1. 4.6: Usability Test Results & Planning for Future Tests
2. Learning Goals

* Interpret and classify usability testing results

 Estimated Read Time: 50 Minutes.

1. Introduction

Welcome back! How did your usability tests from the last Exercise go? Did you find any big usability errors or mistakes in your design? If so, then great! That’s exactly what the test is for. Prototypes aren’t meant to be *perfect*—they’re meant to be perfected. Now, however, comes the billion-dollar question: How do we take all this real human feedback, analyze the information, and decide what to do next?

In this Exercise, we’ll be discussing how to download and synthesize the information you garnered from your usability tests. It’s important that we first make sense of what happened before we make design decisions or communicate results to a wider team of clients or colleagues. We’ll do that by methodically sorting our results using affinity mapping and a tool called the Rainbow Spreadsheet.

As a quick note before we start, we’d like to warn you that this process can be quite lengthy. You’ll be going through all the audio and/or video footage you recorded from your testing sessions, which can take considerable time. Be sure to set aside several hours for this Exercise’s Task or work on it over a period of two to three days.

1. Affinity Maps for Usability Testing

One of the best ways to start interpreting information from your usability tests is via an affinity map, also referred to as an “affinity diagram,” the “K-J method," and various other titles. We tried out some affinity mapping back in Achievement 2 while we were sorting through our user interview data. Now, let’s dig in a bit deeper and see how we can use the process for our usability test data, as well.

“Affinity,” from a scientific standpoint, is defined as the degree to which a substance tends to combine with another. The point of making an affinity map is to isolate the varied bits of information, extract them from your video or audio, and transfer them onto paper. Once these bits of information have been isolated, they can be more easily grouped, allowing you to tease out patterns of behavior or sentiment across multiple participants.

The basic concept of an **affinity map** is simple: isolate information and determine how it relates to *other* bits of information. This is a flexible framework that was originally developed by cultural anthropologist Jiro Kawakita in the 1960s. Note that it’s not exclusive to usability testing (or even UX design!) as the process itself is applicable to many other types of generative research and brainstorming, as well. Fundamentally, affinity maps help designers:

* Isolate concepts and information quickly
* Visualize bits of information so that previously unseen patterns and hierarchies can emerge
* Extract and group information collaboratively within a team

TIP!  
Affinity mapping is a great way to visualize information in a fun and collaborative manner. Check out this [Affinity Mapping Timelapse](https://www.youtube.com/watch?v=swsat1e6yFk) to see what it would look like in a big group.

At this point, you should have conducted six usability tests and have around an hour of video and/or audio. Hidden within this hour of video and audio are bits of important information. In Achievement 2, we sorted our interview results by clustering similar ideas and insights together with sticky notes. We’ll repeat that process here, with the added benefit of some top-level categories to help guide us.

The rest of this Exercise will be spent walking through the procedure for downloading and grouping information from a usability test. We’ll start with affinity mapping, then transfer our information to the Rainbow Spreadsheet. You may wish to do a quick read through of the Exercise first to get a good overview of what you’ll be doing, then walk back through the process and try to apply it to your own results.

1. 1. What You’ll Need



Before you get started, gather the materials you’ll need to help you map. Here’s a quick list:

* **Sticky notes** work great as a tactile, mobile, and flexible option, but are not required. Alternatively, **whiteboards** provide a lot of good space and are easy to erase.
* If you don’t have access to sticky notes or a whiteboard, [Note.ly](http://note.ly/), PowerPoint, or Keynote are nice digital options.
* A **Sharpie**, **marker**, or, if you’re typing, a **large font**. Larger writing forces you to simplify information and keeps you from getting lost in the details. Most observations should be one or two lines max. If you have a longer observation, consider simplifying it or breaking it up into multiple observations.

All set? Let’s go ahead and set up some categories.

1. 2. Create Top-Level Categories

You’ll want to create top-level categories to help you sort your observations. In our generative research phase, we wanted to observe trends before making design decisions, which is why we performed more of an "open card sort" affinity map. Here in our evaluative research phase, however, we want to refine an *already existing* design and keep our insights more focused. Thus, we’ll be taking a “closed card sort” approach and use set, top-level categories to help sort our observations.

