of the overall usability of the system being studied

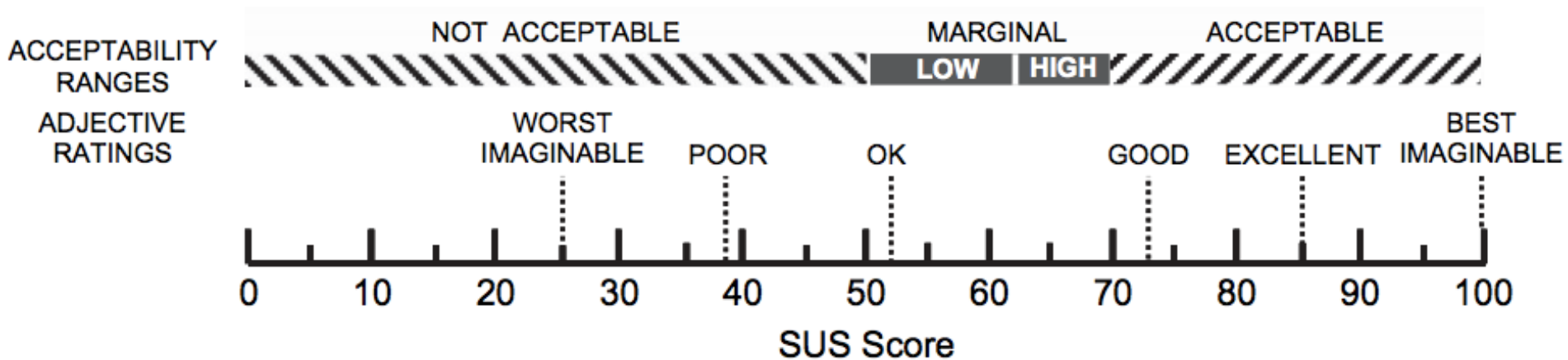
	Strongly disagree				Strongly agree
1. I think that I would like to use this system frequently					✓
	1	2	3	4	5
2. I found the system unnecessarily complex				✓	
	1	2	3	4	5
3. I thought the system was easy to use		✓			
	1	2	3	4	5
4. I think I would need the support of a technical person to be able to use this system	✓				
	1	2	3	4	5
5. I found the various functions in this system were well integrated		✓			
	1	2	3	4	5
6. I thought there was too much inconsistency in this system			✓		
	1	2	3	4	5
7. I would imagine that most people would learn to use this system very quickly		✓			
	1	2	3	4	5
8. I found the system very cumbersome to use				✓	
	1	2	3	4	5
9. I felt very confident using the system					✓
	1	2	3	4	5
10. I needed to learn a lot of things before I could get going with this system		✓			
	1	2	3	4	5

Calculating the SUS score

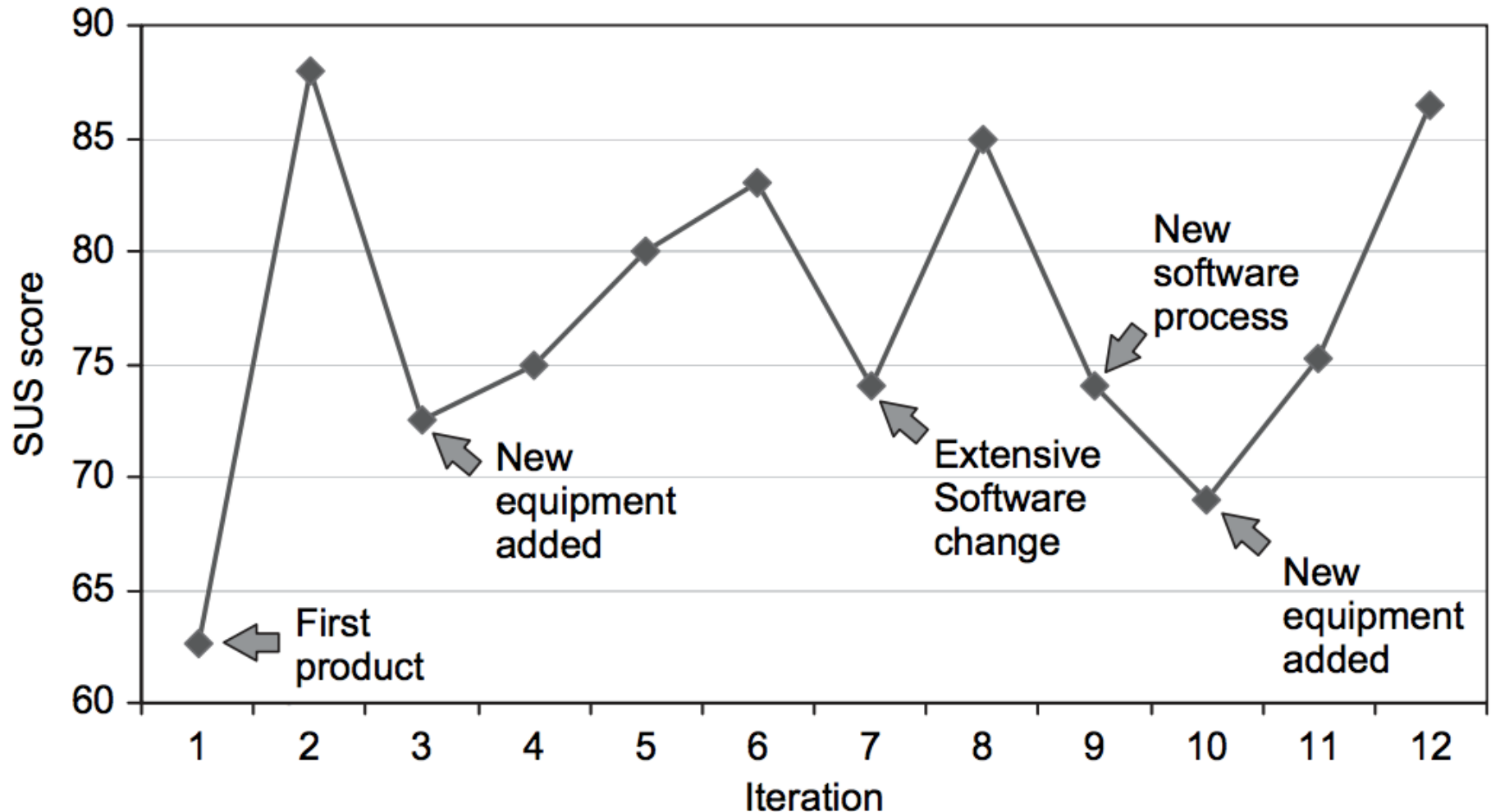
1. For odd items: subtract one from the user response.
2. For even-numbered items: subtract the user responses from 5.
3. This scales all values from 0 to 4 (with four being the most positive response).
4. Add up the converted responses for each user and multiply that total by 2.5.
This converts the range of possible values from 0 to 100 instead of from 0 to 40.
5. Average across all users.

