

The People Nerd's Guide to

Usability Testing for UX Designers

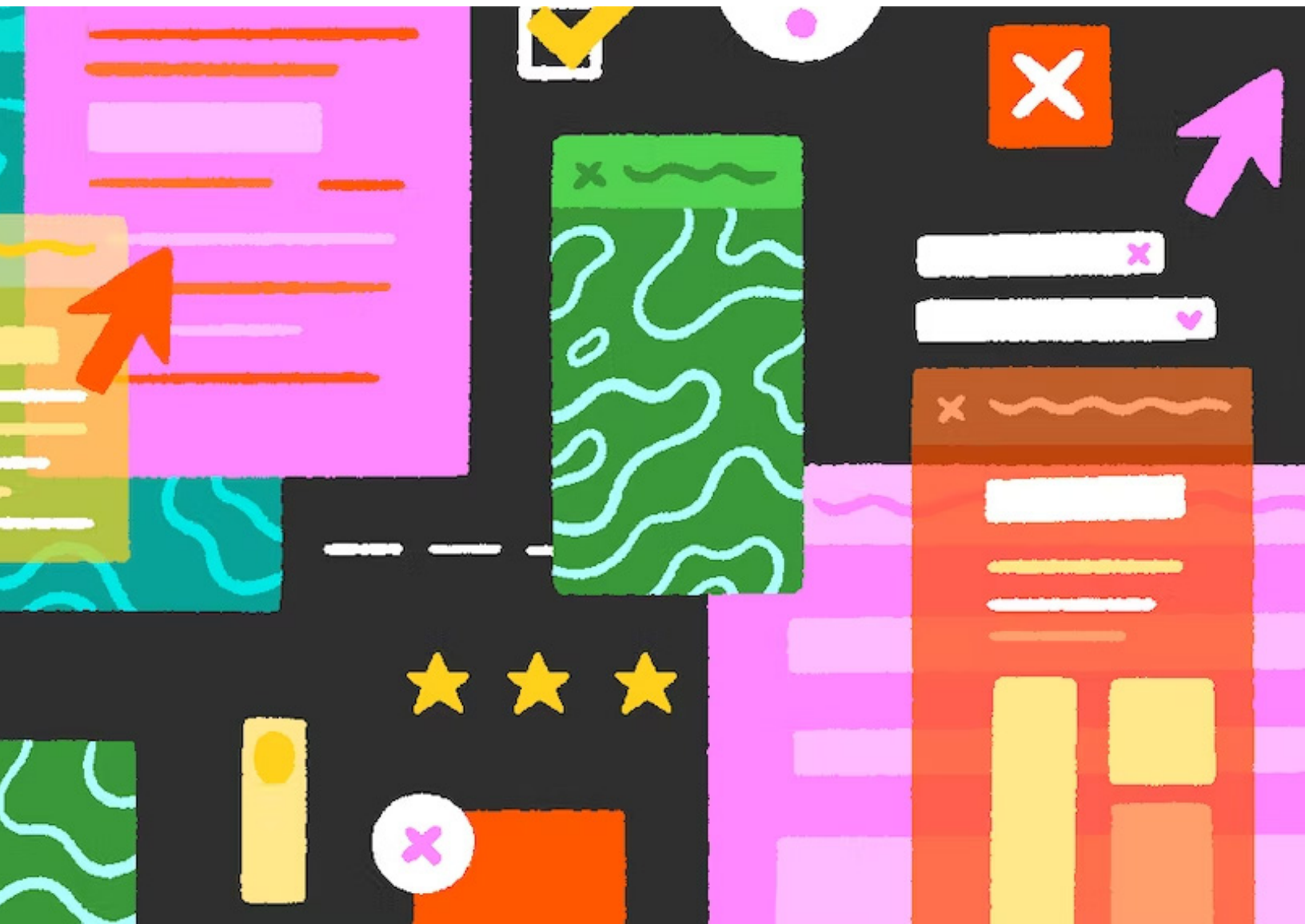


TABLE OF CONTENTS

- 3** Introduction:
How to use this guide

- 4** Chapter 1:
Planning and scoping

- 17** Chapter 2:
Become a usability task writing pro

- 31** Chapter 3:
Analysis and shareout

- 46** Chapter 4:
How Dropbox approaches usability

INTRODUCTION

How to use this guide

Can someone on our website subscribe to our newsletter?

Can our customers easily add an item to their cart?

Are there any bugs on our platform that we aren't seeing?

Was the task we asked them to complete confusing?

How long did it take?

The answer? Usability testing.

Usability testing is one of the staple methods in a user researcher and UX designer's toolkit. It can support a multitude of organizational goals, from learning how users interact with the product to surfacing where bugs and complicated friction points are hiding.

When set up correctly, usability testing can reveal invaluable insights. When conducted poorly, it can result in more confusion and potentially costly product updates.

In this guide, we explore what exactly “usability” means and what “usability testing” is, along with the goals they can (and can't) help you accomplish. That includes how to set up these projects and write tasks, and how to analyze and share your findings.

Whether you work in the design, product, engineering, or are someone eager to get involved in the research process—take this guide and approach usability testing with confidence!

CHAPTER 1

Planning and scoping

Usability testing is a fundamental part of the product development process and a critical skill for anyone looking to evaluate an idea or prototype.

Although usability testing is one of the first research methods many people learn, it takes a lot of time and practice to nail the skill down. From planning to writing tasks to [analyzing data](#), usability testing is a fantastic method to have in your toolkit. It can help your team move forward and get “unstuck” in certain situations.

On the flip side, usability testing can be overused or used in the wrong scenarios. It’s incredibly important that, when planning, we ensure [the methods we choose](#) are the best for getting the information we need at the end of the study.

There are a few times I went about this incorrectly, rushing to conduct a usability test when it wasn’t the correct [method for the goals](#) my team was trying to achieve. Because they are so easy (or seem to be), we can get stuck saying, “Let’s run a usability test,” without being sure that’s the best approach.

To take away that uncertainty that I once felt, let’s go through the end-to-end planning process of a usability test to ensure when you pick it, it’s the best method for what you need!



What is usability and usability testing?

Usability is the ability for someone to:

- Access your product or service
- Complete the tasks they need
- Achieve whatever goal or outcome they expect

For example, with a clothing company, a user would go to the website with the expectation of being able to purchase clothing. But just because a user can buy clothing doesn't mean it's an easy or satisfactory experience.

So we can break usability down further into three areas:

Effectiveness

Whether a user can accurately complete tasks and an overarching goal

Efficiency

How much effort and time it takes for the user to complete tasks and an overarching goal accurately

Satisfaction

How comfortable and satisfied a user is with completing the tasks and goal

Usability testing—whether you use metrics or run a [qualitative usability test](#) (more on that later)—looks at these three factors to determine whether or not a product or service is usable.

So, what about those who browse e-commerce out of boredom (and no purchase intent)?

This leads to the next question...

What are we testing? And how do we know?

As I mentioned above, there are many aspects you could test (even with a simple product!). But the point of a usability test is to ensure that users can complete their most common tasks to achieve their goals.

Again, one of a clothing website's main goals is to purchase clothing. You can have smaller goals within that larger goal, such as comparing clothing or buying a gift for someone.

Then, you can break each larger goal into the tasks people must do to achieve those goals.

For instance:

- Searching for specific types of clothing with keywords
- Filtering through colors, brands, sizes
- Sorting by reviews, prices
- Opening multiple windows to compare different options
- Saving clothing to a favorites list
- Reading (and understanding) the size and fit of clothes
- Adding a piece of clothing to a basket
- Checking out and paying for the clothing
- Receiving a confirmation of purchase
- Receiving the clothing

These are all tasks associated with the larger goal of purchasing clothing. With [usability testing](#), we ask people to do these critical tasks to assess whether or not they can achieve them and the larger goal efficiently, effectively, and satisfactorily. We also get their feedback on the experience we put them through.