* On a whiteboard, with sticky notes, or in your document, write the following categories: **Observations**, **Positive Quotes**, **Negative Quotes**, and **Errors**. Give yourself space to work, as you’ll end up with multiple observations in each category.



1. 3. Watch the Video and Listen to the Audio

This is the time-consuming step. Make sure you set aside plenty of time to go through your video and audio thoroughly. If you do decide to split your analysis up into segments, try not to space out the sessions too much. You need to be thinking of patterns between the various bits of information, so you’ll want everything to be fresh in your mind.

Watch or listen to all of your video or audio and record important observations as you go along. Follow the process below. Don’t worry about classifying the information yet. For now, simply concentrate on transferring the raw video and audio into concisely written notes.

* Give each of your participants a label (we recommend P1, P2, P3, P4, P5, and P6 to keep things simple). On each sticky note, write the label of the participant involved, as well as the version they were testing when they made the observation. Each sticky note should include a short description such as the following in addition to the actual observation: P2 Mobile, P5 Desktop, P3 Mobile, etc.
* Jot down the errors people make, any observations you can construe regarding what participants are doing, thinking, and feeling, as well as important quotes that bring the participants to life.
* Only make one observation at a time. Don’t make multiple observations on the same note.
* Try to use verbs when making observations, for example: “Feeling frustrated during first step of checkout,” “Thought this was a back button,” “Expected the logo to go home,” etc. Using active verbs helps bring participants to life, allowing you to better record and communicate findings.

1. 4. Pay Special Attention to Errors

Errors are particularly important to classify and address, so you’ll want to be sure to record any errors your participants made throughout the test.

* Write down any errors participants make when trying to complete the tasks.
* Once you begin organizing your information, you’ll want to classify the errors according to severity (more on this in the next section). If you’d like to start doing this in advance, you could also do so here.
* Note why you think your participant made a particular error (or, even better, note down why they said they made a particular error).

1. 5. Do “The Big Sort”

* Put all the concepts you’ve written down into their appropriate categories (**Observations**, **Positive Quotes**, **Negative Quotes**, and **Errors**). If you find it helpful, you can further break down “Observations” into what the participant was “Thinking,” “Doing,” and “Feeling.”
* If there are duplicate observations for multiple participants, feel free to combine the concepts by grouping the notes together or combining them into a single note, but be sure to make note of how many occurrences of the observation or error there were and which participant(s) you’re referencing.
* You should end up with 20-25 core observations that you’d like to follow up on. If you have fewer than that number, try reviewing the interviews again or check your notes from the sessions. If you have more than this, try combining related concepts as noted above.

Throughout the process, feel free to rip up sticky notes and/or erase and rewrite observations. This process is all about extracting observations and recording them in a way that makes the most sense to you and your team. It’s meant to be messy. It’s meant to be fun. Fundamentally, the goal is to get the thoughts and behaviors of all the participants onto the same page so we can tease out meaningful patterns.