	Strongly disagree				Strongly agree
1. I think that I would like to use this system frequently					✓
	1	2	3	4	5
2. I found the system unnecessarily complex				✓	
	1	2	3	4	5
3. I thought the system was easy to use		✓			
	1	2	3	4	5
4. I think I would need the support of a technical person to be able to use this system	✓				
	1	2	3	4	5
5. I found the various functions in this system were well integrated		✓			
	1	2	3	4	5
6. I thought there was too much inconsistency in this system			✓		
	1	2	3	4	5
7. I would imagine that most people would learn to use this system very quickly		✓			
	1	2	3	4	5
8. I found the system very cumbersome to use				✓	
	1	2	3	4	5
9. I felt very confident using the system					✓
	1	2	3	4	5
10. I needed to learn a lot of things before I could get going with this system		✓			
	1	2	3	4	5

What is a good SUS score?



SUS scores and their relationship to critical events in the product lifecycle process



Last remark

Usability tests do not replace user research

User Research: Data acquisition and analysis to develop a certain product

Usability Testing (UT): Evaluation of an existing product under certain criteria

Challenging aspects of Usability Testing as an analysis-tool

- UT is a reactive analysis
- UT helps to identify problems with the design, but does neither show the exact problem nor a way to resolve them
- One can only observe the user's reactions; the causes or reasons for them remain unclear
- The results of a UT are limited to the current task and can not be generalized to the whole application

Reader

Preparing for usability testing

Essential test planning steps

1. Scheduling the planning meeting, in which you:
 - » establish test goals
 - » determine how to test the product
 - » agree on user subgroup(s)
 - » determine participant incentive
 - » draft the screener(s) for recruiting participants
 - » create scenarios based on tasks that match test goals
 - » determine quantitative and qualitative feedback methods
 - » set dates for testing and deliverables

2. Writing the test plan
 - » writing an informal test plan
 - » writing a formal test plan

Agree on user groups

Whether you have budget or time for only one day of testing or for several days, you need to decide on the subgroup or subgroups you want represented in your study.

The goal is to define a list of characteristics for each subgroup. Let's call this list of characteristics for a particular subgroup a user profile.

Did you hear about user profiles already?

Create scenarios based on tasks that match test goals

The goal of this step is to think about the questions you want your users to answer.

Let's look at an example...

Question:

Will users look at the top navigation bar to start their search for information?

Task:


Seeking information about online programs for military personnel. Correct choice is *Featured Degrees* in top navigation bar. Users can also find a link to programs for military personnel in the description of featured programs in the center of the homepage, but it may be below the fold on their computer screen.

Scenario:

You have a friend in the military who wants to enroll in college courses while serving. You want to see if there are any online programs your friend could apply for. How would you go about doing this on this website?

