Rocket Surgery Made Easy

Chapters 1 - 7

What is Usability Testing?

- Evaluating a new or existing design to see if the purpose and interactions with the design are well-understood by users.

- Ideally, we want to either:
  - Make something easier to use (Qualitative)
  - ***Prove*** that something is easy to use (Quantitative)

- Quantitative studies:
  - Are Formal and Rigorous
  - Measure something and are Data-driven
  - Minimize interaction with participants to avoid influencing results
  - Require a significant sample to show generalizability

DIY Usability Testing

- Qualitative
  - Less rigor / formal
  - Focused on outcomes and improvement rather than formally “proving” something
- “Less” Scientific
- Requires fewer participants for feedback
- The protocol (study) can change between or during testing

The General Process

- Identify areas / features of the product to test
- Create scripted instructions for the participant
- Recruit Participants
- Gather developers / stakeholder to be observers
- Run the usability study
  - Observers collect notes
- Debriefing session
  - Discuss changes

It just works!

- But why?

It just works!

- But why?

- All sites/products have usability problems.
  - Nothing is perfect

- Serious problems are usually easy to find.
  - We want to avoid these and address them first

- Watching users makes you a better designer.

If it’s so great, why is so little done?

- Lacking firsthand experience with usability
  - The value prospect is not realized
- Lack of time
- Reluctance to show work that isn’t “done”

If it’s so great, why is so little done?

- Lacking firsthand experience with usability
  - The value prospect is not realized
- Lack of time
  - Why not only usage analytics?
  - Analytics data tells you “what” people are doing, but not “why.”
- Reluctance to show work that isn’t “done”
  - Fun Fact: Software is not “done” until it is no longer supported.

Watch the video to see how it’s done!

https://youtu.be/1UCDUOB_aS8

A morning a month.

It’s all we ask….

“A **Morning** a **Month**”

- Testing is done **once a month** with **three participants**.

- Why a month?
  - It’s a reasonable time frame for developers to accomplish meaningful work
  - The time between the following month’s tests can be used to fix the findings of the previous test.

- Why a morning?
  - The goal is for the process to take roughly half a day
  - If done in the morning, a debrief can be done over lunch, and testing for the month is done by late afternoon.

“A **Morning** a **Month”** is a guideline

- Does it actually only take a month?
  - For most of the team, yes.
  - For the facilitator of the test, no.
  - Preparations need to be made in advance of the test.
- Can you do it more than once a month?
  - Yup. One is the minimum. The key is consistency.
  - Agile sprints may not align with once-a-month testing
  - The key is keeping it a routine and not a decision, or it won’t get done.
- I’m not a morning person….
  - Fine, do it in the afternoon or evening.
  - The goal is to have the testing take no more than half a day

This doesn’t seem very scientific…

- It isn’t…and the results will not be statistically valid with only three people.

- If you need that, run a quantitative (“big honkin’”) test
  - See the big tables of differences between DIY-Usability Testing and the Big Honkin’ Test in Chapter 3 of the book.

- The goal is to catch apparent issues with a lightweight, low-cost, low-time commitment testing method.

Test Earlier Than You Think Makes Sense!

What to Test?

- Existing product/site
  - A product undergoing redesign or enhancements
  - Learn from your own mistakes
- Competitor or similar product/site
  - Learn from the mistakes of others
- Sketch
  - Can people figure out what you're making
- Wireframes
  - Find issues with terminology, layout, navigation
- Site/Product Design
  - Has the design affected usability?
- Prototypes, Pre-Releases, Releases, etc.
  - Usability issues that become present as the software evolves
  - Improvements to the product

Usability Study Participants

- What kinds of people do you test with?
- How many do we need?
- How do you find them?
- How do you compensate them for their time?