1. The Rainbow Spreadsheet

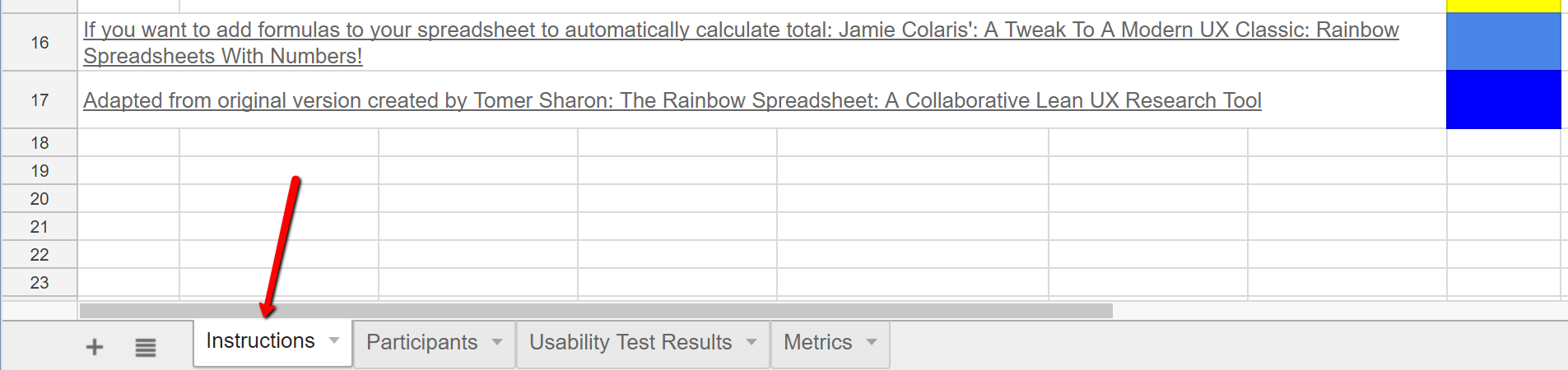
At this point, you should have a big mess of observations sorted into categories. Now, it’s time to make even more sense of this mess and organize your observations in a way that will be useful for other stakeholders to understand (and that will help you make decisions about next steps, as well). As you do more usability tests, feel free to explore other techniques or develop your own. Each project is different, and everyone has a unique way of weaving together patterns to create meaning.

The **Rainbow Spreadsheet**, originally created by UX Researcher and author [Tomer Sharon](https://www.smashingmagazine.com/2013/04/rainbow-spreadsheet-collaborative-ux-research-tool/) and since adapted by many designers, is a great way to organize information from your usability tests and create visual patterns. Though it may look a little intimidating at first, it’s a great way to keep track of your information, especially when you have more than a handful of participants. Here’s a step-by-step guide on how to use the spreadsheet to organize the information from your usability tests and subsequent affinity map. Again, you may wish to read all the way through the steps first, then come back and try to complete the steps with your own data on a second readthrough.

1. 1. Make a copy of the master Rainbow Spreadsheet.

You can find a [ready-made template for the Rainbow Spreadsheet](https://docs.google.com/spreadsheets/d/1_0U6y3fX6ejuTB37TEyzaA-WBmphjpxfKkB2mxeS2kU/edit?usp=sharing). Copy this template into your own Google Drive or download it as an Excel document and use it as a starting point to make your own.

The spreadsheet can look a little overwhelming at first, so let’s take a short tour. You’ll see that there are four tabs across the bottom of the sheet—“Instructions,” “Participants,” “Usability Test Results,” and “Metrics.” We’ll walk through how to use each of these tabs in the following steps.



The tab you’re starting on, “Instructions,” is, as the name suggests, a set of basic directions. It also has a place for you to put your name and date, as well as the name of your usability test from your test plan. Now, let’s move on to the next tab.

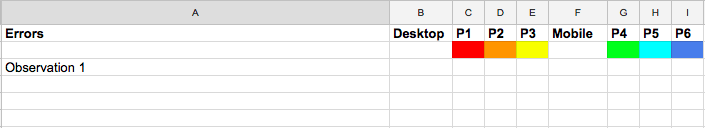
1. 2. Enter participant info under the “Participants Tab.”

Remember, this information is sensitive and should only remain visible to you and other stakeholders. You may have specified how you’ll be using your participants’ information in your informed consent form. Don’t forget to revisit that. If you’re going to share results with others, you may wish to remove the contact information, use pseudonyms, or remove this tab altogether in the shareable version of the spreadsheet. For now, since this information is for your version of the spreadsheet, go ahead and fill in the details. This page is for your records and can be helpful when looking for similarities in participant characteristics and the types of observations you made.