GEORGIA
ONMYLINE

My Learning • My Schedule • My Education



Earn an Accredited Degree - on Your Schedule

HomeAboutFeatured DegreesInstitutionsDegreesCoursesQuestionsContact Us

Building on the strength of the 35 institutions of The University System of Georgia

Georgia ONmyLINE provides access to a full array of online and distance education offerings from the 35 colleges and universities in the University System of Georgia. Through this site prospective and enrolled students can quickly find degrees and/or courses that meet their needs. They can apply for admission to colleges and universities, and once accepted, register for courses. The offerings listed in this site will serve both students who seek fully online solutions for their education needs, as well as on-campus students who seek occasional online formats to meet their content, time, and scheduling needs. [read more...](#)

Search for Degrees

Search for Courses

Featured Degree Programs

You have a wide variety of choices when it comes to getting an education online. So why not choose a degree from the University System of Georgia? The accredited degree programs below are designed to meet your educational needs when, where, and how you need them.

[Online graduate teacher education degrees](#)

[Online degrees and certificates for military personnel to pursue](#)

[Online criminal justice programs at the associate's, bachelor's, and master's level](#)

[Other new featured degrees coming soon!](#)

eCore is here!

eCore (electronic core-curriculum) is not a formal degree program but is a convenient and adaptable option for either getting started in college or for resuming college after a break or pause. The core curriculum primarily consists of courses required during the first two years of college for a given degree. Check out [eCore: Georgia's college core curriculum... online](#)

Questions

Below are frequently asked questions about Georgia ONmyLINE. If the following do not answer your question, feel free to [contact us](#).

[How are financial aid, tuition, and fees handled?](#)

[How do I apply for admission?](#)

[How do I transfer credit?](#)

[Will my degree reflect that I took my courses online?](#)

[How are online courses different from on-campus courses?](#)


[What can I expect an online course to be like?](#)

[How do I know if online courses are a good choice for me?](#)

[What about technical support?](#)

Related Links

- [Georgia ONmyLINE Video ad](#)
- [Academic Calendar](#)
- [Institution Map](#)
- [Are You READY for Online Learning?](#)
- [GAcollege411](#)
- [Electronic Campus](#)
- [eCore: Georgia's College Core Curriculum Online](#)



Hosted by:
Board of Regents of the University System of Georgia
"Creating A More Educated Georgia"

Privacy | Contact Us
©2010 Georgia ONmyLINE
Last modified: November 10, 2009

Decide on your first scenario

Deciding on your first scenario should be based on the answers to the following questions:

- » How do you want your users to first experience your product?
- » What's the best starting point?

At whatever point you start the test, it's often a goal to capture users' first impressions, especially when the study is formative.

Possible questions:

- » What type of site or application is this?
- » What do you think you can do here?
- » Look over the tabs or links and share what you think these mean and what you think will happen if you click on them.
- » What would be the first action you would take?
- » Are there any words or labels that you don't understand?

Organization of the “other” scenarios

- » A sequence of tasks that has to be performed in a particular order
- » Most frequently performed tasks
- » Tasks getting the most calls to the help desk
- » The most critical tasks (which may not be tasks that are performed most frequently)
- » New tasks/new features
- » Tasks or task flows that are the subject of internal debate (the user
- » Can show the team what works, what doesn't)
- » Comparative tasks using alternate designs of your product or your product and a competitor's product

Determine quantitative and qualitative feedback methods

Depending on whether you are conducting a formative or summative evaluation and what your goals are for the study, you may want to focus on one type of data collection or another.

Quantitative data are performance and preference data.

- » Performance data are based on measurements of users' actions, such as time on task; number of errors; recovery from errors; success or failure at task completion; use of help, documentation, or embedded assistance.
- » Preference data are based on users' responses to questions on post-task and post-test questionnaires.

Observations of your participants yield rich qualitative feedback. Qualitative feedback is gathered by noticing what participants do while they are engaged with the product.

Needed activities to prepare for testing

Recruiting participants

Assigning team roles and responsibilities

- » Developing team checklists
- » Writing the moderator's script

Preparing or using other forms

- » Preparing a video consent form
- » Preparing a special consent form when testing with a minor
- » Using a non-disclosure agreement
- » Preparing an observer form

The moderator's checklist

- » Welcome the participant.
- » State the purpose of the study.
- » Provide forms required for participation, unless these have been completed in advance.
- » Explain the testing process.
- » Describe thinking out loud.
- » Ask the participant to share any questions or concerns.
- » Start the study.

Preparing or using other forms

Preparing a video consent form

- » Participants will need to sign a consent form giving you permission to record them for the test

Using a non-disclosure agreement

- » Whenever you are working on a product in development, you are likely to need a non-disclosure agreement (NDA)

Preparing an observer form

- » If you have observers who will be contributing to the discussion of the findings, you can create an observer form to standardize the method they use to take notes.

Participant number/name:	Date/time of session:	
Scenario/task	Issue/problem	Observer's name and comment

Creating questionnaires

The number of questionnaires you will need to create depends on the format of your study.

If your study is being conducted very early in product development, such as at the paper prototyping stage, you may not need to create any questionnaires.

If, however, you want to get structured feedback from participants before, and at the end of the test, you will want to create questionnaires. The types of questionnaires you will want to consider developing include:

- » pre-test questionnaire
- » post-test questionnaire