Test Participants

- Target participants that reflect your audience, but…
- **“Recruit loosely and grade on a curve.”**
- Getting “representative users” is problematic as individuals differ even from the same domain (field of work).
  - “Novice” vs. “expert”
  - Some knowledge may not be as “common” as assumed.
- Serious usability problems often don’t require specific expertise.
- You can always add one target user per test to get “niche” feedback.
- Don’t reuse participants for testing the same product

Three Participants is Enough

- They are likely to encounter any major issues
  - They will miss things, too; this is why the testing is done regularly.
- It's less effort than finding more people
- It's more important to do more regular rounds of testing than large-scale testing.
- Three users allow for testing and debriefing in the same day
- It is easier to encourage observers to a shorter testing session
  - Longer sessions have diminishing returns, and people's attention fades
- Too much feedback
  - More notes than can be processed during a debrief
  - More "nitpicking" than serious issues
  - How to prioritize resolutions?

How to Recruit

- Go where the people are depending on your demographic
  - Colleges/Universities
  - Senior Centers
  - Conferences
- Personal Acquaintances
  - Family and Friends
- Remote Studies
  - Expands the pool to anyone with internet
- Blanket Invitations
  - Message boards, Emails, or Social Media
- **DO NOT**
  - Recruit people from the company who are familiar with the product
  - Let marketing recruit participants

What to Do With the Recruits

- Screen them with a simple phone interview
  - Are they available on the test day?
  - Do they meet your qualifications?
  - Explain the study and expectations.
  - Describe compensation

- If they seem to be a good candidate, schedule them.

- Follow up with an email with details after the call and another follow-up a few days before the testing to confirm their attendance.

Compensation

- Depends on the participants
- A thank you letter may be sufficient
  - Some people cannot accept gifts
- Company “swag” as a memento
- Monetary Offerings
  - Gift Cards are the easiest method of compensation
  - Common is ~$50 for an hour
  - The more expertise, the more compensation
  - Higher than the going rate can encourage more reliable participation

Recruiting Sounds Like a Lot of Work?!

- Yep

- You can also pay someone else to do it for you.
  - Focus group rentals
  - Market research

- Outsourcing recruiting will widen your participant pool and save you time…for a modest price.

What if Someone Doesn’t Show?

- Always have a remote standby that you can use as a fill-in.

- If that fails as well, take almost anyone.
  - It is better to fill the slot than lose the feedback and observers
  - Also, you will do multiple rounds of testing, so one ”imperfect” scenario isn’t too large of an issue

Creating Testing Tasks

- List five to ten of the most important things people need to do with your product.
  - Make sure these are user goals, not your goals…
- Which things are most critical?
- Which ones do you think people struggle with?
- Look into other feedback
  - Common customer service issues
  - Red Flags from Web Analytics

Creating the Testing “Scripts”

- Convert each task to a readable script that explains to the participant
  - The script frames the task as a short “story,” giving context and the goal of the activity
  - Include any needed information the user may not have (login credentials, etc.)
- Don’t give any hints in the scenarios.
- It is okay to place restrictions on users, such as not using search or opening other applications, etc.
- Pilot the test scripts with anyone (family, friends, etc.)

Rocket Surgery Made Easy

Chapters 8 - 12

The Role of the Facilitator

- **Tour Guide**
  - Keeping them moving through the process
  - Keep them happy (encouragement, etc.)
  - NO HELP OR HINTS/TIPS ON SCENARIO COMPLETION

- **Therapist**
  - Encourage thinking out loud (literally called the “think aloud protocol”)
  - Have them narrate events
  - Verbalizing feelings
  - Avoid influencing them
  - Keep in mind ethical responsibilities

- Effectiveness in these roles yields design insights unavailable to other methods.

The Testing Room

- Computer and peripherals
  - Necessary to use the product
  - Stick to standard interaction devices unless your product requires special devices
    - Avoids distractions
  - Use displays with a common resolution
    - Perhaps 1080 vs 4K+
- A high-quality microphone
  - If you can’t hear them, you can’t get good data
- Speakerphone
  - You want observers to be able to listen as well without having to crowd everyone around your participant.

TEST ALL YOUR TOOLS TO MAKE SURE EVERYTHING WORKS!