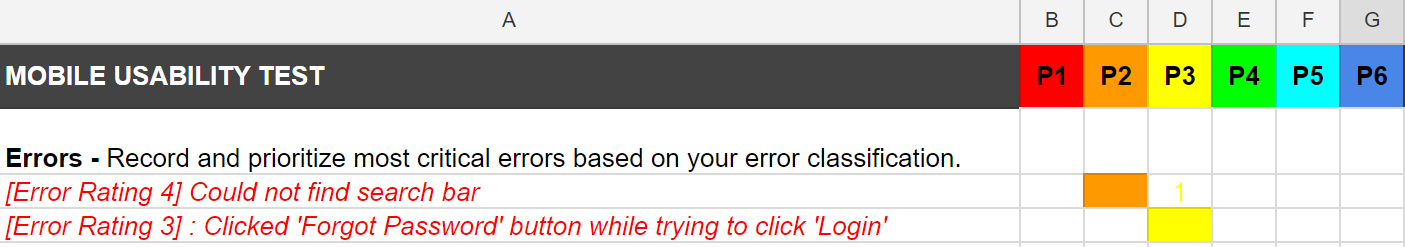
1. 3. Assign each participant a color and customize the number of participants to match your study.

You should have conducted your usability tests with at least six participants (three on desktop and three on mobile). Each participant is assigned a color. Feel free to customize the colors to your heart’s content—just make sure the shades are different enough that you can tell them apart. You’ll use these colors in the “Participants” tab, as well as in the “Usability Test Results” tab. If you colored your participants similar to the example above, here’s an example of how your Usability Test Results sheet might look:

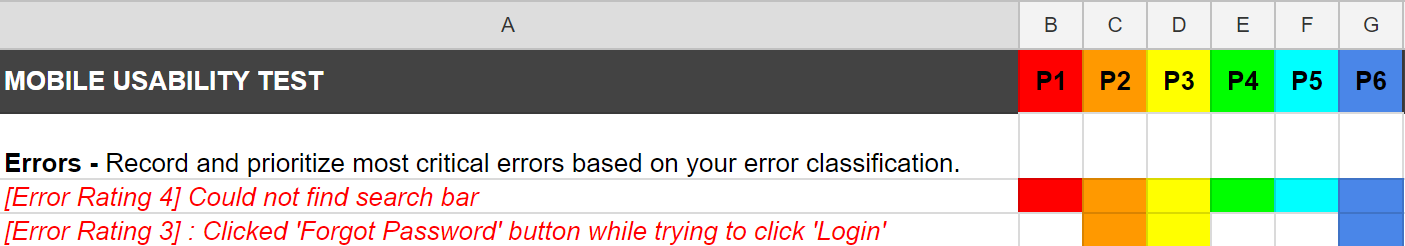


1. 4. Enter and organize your affinity map observations into the “Usability Test Results” tab.

For each observation, fill color in the appropriate box. For example, if Participant 1 (P1) made observation 1, Participant 2 (P2) made Observation 3, and Participant 3 (P3) made Observation 2, it would look like this:

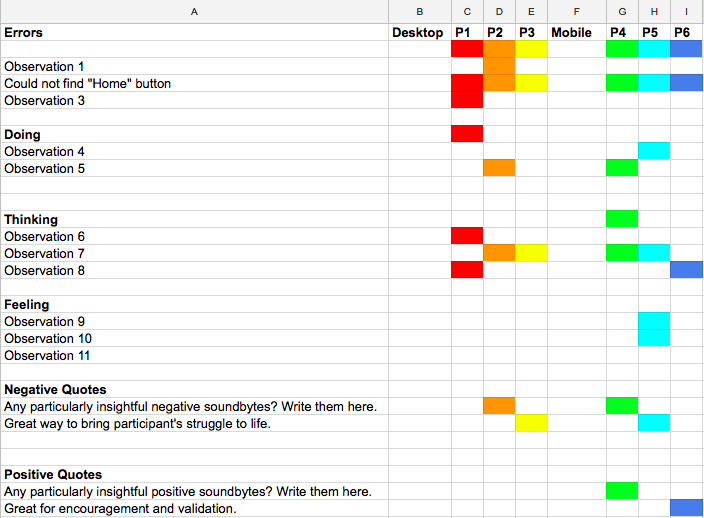


If more than one participant made the same observation, you simply fill in the appropriate colors for each person relevant to the observation, like so:



As mentioned above, you should end up with 20-25 core observations that you’d like to follow up on. If you have fewer than that number, try reviewing the interviews again or check your notes from the sessions. If you have more than this, try combining related concepts as noted above. Use the “Notes” column to the far right of this sheet to record anything of interest that doesn’t fit in the concise observation description.