If someone can do these tasks, they can achieve their expected outcome. However, if efficiency, effectiveness, or satisfaction suffer during this process, they may get frustrated and give up or go to a different website.

We've all encountered this—an infuriating user experience that made us rage-click, throw our phones (against a soft surface, of course), and give up on a product or service.

This is why usability is so vital. It can give us a clear understanding of how our product is aligning (or not aligning) with users' mental models of the experience.

Qualitative versus quantitative usability testing

There is an essential distinction between qualitative and quantitative usability testing. I put usability testing into one category for a while, rather than thinking through which was most applicable. I also sometimes tried to run a hybrid test, which made for an inefficient use of the method.

Qualitative usability testing is looking to get actual feedback from the user on the idea, prototype, or flow. For this, you show participants screens or images and ask them to provide qualitative feedback to understand how they perceive what's in front of them. This route means the session will be a discussion or conversation.

Conversely, **quantitative usability testing** strictly assesses the above three cornerstones of usability: effectiveness, efficiency, and satisfaction. Through this session, we evaluate how participants move through the experience we put in front of them [through metrics](#). In this session, you aren't asking for qualitative feedback (as that would skew the metrics). Instead, you're trying to mimic a real-life experience as much as possible.

Is usability testing the correct method for your study?

Knowing what a method is and when to apply it are two distinct things. As I mentioned above, I knew what usability testing was.

I was reasonably confident in my skills, so I overused the method for studies that would have been better off with a different approach. This happens quite often—we get comfortable with and stick with a method, even when it may not be ideal.

However, there are ways to determine if usability testing is the proper method for your study. I do this by reverse engineering the information we need by the end of the study. Start with research goals by [talking to stakeholders](#) about the questions they need answered by the study.

Creating research goals for usability testing

Before we dive into the main goals for usability tests, it's great to understand where usability testing shines and what it can help us understand:

- How users perceive a product
- How well a user can use a product for its intended function
- How well a product allows (or doesn't allow) a user to reach their goals
- How a user uses a product, separate from how we think a user should use a product
- How a product functions when placed in front of a human
- Where bugs and complicated user experiences lie within a product

On the flip side, usability testing will NOT tell us:

- The emotions a user is feeling outside of their immediate actions (instead, try [generative research](#))
- Statistically significant quantitative data on usage patterns or trends (instead, try [product analytics](#))
- Preferences between two versions of a design (instead, try [A/B testing](#) or [comparative usability testing](#))
- The desirability of a product (instead, try [market research](#))
- Complete market demand and value (instead, try [market research](#))
- What people will pay for a product (instead, try [pricing research](#))

Reverse engineering goals

As mentioned above, I love to reverse engineer my research goals by basing them on my stakeholders' questions and the information they need to make a clear and straightforward user-centric decision moving forward.

I ask my stakeholders questions when they come to me with a particular research request. Some of these questions include:

- What type of information do you need at the end of the project?
- What decisions do you want to be able to make?
- What are the top three questions you need answered?
- What is your definition of success for this project?

Another way to do this that I've used is to have them fill out this mad lib:

I need [X information] to answer [Y question] to make [Z decision] by [timeline].

Once you have this information, it's easy to start creating research goals that help you determine the best methodology. I do this by writing these in a [research plan](#), which is a great way to keep everyone aligned during the research project.

I will split the goals up by qualitative and quantitative usability testing.

Qualitative usability testing goals

Taking the above questions, let's say that our stakeholders gave us the following answers about the information they need:

- To understand how participants react to and perceive the prototype/experience
- To understand how well our idea aligns with their mental models or expectations
- To get some feedback on how we could improve the prototype/experience

And about the decisions they are trying to make at the end:

- To make iterations on the experience so that it better aligns with users' mental models
- To improve the flow/experience based on feedback
- To feel more confident they are going in the right direction with the experience/flow

With this, it's clear that they want qualitative feedback from participants to better understand the current experience and how they can improve it. We seek [qualitative feedback](#) whenever we want to know how people react, feel, and perceive.

Additionally, the team wants to improve the experience so qualitative feedback would be hugely helpful.

For this, I would say the goals are:

- Discover participants' reactions to and perceptions of the current experience
- Learn about participants' current pain points, frustrations, and barriers to the prototype/experience
- Evaluate how participants work through the prototype/flow and their feedback

Quantitative usability testing goals

Instead, let's say our stakeholders answered that they wanted the following information from the study:

- To measure how effective and efficient the product is
- To evaluate how the product works when put in front of a human
- To benchmark this current experience against a future one

And they wanted to make the following decisions:

- To improve the efficiency and effectiveness of the product
- To make any final changes before shipping the product

With this, we are looking at concrete information regarding effectiveness and efficiency. The best way to get this type of information is through quantitative testing. With this information, qualitative feedback wouldn't give us the necessary results. Quantitative usability testing would enable us to measure the metrics and improve them by the end of the study.

A note on fidelity and order

For quite some time, I made quite a significant mistake with my quantitative usability tests. I tested [low-fidelity prototypes](#) with quantitative usability testing.

Low-fidelity prototypes often have a "happy path." The point of making a low-fidelity prototype is not to design everything, including the "unhappy path." With this, when it came to measuring the most common metrics, such as time on task or task success, I couldn't accurately measure them.

If there was only one path, how would I know if someone might be unsuccessful? How would I know how lost someone would get? Or how much time would it take them if there were all the distractions there typically are on an interface?

Low-fidelity prototypes aren't ideal for quantitative usability tests because they don't mimic real life in the way high-fidelity or live products are better suited to. There might not be opportunities for participants to fail or get lost.

With that in mind, I always recommend using qualitative usability testing with low fidelity so you can hone it into a high-fidelity or [working prototype](#) that you then test with a quantitative usability test.

Picking metrics (for quantitative usability tests)

Since quantitative usability tests require metrics to measure, it's best to pick those earlier on in the process. I base these metrics on the three cornerstones of usability we already mentioned:

Effectiveness

Whether or not a user can accurately complete a task that allows them to achieve their goals. Can a user complete a task? Can a user complete a task without making errors?

Efficiency

The amount of cognitive resources it takes for a user to complete tasks. How long does it take a user to complete a task? Do users have to expend much mental energy when completing a task?

Satisfaction

The comfort and acceptability of a given website/app/product. Is the customer satisfied with the task?

Combining these metrics can help you highlight high-priority problem areas. For example, suppose participants respond confidently that they completed a task, yet most fail.

In that case, there is a vast discrepancy in how participants use the product, leading to problems. Let's break up the metrics by area of usability testing:

Effectiveness

Task success

This simple metric tells you if a user can complete a given task (0 = Fail, 1 = Pass). You can get fancier with this by assigning more numbers that denote users' difficulty with the task, but you need to determine the levels with your team before the study.

The number of errors

This task gives you the number of errors a user committed while trying to complete a task. You can also gain insight into common mistakes users encounter while attempting to complete the task. If any of your users want to complete a task differently, a common trend of errors may occur.

Single Ease Question (SEQ)

The [SEQ](#) is one question (on a seven-point scale) measuring the participant's perceived task ease. Ask the SEQ after each completed (or failed) task.

Confidence

Confidence is a seven-point scale that asks users to rate how confident they were that they completed the task successfully.