The Testing Schedule

- Pre-Test Prep (60 min)
- Welcome (4 min)
- Pre-Scenario Questions (2 min)
- Main screen tour (3 min)
- The Tasks/Scenarios (35 min)
- Probing Questions (5 min)
- Wrapping Up (5 min)
- Prepare for the Next Test (10 min)

Ethics

- People should leave the room in no worse shape than when they entered.
- They can leave whenever they want WITHOUT penalty.
- Do NOT use identifying information about them during the test
  - No face recording, Last names, etc.
- Keep the records under your control and dispose of them when they are not needed.
  - If you need to distribute the data, redact any personal information
- In academia, a study may require IRB approval. **DO NOT SKIP THIS PART!**

Observers

- Make it a “spectator sport”
- Everyone is invited
  - The more people involved, the more people can see the benefits
- You need to:
  - Make it easy to attend
  - Advertise
  - Present the benefits/value
  - Get management (or higher) involved
  - Provide the good snacks

Observer Responsibilities

- Watch, learn, and take notes
- After each session, write down the three most important usability issues they noticed
- Suggest questions for the facilitator
- Enjoy snacks
- Come to the debriefing session

Observer Room

- A computer screen casting the product being used during the test
- A projector or large display so people can see
- Speakers that can get decently loud (depending on the number of people and the room size).
- Snacks…seriously, it’s important. Get ‘em what they like.
- A speakerphone to hear the participant and facilitator.
- DO NOT PUT THE ROOM NEAR THE TESTING ROOM
- MUTE THE SPEAKERPHONE IN THE OBSERVER ROOM

The Surprise Hero

- Have someone help you to manage the observers.
  - Your “Hall Monitor”
- Make sure everyone
  - can see and hear the test
  - gets a copy of the observer instructions, test script, and scenarios
  - has something to take notes on
- Keep them on-topic
- Let them know where and how to reach you
- Have people step out if they need to make phone calls
- Remind everyone to take notes

Debriefing

- Only people who attend at least one test session may take part in the debrief
- Focus ruthlessly on only the most serious problems
  - You won't have the time or resources to do everything
  - Easy to get caught up in the minutia
- Is a problem "serious"?
  - Will a lot of people experience the problem?
  - Will it cause a serious problem for people who experience it, or is it just an inconvenience?
    - Severity is a judgment call.

Running the Debrief

- Everyone should look at their notes and pick their three most serious issues.
- Everyone goes around the room and reads the issues aloud
  - All issues are written down on an easel, whiteboard, chalkboard, etc., so everyone can see
- Rank the top 10 issues
  - This may or may not require discussion/voting
- Write a fix for each of the top ten
  - Do not skip any
  - Do not let the fix take more than one month. It will “slip” and never get done.
- Product a short email summarizing
  - What you tested
  - The list of tasks
  - List of problems to fix in the next month based on observations
  - Where people can see the recordings and when the next test will be

Fixing Usability Issues

- “What’s the smallest, simplest change we can make that’s likely to keep people from having the problem we observed?”
- Help users now! You can iterate on the solution to make it better.
  - Done > Perfect
- If it’s a “core issue,” look for mitigations until a better solution can be implemented
- Even if it will change eventually, implement a fix
- Do the least you can do!

Tweak, Don’t Redesign

- Costs less
- Less work
- Minimal impact on team members
- Small changes can be made sooner
- Small changes are more likely to happen
- Redesigns have inherent complexities and risks (lots of change at once)
  - Large changes are more likely to break other working things
- People resist change, so redesigns annoy them
- Resigns involve more coordination with people and more meetings

Take Something Away

- It’s easy to add things
  - More text to explain
  - More color, emphasis, size, etc.

- Sometimes, less is more.

- People may be overwhelmed or confused by how many options/features/information they need to wade through to perform a task.

- “A designer knows he has achieved perfection not when there is nothing left to add, but when there is nothing left to take away.”
  - Antoine de Saint-Exupéry

Common Issues

- Getting off on the wrong foot
  - This can happen with the first page/screen of an application
  - A mistake from starting with the assumptions/understanding
  - We need to clarify the purpose and make guidance, navigation, and feature affordances clear

- Failure to shout
  - Subtlety is less effective for usability
  - Make your calls to action clear and prominent

Common Organizational Issues

- Change of management, direction, or both
- Putting things off
  - technical debt
- Lack of commitment from the right people
- Sabotage
- Overzealous resolutions
- Deep-rooted issues for usability may be a symptom of unresolved conflict or organizational dysfunction.