Once you’ve filled in all your observations, your sheet will look similar to the below diagram. The visual blocks make it easier to identify patterns. For example, Observation 2 (Could not find “Home” button) is filled in for all six participants across desktop and mobile. This is a strong indication that your design is not allowing people to easily go “Home.”



Add up the total number of times an error, observation, or quote occurs (how many colored boxes are filled in next to it) and fill in the “Total” column.



1. 5. Classify your errors.

Phew! You’ve already done a lot of work, but it’s not over yet. It’s now time to rate the errors that came up in your test. As discussed in previous Exercises, one framework for examining usability is Jakob Nielsen’s Five Components on Usability, which covers learnability, efficiency, memorability, satisfaction, and errors.

Some of the components require testing multiple times (such as memorability, which looks at whether your users are able to reestablish proficiency with your product after a period of time away) or can be tricky to measure (such as satisfaction). For this reason, we’ll be focusing most closely on errors for the purpose of your current test. If you completed the bonus task as part of the test script Exercise, you can also follow up on the results of your satisfaction questions and use the “Metrics” tab to classify your results. Otherwise, let’s simply focus on classifying your errors.

In isolated usability tests, errors are of particular interest as they’re easy to catch and important to fix. Here are some things to consider when classifying errors:

* How many errors do users make? What’s the frequency of errors?
* How severe are the errors?
* How easily can users recover from the errors?
* How much does an error compromise your ideal user flow for completing the specified important tasks?

There are various ways to categorize and rank the severity of errors, but, for consistency and clarity’s sake, we’ll refer to an adapted version of Jakob Nielsen’s four-step rating scale:

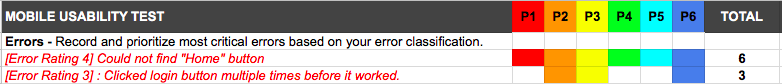
* 0 = I don't agree that this is a usability problem at all
* 1 = **Cosmetic problem only:** need not be fixed unless extra time is available on project
* 2 = **Minor usability problem:** fixing this should be given low priority
* 3 = **Major usability problem:** important to fix and should be given high priority
* 4 = **Usability catastrophe:** imperative to fix before product can be released

Cosmetic and minor problems can create disquietude in the look, feel, and structure of an app, but don’t hinder test participants from completing prescribed tasks. If participants continually struggle with the same task, it’s likely there’s a major or catastrophic error, such as an issue with navigation that prevents users from locating important information or features.

You already know the **frequency** of errors due to the number of colored blocks next to each error observation. Now, it’s time to rate the **severity**. Frequency and severity will help you prioritize what to fix.

As you can see, an “Error Rating” is assigned at the beginning of each of the errors you noted in your sheet. Prioritize the most critical errors at the top of your document.

Let’s look at the observation we made regarding the “Home” button in the screenshot of the example Rainbow spreadsheet below. All six participants reported not being able to find the “Home” button, making it a high-frequency issue, and it also causes significant usability problems (versus a cosmetic issue). For this reason, we’ve ranked that observed error as a “4” in severity. In contrast, three participants, had difficulties clicking the login button. As a “Error Rating 3”, this is a “major usability problem” as it’s highly inconvenient, but not necessarily a “usability catastrophe” as the participant was still able to proceed after some effort.



When you are done, you may wish to sort your errors based on the error rating, with the highest ratings at the top.