Efficiency

Time on task

This metric measures how long participants can complete or fail a task. This metric can give you a few different options to report on, where you can provide the data on average task completion time, average task failure time, or overall average task time (of both completed and failed tasks)

Subjective Mental Effort Question (SMEQ)

The SMEQ allows the users to rate how mentally tricky a task was to complete.

Satisfaction

System Usability Scale (SUS)

The [SUS](#) has become an industry standard and measures the perceived usability of user experience. Because of its popularity, you can reference published statistics (for example, the average SUS score is 68).

Usability Metric for User Experience (UMUX or UMUX-Lite)

The UMUX/UMUX-Lite is the “newer” version of the SUS and measures the experience's perceived usability. The advantage is the UMUX is only four items, and the UMUX-Lite is only two, making it easier for participants to take.

When choosing metrics for a quantitative usability test, especially if it is moderated (more on that in a bit), make sure not to pick so many that it is difficult to manage. The first time I conducted a quantitative usability test, I got very excited. I decided to measure time on task, task success, confidence, the SUS, and the number of errors.

It was way too much for me to properly keep track of during the test, even though I had a notetaker helping me. Of course, if you're using an unmoderated tool, some metrics—typically task success and time on task—are already recorded, so you don't have to worry about tracking them.

But also keep in mind, when choosing metrics, the number of surveys or items you're asking your participant to respond to. For instance, in one study I ran once asked after each task:

- Single ease questionnaire
- Confidence
- Subjective mental ease questionnaire

And then, on top of that, I asked the SUS after the end of the test, as well as some general questions on satisfaction. All of these measures were a lot to ask the participants, and I felt them fatiguing after a few tasks.

Now, I typically choose the following metrics for each task:

- Time on task
- task success
- SEQ

And for the end of the test, I use:

- UMUX or UMUX-Lite

If we measure particularly complex tasks, I'll also ask about confidence because I wonder if people believe they're completing tasks correctly. Additionally, if we're curious about potential unhappy paths, I add the number of errors and then report on the most common unhappy paths people take.

However, when it comes to quantitative usability testing, I always operate by keeping it as simple as possible. I also always talk with my stakeholders for their input on the decision.

Who should you talk to? And how many of them?

The next step when planning a usability test is considering which participants to talk to. Recruitment can make or break a study, so I recommend going into this step as thoughtfully as possible. Trust me, it's challenging to fill a 60-minute usability test with the wrong participants, plus it's awkward.

The best way to do this is through a [screeener survey](#). You use these short surveys to qualify participants and ensure you get the best fit for the information you need.

One big mistake I used to make was only asking for demographic information in my screener surveys. Since I came from an academic background, demographic information was a must in many projects. I didn't focus on behavior or habits in those projects.

Unfortunately, when I just asked for demographic information (ex: gender, age, location), I landed in a terrible situation: the participants fit the demographic data, but they couldn't give me the information I needed.

For example, I was investigating the experience of a new flow for people to better estimate their jeans size before buying. Unfortunately, I focused too much on demographics instead of habits and recruited people who didn't purchase jeans online or had no problem estimating their sizes.

Once this happened (a few times), I realized that it was essential to write good screener surveys for a few different reasons:

- Finding and talking to the most relevant participants with the characteristics, habits, and behaviors you need to understand better.
- Hitting the [correct sample size](#) by segmenting your participants into different buckets, you can later create meaningful deliverables that lead to action.
- Ensuring the return on investment for your research will be as high as possible and avoiding wasted time and money on suboptimal participants.
- Avoiding burning out your participant list by asking all users to participate constantly.

For usability tests, I ask myself the following questions to help understand the screener questions I should ask:

- What are the particular behaviors I am looking for?
- Have they needed to use the product? And in what timeframe?
- What goals are essential to our users?
- What habits might they have?

So, if we were to redo the failed screener about better estimating your size when purchasing jeans online, I might need to hone in on the following criteria:

- People who have purchased jeans online in the past three months
- People who have struggled with jean sizes online in the past three months
- People who have returned jeans after buying the wrong size

By targeting these criteria, I would have been better able to test the flow with someone who had these painful experiences in the past, getting me much more valid and reliable data.

How many people?

One note about the [sample size](#) for usability testing: A general idea for evaluative research is testing with five people. When I review research plans, I often see five participants as the number for an evaluative study.

While this could be correct, it isn't a hard and fast rule. Like the above, if you [pick five random people](#) off the street to test your product, you likely won't find 85% of the issues.

The fine print behind five people per usability study is that it means five users per segment. So always make sure you're thinking about segmentation and aren't asking five completely different and random participants.

Session logistics

There are a few other areas to cover when scoping and planning your usability test:

Unmoderated versus moderated usability test

I typically default to moderated usability tests because that's what I grew up on; however, unmoderated testing has some immense benefits.

Typically, I use moderated testing for qualitative usability tests because it allows me to discuss better and dig deeper with the participant. For instance, if participants struggle with certain areas of the experience, I can understand why in a moderated test. It allows me to converse with participants that can give the team more direction with additional and more profound feedback.

However, with a quantitative usability test, you aren't looking for this depth of feedback or to understand why people are struggling. Instead, you are looking specifically to measure your chosen metrics. With this, unmoderated testing can be hugely beneficial by giving quicker responses and a larger sample size.

Session length

Regarding the length of usability tests, I have seen them run anywhere from 15 to 90 minutes, depending on the complexity of the tasks and the type of information you are trying to get.

For instance, when it comes to unmoderated quantitative usability tests, you could be looking at a session length of 15-20 minutes since you are strictly measuring the metrics and not going deep into qualitative feedback. Going toward a moderated qualitative test might run closer to 45-60 minutes because you need time to ask follow-up questions and discuss these with the participant. I always default to 60 minutes during a moderated qualitative usability test because it's better to end early than to run late, and this timing gives me a chance to understand the "why's" behind people's feedback. For unmoderated quantitative tests, it depends on the number of tasks (of which I recommend no more than 10, if possible), but I usually choose 20-30 minutes.

And finally, for complex and moderated quantitative tests, I default to 45-60 minutes to go through the necessary tasks and measurements.

Now that you're familiar with what usability testing is (and isn't)—and got a first look at how to prepare for one—we'll dive into the nitty gritty of creating an effective usability test, task writing.

CHAPTER 2

Become a usability task writing pro

Let's quickly see if people can "do stuff" on our product. All you have to do is write some tasks. Easy peasy...right?

Usability testing and task writing can feel deceptively easy. Trust me. I've fallen into the trap of quickly [writing usability tasks](#) that confused participants and led to useless data that was entirely unhelpful to the team.

There is an art and science to crafting usability testing tasks that ensure your participants get the necessary information, and your team gets the information they need to make better decisions.

In this step-by-step guide, we'll review all the components that go into writing an excellent usability testing task and how to do it.



What is a task?

Before diving into the how-tos, paying attention to something often overlooked when it comes to usability testing is essential.

What is a usability testing task? We use the word as if we know exactly what it means. And, while task is a common word, what does it mean in the context of usability testing?

A usability testing task is a series of steps the user has to perform to accomplish a given goal.

This definition is hugely important to pay attention to because one of the biggest mistakes I see in usability tests is letting a participant land on a page and explore it with no goal.

It's rare for people to go to a product or service with no goal. How often have you gone to a website homepage just to explore it? How frequently have you visited a product or service with no goal?

When we don't have clear goals for our [usability test tasks](#), even if they are super open, we end up with data that can be biased and unreliable. The information you get from participants might not be realistic for their real-world experiences.

So, our tasks have to have relevant goals.

How to select the tasks

Usability testing is about observing and sometimes measuring participants completing tasks on a product or service. To get the most relevant data, your tasks should be as close to real-world experience and usage as possible.

For example, if I love to read and am looking for a new book and land on your product, I'm likely there to browse and purchase a book. So, suppose you ask me to do some auxiliary tasks like finding and signing up for your newsletter, finding careers available at your bookstore, or looking up your special bookstore credit card. In that case, you won't get the most essential data from me.

Sure, there might be a use case for those tasks, but depending on [the persona](#) (or group of users) you are optimizing your experience for, you must pick tasks carefully.

How do you pick the best tasks for your usability test?

Define who you're testing

The first step in picking the right usability tasks is knowing who your users are. Like the example above, if I came to your website to find and purchase a book, you would want to ensure you tested my particular flow. However, if I were seeking a career at a bookstore, I would have a completely different use case.

If you have different [personas or groups of users](#), pick who you're targeting with your usability test, because it will allow you to create tasks tailored to their specific needs and goals.

List the users' top goals and tasks

Based on who you're testing, list the potential tasks your chosen user group might perform while interacting with your product or service.

The best way to create this list is by [triangulating data](#) from:

- Analyzing the typical user flows within your product to see the most common pathways users take
- Analytics that show usage data of different features or areas of your product or service

- Customer support tickets that include common complaints or problematic flows your users encounter
- Previous research that demonstrates the users' goals for your product or service

Using the book website as an example, you might list out tasks such as:

- Searching for a specific book
- Browsing by genre
- Reading book reviews
- Filtering and sorting book results
- Adding books to your wishlist
- Adding books to your cart
- Purchasing a book
- Tracking a book order
- Purchasing a gift voucher
- Giving a review of a book
- Using a book recommendation tool
- Finding a store

Prioritize the task list

As you can see, there are many tasks on the above list. Not all tasks are equally important, so defining who you're talking to before writing your usability testing tasks is essential.

When picking tasks, I recommend choosing the critical tasks for users to complete that get them to achieve their primary goal.

So, if my main goal is to come on to your website and purchase a book, then the essential tasks for me would include:

- Searching for a specific book
- Adding the book to my cart
- Purchasing the book

Depending on if I knew which book I wanted to buy, you could also have me browse by genre or popularity.

In terms of how many tasks you should have in your usability test, it depends on how long you have with your user and how complex the tasks are. If you have 60-90 minutes with a user for a [moderated test](#), aim for 10-15 tasks. If you have less time, you'll have to cut the number of tasks.

How to write a task

Now that we understand how to pick the top tasks we want the user to go through, it's time to write them!

Before we get into writing the tasks, let's look at what we should avoid when writing our usability testing tasks.

I see a few recurring mistakes during many usability tests I have observed. I have to actively tell myself to avoid these mistakes (which is where practice comes in) as they are easy to slip into.

Task writing mistakes to avoid

Using words in the interface

If you're trying to get users to do an action and those words are in your interface, don't include them in your task. Using words in your interface makes tasks easier for participants as it leads them to the correct answer.

Instead, use synonyms of the words. If you're trying to get someone to subscribe to your newsletter, ask them, "How would you get more information via email?"

Creating elaborate scenarios

I love fiction writing and am guilty of this mistake. Sometimes, I get carried away with my scenarios, and suddenly, the participant has been leading this unbelievable life that has brought them to this product.

The participant has to read through the scenario details to complete it, so including elaborate details that aren't conducive to the task can skew your data. For example, I was asking someone to demonstrate how they would purchase train tickets in the past.

I came up with this scenario: "They wanted to go on holiday to Spain to sit on the beach because work and life were stressful, but they couldn't find the perfect connection, etc."

Although most participants laughed at the scenario, which might have been relatable, it took us out of the real-world situation.

Offending or triggering the participant

I have made this mistake several times. Avoid offensive details like religion, weight loss, politics, and health. Instead of making the scenario about the participant, make it about the subject.

Avoid scenarios of “blaming” the participants or putting them in an awkward situation. Also, consider that you don’t know the participant’s history, so try to avoid holidays such as Mother’s Day or Father’s Day, as [these could also be triggering to participants](#).

Biasing with content

There are usability tests I have observed where the person gives some added “content” to what they want the participant to do. For example, telling the participant to choose the option that gives them an “awesome discount.” Try to keep your words and descriptions [as neutral as possible](#).

The components of a task

As we’ve seen, an essential part of a usability testing task is having a goal, but what other components go into writing an effective usability testing task?

When I write and teach usability tasks, I use the following formula:

Action Verb + Object + Context + Goal + (Optional) Constraints + Endpoint

- **Action** - This is the specific action you want the participant to perform.
- **Object** - What the participant interacts with.
- **Context** - The scenario or context in which the action is taking place. This sets the stage for the task.
- **Goal** - The purpose or objective of the task. What do you want the participant to achieve or find?
- **Constraints (Optional)** - Any limitations or conditions that apply to the task.
- **End point** - Where do you want the user to end?

I use this formula for each task I identified and prioritized in the above list. So, let's build some examples based on the prioritized list:

- Searching for a specific book
- Adding the book to the cart
- Purchasing the book

Example 1: Searching for a specific book

- **Action** - Search for the book
- **Object** - Barnes & Noble homepage
- **Context** - They enjoy reading horror and are looking for Stephen King's new book, Holly
- **Goal** - Find the latest book
- **Constraints (Optional)** - None
- **End point** - Search results page

You love reading horror fiction and just heard Stephen King recently released his new book, "Holly" (context). Using www.barnesandnoble.com (object), search for (action) his newest book.

Example 2: Adding the book to the cart

- **Action** - Add the book to the cart
- **Object** - Barnes & Noble search results page
- **Context** - They want to purchase the book
- **Goal** - Have the book in the cart
- **Constraints (Optional)** - None
- **End point** - Their cart

You've decided to purchase "Holly" (context). Using this page (object), add the book to your cart (action).

Example 3: Purchasing the book

- **Action** - Purchase the book
- **Object** - Cart page
- **Context** - You are ready to checkout
- **Goal** - Complete the purchase
- **Constraints (Optional)** - Stop before payment (unless there are dummy payment details you provide)
- **End point** - Right before payment//after payment (with dummy payment details)

You're ready to complete your purchase (context) of "Holly." From your cart (object), begin the checkout process (action), but please stop just before entering payment details (constraint).

With this formula, you can easily ensure you give the participant relevant context and that your usability tasks have a goal and intention.

Let's look at a few more examples outside of Barnes & Noble.

More examples

Example 4: Clothing website

Task: Using the website www.asos.com (object), imagine you are looking for a red summer dress for a wedding in a month (context). Find the most suitable dress for this occasion (goal) and add it to your cart (action), but do not make a purchase (constraint).

In this task, you could give additional context, such as size, price, or style, if you wanted to see the participant play with filters.

Example 5: Notification

Task: You would like to receive notifications for new meditation exercises that come up (context). Locate (action) the area in the Headspace app (object) that will allow you to change your notification preferences (goal).

Example 6: Comment on Google Doc

Task: You've just received an email with a Google Doc link (context). Click (action) on the link (object) and give your colleague the feedback that the date on page three needs to be changed to October 31st (goal).

Bad tasks, rewritten

Now that we've seen some suitable tasks, reviewing bad tasks, why they are bad, and how to rewrite them is always helpful.

Bad Example 1:

Task: "Please test our website at Airbnb.com and tell us your thoughts."

Why it's bad: This task is overly broad and lacks the most crucial part of a task: the goal. It doesn't give any context or direction and could lead to vague and unreliable data.