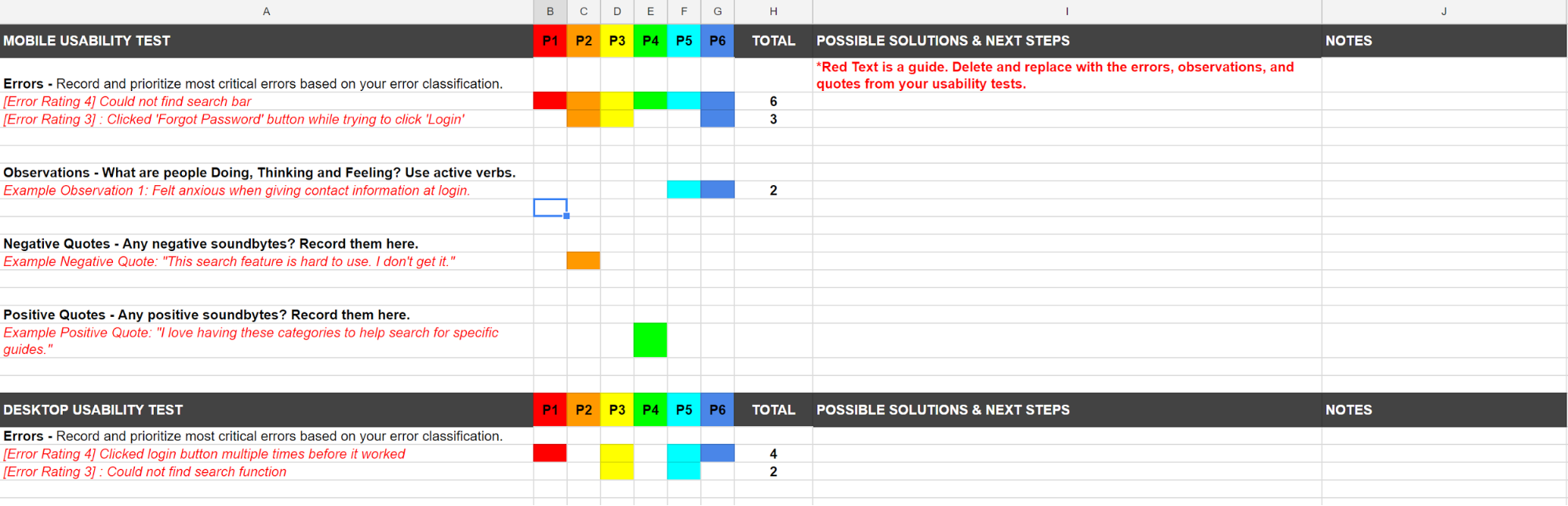
OTHER RATING SCALES  
Usability researchers have come up with many different scales to rank and prioritize usability issues and errors. For a more comprehensive list and explanation, check out Jeff Sauro’s “[Rating The Severity Of Usability Problems](https://measuringu.com/rating-severity/)." If you feel inspired, feel free to use a different scale or come up with one of your own!

As a final note, we want you to enjoy the process of downloading and organizing information. Think of it as a way of arranging your participants’ thoughts with the intent of bringing meaning to you and your team. The end goal is not to create a colorful spreadsheet, but to create something meaningful that can help you make better design decisions.

”UX research is not about producing reports. It’s about answering people’s questions and helping the organization develop empathy for its customers. Deliverables and reports are means to that end…If it doesn’t work for you, then customize it to your team’s needs or develop your own tools. Writing a report for a week will not serve your team’s needs very well. Find creative ways to learn about users together!”  
TOMER SHARON

1. What’s Next? Proposing Possible Solutions

The last step of our spreadsheet helps us elicit solutions from our newly organized data. Use the “Possible Solutions” column to jot down some fixes that will help influence future iterations of your design and prototype. In particular, focus on the errors you indicated as most severe and on observations that applied to more than one participant. You don’t need to have a perfect solution to everything yet, but you should have a good idea of what you’d do to fix an issue (or perhaps do further testing).



While you think about solutions, don’t forget to reference your test plan and test goals. What were the main features and functionality you needed for the app to be considered successful, and what were the business needs? These can help provide focus when it comes to deciding what to address and how.

1. Summary

One of the most critical parts of running a usability test is figuring out what to do with all your results. As with most other parts of UX, there are many ways to tackle this challenge depending on your particular project or situation. One of the easiest ways to sort and classify results is via an affinity map. Affinity mapping helps you condense quotes and observations into concise pieces of information that can then be grouped together to determine patterns.