Improved task: Imagine you are planning a weekend trip to New York City with a friend from October 28 - October 31 (context). Use Airbnb.com (object) to find (action) an apartment to rent (goal) that is under \$200 a night (constraint).

Why it's better: This task touches on the crucial components we covered above and provides the participant with a clear goal and context so they can complete the task.

Bad Example 2:

Task: "Click around the app and see if you encounter any problems."

Why it's bad: This task lacks context and any sort of goal. Users don't come to your product or service to click around and find problems (most of the time, at least).

Improved task: You are creating a new dating profile on our app (context). Open the Bumble app (object) and complete the basic onboarding information, including uploading a profile picture (action).

Why it's better: The improved task provides a precise scenario and a specific user goal, ensuring the participants know what they are doing during the task.

Bad Example 3:

Task: “Test the checkout process on our site.”

Why it’s bad: This task doesn’t specify which actions the participant should take during checkout or give any context. Users don’t go to products or services to test them!

Improved task: You’ve added a laptop and a mouse to your cart (context). From this screen (object), complete your purchase (action). Please stop right before confirming the payment (constraint).

Why it’s better: The new task provides a clear context, a defined goal (completing a purchase), and specific instructions on what the participant should focus on—resulting in more actionable feedback.

Quantitative versus qualitative tasks

There are a few key differences between qualitative and quantitative tasks, and they mostly come up in the structure and flow of the tasks.

Qualitative tests

With qualitative usability testing, you can:

- Ask the participant to think out loud
- Probe for reactions or perceptions
- Ask [follow-up questions](#) after (or during) the task

Example of a qualitative usability testing script:

Introduction: Hi, I'm Nikki! Thank you for participating in this research session for BestBuy.com. We want to understand your experience while interacting with certain parts of our website. I will give you a few activities to do, and while you are doing them, please think out loud so I can understand your entire process! Remember that this isn't a test; we genuinely want to understand your personal experience. Let me know if you have any questions.

Task 1: Finding a 55-inch smart TV

Scenario: Imagine you are in the market for a 55-inch smart TV. Your budget is \$600. Using BestBuy.com, find a 55-inch smart TV within this budget.

Follow-up questions:

- Describe your overall impression of that process.
- Talk me through what was most confusing.
- Explain one thing that was missing.

Task 2: Checking the return policy for TVs

Scenario: Before you decide to purchase the TV, you want to check the return policy of the TV you're interested in. Using BestBuy.com, find the return policy for TVs.

Follow-up questions:

- Summarize what you found in the return policy for TVs.
- Describe how you would change the experience.

End of session questions:

- Talk me through your overall impression of the entire experience.
- Describe how you felt about the overall experience.
- Walk me through one thing you would change or improve.

As you can see, in qualitative usability testing, you look for the participant's constant feedback through their thinking out loud and explaining their thoughts. You can probe during the task, asking them why they performed a particular action or clicked on a specific area and, after each task, you're asking follow-up questions to understand their experience better.

Quantitative usability test

On the other hand, with quantitative usability testing, you want to:

- Have the user focus on the task in silence so you can measure your metrics reliably, like time on task and task success. Having the user talk while you are trying to measure these metrics can hugely skew the data
- Ask for follow-up metrics such as the [SEQ](#), confidence, and UMUX
- Probe into reactions and perceptions at the very end of the test

Here is that same script geared for quantitative usability testing:

Introduction: Hi, I'm Nikki! Thank you for participating in this research session for BestBuy.com. We want to understand your experience while interacting with certain parts of our website. For the next 60 minutes, I will ask you to perform five different activities. I will give you all the relevant information you need for the activities. You can tell me when you're done with the activity, and I will ask you to rank your experience. Remember that this isn't a test, so there is no one right way to do anything. Let me know if you have any questions.

Task 1: Finding a 55-inch smart TV

Scenario: Imagine you are in the market for a 55-inch smart TV. Your budget is \$600. Using BestBuy.com, find a 55-inch smart TV within this budget.

SEQ: Overall, how difficult or easy was the task to complete?

- 1 = very difficult, 7 = very easy

Task 2: Checking the return policy for TVs

Scenario: Before you decide to purchase the TV, you want to check the return policy of the TV you're interested in. Using BestBuy.com, find the return policy for TVs.

SEQ: Overall, how difficult or easy was the task to complete?

- 1 = very difficult, 7 = very easy

End of session UMUX survey:

- This website/product/tool/software/prototype capabilities meet my requirements.
 - 1 = strongly disagree, 7 = strongly agree
- Using this website/product/tool/software/prototype is a frustrating experience.
 - 1 = strongly disagree, 7 = strongly agree
- This website/product/tool/software/prototype is easy to use.
 - 1 = strongly disagree, 7 = strongly agree
- I have to spend too much time correcting things with this website/product/tool/software/prototype.
 - 1 = strongly disagree, 7 = strongly agree

Follow-up questions:

- Talk me through your overall impression of the entire experience.
- Describe how you felt about the overall experience.
- Explain the most confusing part of the experience.
- Describe the number one thing that was missing from the experience.
- Walk me through one thing you would change or improve.

In the [quantitative usability test](#), we ensure the user is focused on the task so we can reliably measure any metrics you defined, such as time on task or task success. If you have follow-up questions, ask them at the end of the session!

Should users pick their tasks?

One common question I always get asked is if users should pick their tasks to make the scenarios as realistic as possible. If this is your goal, I recommend conducting a [walk-the-store interview](#) to see how participants have used a product or service based on a real-life scenario.

This approach will enable you to see how they use a system “in the real world” and watch the participant work more “naturally” compared to the scenarios and tasks of usability testing.

However, you can still capture that evaluative component in Walk-the-Store interviews. You are watching them complete their tasks and where they run into pain points. When you see this pain point, you dig in to understand it better, as you might in a qualitative usability test.

Practice makes perfect

The best thing you can do is practice writing usability tests with the formula above and conduct dry runs with your colleagues—that is the number one way I built my confidence as a usability tester, and how I teach others usability testing! It takes time to master the science of the usability task, but with time, you can become an expert.

Now that you’ve written your tasks and started gathering data, how do you go about analyzing and distilling your findings? In part three, we’ll go over strategies for analyzing usability tests and sharing what you’ve found.

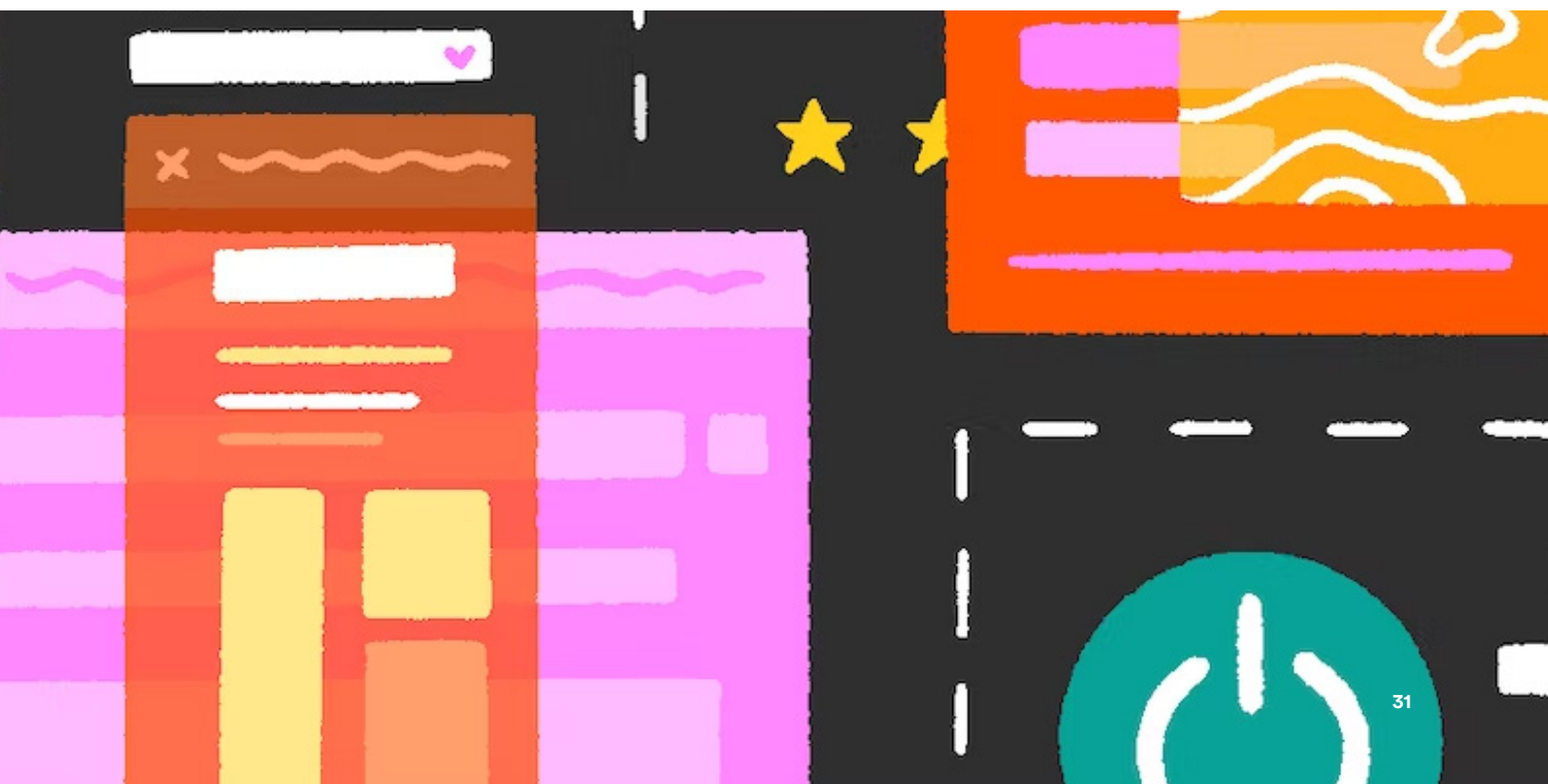
CHAPTER 3

Analysis and shareout

Watching your research data come to life is incredibly gratifying, and the analysis portion of the research process is one of my favorites. You get to see everything you planned come together—and get a peek into patterns and trends.

It's hugely exciting to see the potential of your project come to life, and see where teams could gain so much insight and action from your work.

So, why does analysis often feel overwhelming and confusing?



The challenges of analysis

When I started researching, I struggled so much in the analysis phase of my projects. Analysis was like this veiled mystical journey where one entered with data and magically came out the other side with stunningly [actionable and valuable insights](#).

Let me tell you, that wasn't my experience. I had no idea how to turn the words I got from participants into something my team could appreciate and use. I wasn't sure how to make it [helpful, useful, or actionable](#).

Over time—and with a lot of practice (aka mistakes),—I found my rhythm with synthesis. I learned how to take my research sessions and turn them into insights, gathering information and turning it into outcomes or deliverables the team could meaningfully use.

When it comes to usability testing, sometimes analysis seems simple. Something didn't work, so we should fix it. However, there are steps you can take to make your [usability insights pop](#) beyond the basics.

Get to know what the team needs

One of the problems I encountered when starting to analyze data was that everything felt important. I wanted to include every detail in my report because I feared missing something fundamental. And, to be honest, everything I heard sounded interesting.

So before we dive into analysis, it's essential to deeply understand what your team needs from you. That doesn't mean you can't report on interesting findings outside of that information, but ensuring you get the most relevant data for them is critical. It's what makes a project [successful and actionable](#).

We discussed this in part one, but I wanted to re-emphasize how useful it can be to reverse engineer your goals [based on stakeholders' needs](#). This process will enable us to highlight the most relevant information for them to make the decisions they need to move forward.

I [ask my stakeholders](#) questions when they come to me with a particular research request.

As a reminder, some of these questions include:

- What type of information do you need at the end of the project?
- What decisions do you want to be able to make?
- What is the number one gap in the knowledge you need to move forward?
- What are the top three questions you need to answer?
- What is your definition of success for this project?

Here's one way to do this: I use the formula below to have them fill out a mad lib:

I need [X information] to answer [Y question] to make [Z decision] by [timeline].

Using this information, you can ensure that you focus on the information teams need to make better decisions, move forward, and utilize any outside or additional insights to apply to different projects or continue with more research.

Now, let's dive into the analysis.

Analyzing and sharing qualitative usability tests

Since qualitative research is my favorite, let's start with how to analyze qualitative usability tests. This process is more ambiguous than analyzing quantitative research, so let's get into the weeds to make it as transparent as possible.

Qualitative usability testing is looking for feedback from the user on the team's idea, prototype, or flow. For this, you show participants screens or images and ask them to provide qualitative feedback to understand how they perceive what's in front of them.

With qualitative usability testing, you're gathering information to:

- Discover participants' reactions to and perceptions of the current experience
- Learn about participants' current pain points, frustrations, and barriers to the prototype/experience
- Evaluate how participants work through the prototype/flow and their feedback on the experience

If you think about the above common goals for usability testing, this helps structure your analysis. You're digging into their thoughts, perceptions, and emotions, so you'll get that data from them. This means focusing on understanding patterns and trends of words

Steps to analyze qualitative usability tests

For this example, let's imagine that you are working at a candle company that sells candles online. You've decided to do a qualitative usability test on a prototype you've created that helps people understand better what the candle will smell like.

Step 0: Get aligned

Ahead of time, ask your team for the information they need. That includes:

- To understand people's pain points with the prototype when trying to understand a candle scent better
- To discover people's reactions to the idea of purchasing sample candles
- To identify gaps in the experience and any improvements they could make

The outcomes the team is looking for include:

- To make iterations on the experience so that it better aligns with users' needs
- To improve the flow and experience based on feedback
- To know the top pain points users experience with this prototype
- To feel more confident they are going in the right direction with the experience/flow

You don't have to do too much with this information immediately, but you'll see its importance as you begin synthesis.

Step 1: Transcribe or debrief

Depending on time and resources, there are two different ways you can go about this step. The first option is to transcribe each session into a Google Sheet or Excel while highlighting key quotes or moments you want to remember. I typically transcribe my sessions within 24 hours. I do this by listening to the session and writing down the exact word-by-word in Google Sheets—it is a transcript.