Once you have your basic information and patterns identified, the information can be processed using the Rainbow Spreadsheet. The purpose of this sheet is to help you classify important observations and errors by presenting your results in a visually organized fashion. This is particularly useful if you’re running a large usability test with more than 10 participants. As you might imagine, you’d end up with copious amounts of data, which would be hard to sort using only a single document (or even just an affinity map). In the next Exercise, we’ll use what you’ve learned through this complete process to inform your stakeholders of the results of your tests, as well as the next steps for your prototype.

1. Resources

* [A Guide to Carrying Out Usability Reviews](http://www.uxforthemasses.com/usability-reviews/)
* [Sensible](http://www.sensible.com/)
* [Usability Testing Checklists](http://sensible.com/Downloads/checklists.pdf)
* [A Framework for Preparing Your Own Remote User Test Script](http://www.userzoom.co.uk/blog/ux-analytics-testing-tools/framework-preparing-remote-user-test-script/#content-read=true)
* [12 Expert Tips for Moderating Like a Pro](https://www.usertesting.com/blog/2014/09/22/12-expert-tips-for-moderating-like-a-pro/)
* [A Guide To Simple And Painless Mobile User Testing](https://www.smashingmagazine.com/2015/12/simple-and-painless-mobile-user-testing/)
* [Reflector](http://www.airsquirrels.com/reflector/)
* [14 MAR 2014 Recording Mobile Device Usability Testing Sessions – Guerrilla Style](https://www.thoughtworks.com/de/insights/blog/recording-mobile-device-usability-testing-sessions-%E2%80%93-guerrilla-style)
* [10 Usability Heuristics for User Interface Design](https://www.nngroup.com/articles/ten-usability-heuristics/)

You scored 5/5 on this Exercise last time.

Take Quiz AgainSee Details

1. Task

* [DIRECTIONS](https://careerfoundry.com/en/course/become-a-ux-designer/exercise/usability-testing-results#directions)
* [SUBMISSION HISTORY](https://careerfoundry.com/en/course/become-a-ux-designer/exercise/usability-testing-results#step_submission_history)

 Estimated Task Time: 5 Hours.

Remember all the data you obtained from your usability tests? It’s now time to analyze it! Download the audio and video data from your test sessions and organize the information using an affinity map. Then, organize your errors, observations, positive quotes, and negative quotes using the Rainbow Spreadsheet. Analyze the results to come up with possible solutions and changes to improve the next iteration of your project prototype.

**Directions**

1. Following the directions in the Exercise, create an affinity map using the audio and video recordings of your usability tests. Once you’re finished, take a picture of your affinity map to submit along with the Task (or, if you made a digital one, take a screenshot or share your file).
2. Complete your own Rainbow Spreadsheet following the directions above, inputting errors, observations, negative quotes, and positive quotes (go [here](https://docs.google.com/spreadsheets/d/1_0U6y3fX6ejuTB37TEyzaA-WBmphjpxfKkB2mxeS2kU/edit?usp=sharing) for a template if you haven't already). Make sure that you classify all of your errors using the error rating scale from the Exercise. You should come out with 20-25 items (including observations, errors, and quotes) to focus on.
3. Propose “Possible Solutions and Next Steps.”
   * After you’ve recorded your results and prioritized your errors, observations, and quotes, make hypotheses on how to improve your prototype in the “Possible Solutions and Next Steps” column next to each of your 20-25 items.
   * You’ll be thinking more about possible next steps in the next Exercise, so don’t worry about getting them perfect here. Simply use this chance to brainstorm and come up with ideas.
4. Upload your affinity map (either in a digital document or as a photo in a PDF), as well as your completed Rainbow Spreadsheet here. Feel free to share additional thoughts or ask questions in the submission box.

**Bonus Task**

If you did the previous Bonus Tasks and included other metrics beyond error metrics in your usability test, now is your chance to analyze them. Customize the “Metrics” tab with the other important metrics you chose—Success Rate, Time on Task, and/or Satisfaction. You can always go back and record elements such as “Time on Task” by watching the video. This is a great way to benchmark metrics over time across multiple participants. If you decided to use a follow-up survey during your usability test, such as Net Promoter Score (NPS), you could customize the “Metrics” tab to include this information, as well.

Bottom of Form

Top of Form