lab.fromatob.com		
No connection found, but I did find it on the DB website which is very strange	pain point	
something else about DB deals recently, I was looking for a trip from munich to frankfurt to make a reservation in a special train compartment because I have a baby and a stroller, and there is a special baby compartment, but i could not book this on fromatob, so i had to go to DB directly	pain point & need	
i believe i could make a reservation to be in a large space, or by the window, but i couldn't book any special needs, which is what DB allows me to do	need	
other stuff as well, like disabled people, there are special compartments like this as well	need	
i know there are a lot of parents traveling frequently on parental leave		
recommendations here, i wouldn't go for the cheapest which is bus because i am traveling with a baby, and sure fastest connection	motivation	21
i would like to go for a train with the baby compartment, so i would just disable the other ones - specifically for this connection, flying from munich to frankfurt seems silly and not very sustainable from my point of view	motivation	
what is recommended? why is this recommended? this doesn't make sense, why? because the price is the same as the cheapest, the fastest makes sense, so i would be confused about why this particular connection is recommended, but maybe because of the time	pain point	
to be honest, i would rule out the options very often, i would go for a particular time: i want to leave between 9am and 3pm and I don't want to arrive later than 8pm; I don't want to change trains	pain point	23:45
okay, so nothing is found, then i would maybe go, that's unfortunate because that is when i want to travel...still nothing found so I guess i would change the time but keep the direct	pain point	
90EUR for a trip is too expensive so then i would look for pricing options where i could see everything which is under 50EUR, can't change this here	need	25:45:00
now there is a limited choice, but it would be nice to have the price as well; it is really common when you use sites like booking.com or airbnb	need	
unstructured here: i can see the top three, but what would be nice, the recommendations were highlighted a little more by color so you know exactly what they are, but then i see it is organized by price, but i would like to organize the order by timing	need	
here you have price, it would be nice to be able to sort by timing, to travel as early as possible; sorry by departure time, you could argue to go back to the filters but it would be nice to have here too	need	
Didn't return any results: if i already decided I wanted to travel by train, i would go to DB directly, or for flights, I would go to google flights	pain point	
booking.com		30:15:00
the filters on the say, i could say limitation by price, how much do i want to spend per night; the other filters relevant to accomodation, but i don't know if you could translate it to transportation	need	
suitable for bringing along children or filter for sustainable mode of travel to get recommendations, maybe there are some better ways	need	
maybe like convenience as a filter, you already have the fastest way of traveling, so maybe not		
lab.fromatob.com		
another way you could do price filter is the total budget you would want to spend on trip, i like the slider, but maybe there are also people who like to type it in		34:45:00

I must admit that this is time-consuming, but it's an excellent option for you in certain circumstances like:

- You are the only one doing the research analysis/synthesis
- You want to remember each session individually to ensure you won't miss anything
- You have time and patience to do this!

Another way to transcribe is by using a transcription service or paying someone to transcribe your sessions. I typically don't use this because I lose the benefit of transcribing the session to ensure I remember everything. This is a great option if you want the transcriptions as an artifact for your team!

The other option is to conduct a debrief session. I will get into these more in-depth when going into the next step of categorizing and organizing data. In sum, a debrief is a nugget-sized synthesis session, focused only on one session at a time and usually held right after the session.

Debriefs are a wonderful option if you are:

- Synthesizing with multiple team members
- Wanting to speed up your synthesis process by not waiting until the very end to synthesize everything at once
- Interested in teaching people how to synthesize on a smaller level
- Able to hold the space for this session right after the usability test

Those are the two main ways to gather the necessary information to move to the next step.

Step 2: Choose tags/codes

Remember the information you asked from stakeholders in step zero? That is where this comes into play.

The number one goal of synthesis is to make sense of all the data you gathered. Take that raw information and find patterns to highlight the most essential data you found. This step is where tags/codes come in. I will use these two terms interchangeably from here out.

Tags allow us to categorize large amounts of data to uncover the bigger trends and patterns across participants.

There are two ways to choose your codes:

- **Inductive method:** With the inductive method, you don't create any codes until you have reviewed some data. You then find the codes in your data.
- **Deductive method:** You come up with codes before you synthesize your data.

Especially with usability testing, I typically use a deductive method and develop the codes before I go into usability. And that is where the information you gauged from stakeholders becomes critical because it helps us determine tags.

Stakeholders were looking for people's perceptions of the prototype, pain points, experience gaps, and reactions to purchasing sample candles.

You'll need to find data that helps them understand the above points more deeply so they can make better decisions and move forward with the idea.

With that in mind, let's choose the tags to focus on:

- Overall perceptions
- Pain points
- Experience gaps
- Reactions to samples

I usually only pick four because, beyond that, it can start to get messy. However, I always include a fifth area called the “parking lot,” in which I include any other interesting or important data outside the scope of the team’s immediate needs.

Step 3: Categorize the data

It's time to use those codes to categorize the data into the patterns you find and get those larger findings.

Whether you transcribed or have chosen to go with debriefs, this will look slightly different.

If you have transcribed your data in Sheets or Excel, you can code the data right in there. Look through your transcription and what the participant said, and write down any time you find a code that would be applicable.

For example, let's say participant one said:

- “I have to look up what all of these scents mean—no idea what frankincense is—so listing these different odors actually doesn't help me determine anything at all.”
 - This would fit nicely as a pain point for the experience
- “I've bought samples of candles in the past to figure out what I liked/didn't, and whatever I didn't, I just used as a gift. It was very useful before I bought a bigger candle as they're quite expensive.”
 - This would be tagged under “reaction to samples”

notes	tags	timestamp/comments
the last time was when i went to frankfurt the last time, right before my trip to ottawa		02
i usually use the app, not the desktop, the mobile version, i definitely only use incognito in chrome, if i do use the computer because i like to be sly and clever, but i have an account anyways		
i guess i find it really easy to use, where i am, where i am going, and start searching, and that's great		
i was at home, i usually book things at home, kinda last minute, i'm bad that way, within 24 hours or so, it is a bad habit i need to get abhold of		
usually booking in the evening time, my husband and i work during the day, so the only time we have to book things is during lunch hour or in the evening time, usually 8 or 9 at night, sometimes 10, after dinner, you can relax and book things		
i am an serial impulse traveler, which is not great because i have definitely made egregious errors with my timetables in the past, and so i like to have an hour to look around, have my dates correct, and everything		
when we can travel together, often, this time i went by myself, we got married september of last year and are actually having the big celebration in september, so i just went to finish wedding planning - he couldn't come because he had to work, but usually we go together		
i usually sit on the couch, and open up my phone, im the world's worst technology users, when im a grandmother i will be made fun of forever, i open my phone, go to the website, where i am, where i want to go and the date - so for train tickets, unless it is a planned weekend trip to paris or something, and i'm not 100% sure when i want to leave, i very often book one-way tickets at a time. i look for Koln to FrankFURT, see what there is and usually take the one that has the best timing option, and, with this platform, i dont feel the need to switch to my laptop, which i normally feel the need to if i am booking an ICE train, the website is a bit complicated, but fromatob works really well on my phone, i put in my info, and such		5:13
i put the info in, hit go, take a deep breath so i don't make any mistakes like last time, put in my credit card info, take a deep breath and that has been working very well		
usually i have the TV on mute during that, so i don't get distracted		
i don't use blabla car, and i have no interest in it, i really can't hold a conversation for that long, it sounds so painful, so i tend to skipover those...it doesn't matter, but i have no interest, i dont have the ability to just chat for four hours or longer		
its great that blabla car is there, because people use it, my friends do, it is like a student thing so i understand the purpose		
i like buses, i use that as well, they are great. i think if i were on my own with less luggage, i would use buses more often - i do enjoy them and look at them to look at the timing, definitely a consideration, but trains usually work out the best between scheduling and budget		
i travel pretty frequently from frankfurt to koln or london, i am getting back into becoming an artist again, so ive been honing my skills to packing one singular carryon piece that i own, for upwards of 3 weeks at a time, trying to get better at that...and it has been working, so that time i could've easily booked the bus but had already booked the train		
i don't really use a lot of discount or deal travel websites for domestic travel, like within EU, but, if i am going a shoert distance within germany or netherlands or belgium, i will usually not look around, i will use the fromatob platform to see what is on there, and if the timeframe doesn't work there, then i will go through thr ICE website to see what they have...and then i see why they aren't on fromatob because they are expensive...but sometimes i will just buy the more expensive ticket because i need it to be at a certain time, but i do try to be more flexible when i travel		11:15
i do use other websites for flights, often, i have a couple i prefer and we also have a travel agent for our longer flights that are more complicated, because it makes it so much easier		12:28
i use edreams pretty often, it is new, i used aerlingus for a while, but they got a little more expensive, so i stopped...and now i use skyscanner and SAS travel, i use that now, it depends on my destination, on edreams and skyscanner, you can put in the airline you want and that is nice, but that is it...i have never booked through aircanada or lufthansa because there flights are so expensive and i dont need to do that		
edreams looks like a total scam, you open it and you're like 'okay great' but their tickets work...the first time i traveled with them, i got to the airport and thought, 'what are the chances that this is a fraud ticket' but it worked! and i was so relieved. i had a contingency plan in case it was just garbage		
another one i use is expedia.ca, which is canadian, weirdly i find if you find anything from canadian or american websites, it is more expensive than booking it from the european website, which i dont really get but oh well		
we don't use the same person for the travel agent every time, but we use an office in Koln, STA travel, they are great, really friendly and they hire people who enjoy traveling and, you go in and they have people who socialize in certain areas, because that is where their interests are, which is really cool and they make you feel important when you so in, but i have no idea how they set		

If you choose the debrief route, you must run a debrief session after each usability test. Here is my step-by-step process:

① Decide who to invite

Invite the stakeholders who are observing that session (e.g., designers, developers, product managers) because they will be able to participate. If more people are observing outside the team, I allow them to join. But I highly recommend the people working on the project to be present.

② Create timeboxes for each code

For each code you decide on, ensure a specific time for each. For instance, give people five minutes to brainstorm pain points individually and then five minutes to discuss them.

③ Create the board

In-person, use a whiteboard. Remotely, use [Miro](#), Jamboard, Trello, or Mural. Before the debrief session, create a template you can fill out. This template gives the team structure to fill out the appropriate information at the right time but allows for creativity. Each board = one participant.

④ Set the agenda

You want a clear schedule for everyone to follow, which can include the following:

- Explain board: 5 minutes
- Overall perceptions brainstorm: 3 minutes
- Overall perceptions discussion: 5 minutes
- Experience gaps brainstorm: 3 minutes
- Experience gaps discussion: 5 minutes
- Pain points brainstorm: 3 minutes
- Pain points discussion: 5 minutes
- Reactions to samples brainstorm: 3 minutes
- Reactions to samples discussion: 5 minutes
- Wrap-up: 3 minutes

Step 4: Find patterns and trends across participants

Once you've coded each participant's overall perceptions, experience gaps, pain points, and reactions to samples, you'll look across all our participants to find patterns and trends.

This is the most exciting part because you can start to see the emerging findings from the data and what you will eventually share with teams.

To do this, create one giant board with the codes. For each participant, move their data to this bigger board, so now you have a big board with everyone's data.

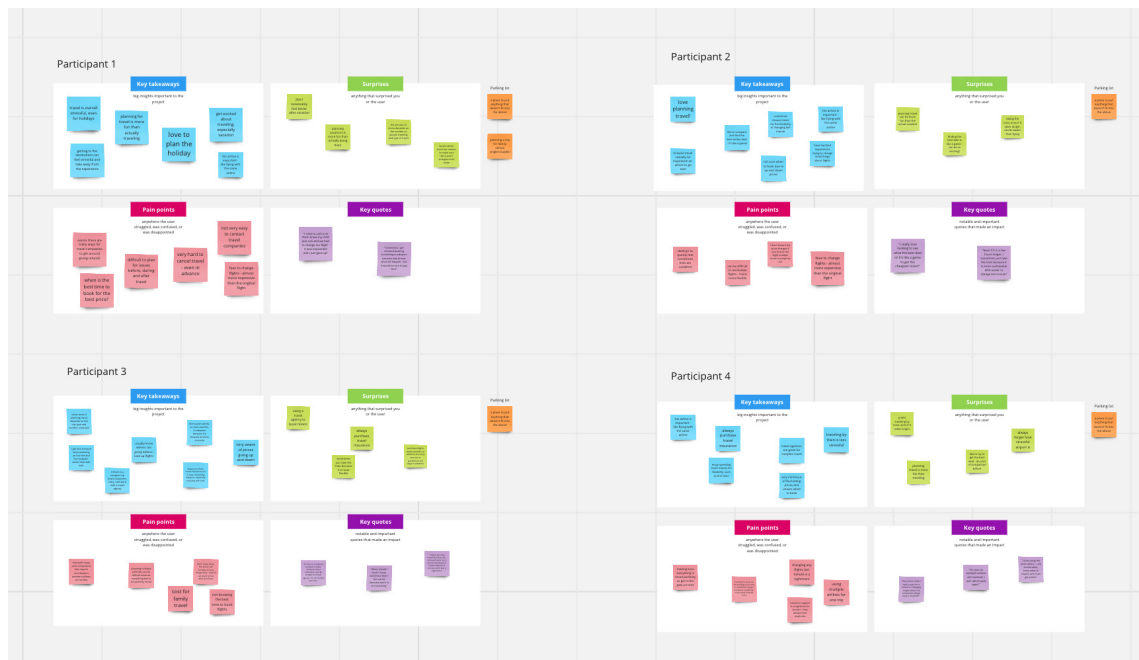
Then, move all the participants' pain points under the pain point section, overall perceptions under that, and so on. Now comes the magic. Take each section and look at all of the different data under it. For instance, you could start with pain points and look at all the various pain points participants experienced.

As you look, cluster things that sound the same. So, going to the earlier example of not knowing what the different scents mean, let's say participants two, four, five, and six said that. Combine all of those into a cluster and name it "difficulty identifying scents."

Usually, I define patterns or trends by the sample size. I typically think of a trend if 1/3 (round up) of the total sample size said something similar.

So, for seven participants, if three or more said they had never bought sample candles in the past, I would categorize that as a pattern. If I spoke to 15, I would consider a pattern after five people said the same thing.

This isn't an exact science and there can be exceptions, so you don't need to use this as a hard-and-fast rule. For example, one-off insights can be compelling. As you advance in your process, you will better judge patterns and trends, but the 1/3 rule is a nice place to start. Do this for all the data in each section, combining similar information and clustering it. These become the patterns that you share with the team.



Step 5: Share the data

After identifying the patterns, it's time to share these with the team and see the research come to life.

Sometimes, the team will be involved in creating the patterns. In this case, I ask them if they need a separate deliverable, such as a report, or if the board you used to synthesize is okay. You might not have to create a separate report depending on their needs.

Once you've identified your patterns, it's time to pick the most impactful and important ones. You can select the patterns to focus on based on the weight of the cluster (how many participants brought up that particular point), or you and your team can discuss your perceived impact of each of them.

I recommend a mix of the two—find the highest weighted patterns, the ones most people spoke about, and then identify the level of importance of each.

Once you have this information, you can put it into a report if your team needs that. I typically take about five of the most important patterns and report on them and write a more detailed finding, which includes:

- The key learning of that finding
- The why behind the finding, such as why someone encountered a problem
- The consequence of the finding, which suggests what might happen if you don't do anything about the finding

Step 6: Activate the data

Depending on your team's maturity and the project's complexity, your team can move forward based on the findings you uncovered.

However, suppose they're having difficulty figuring out the next steps. In that case, I highly recommend an [ideation session](#) to empower them to respond creatively to your research findings.

Analyzing and sharing quantitative usability tests

Since quantitative usability tests are about numbers, analyzing and sharing this information feels much more straightforward.

As a reminder, quantitative usability testing strictly assesses the three cornerstones of usability: effectiveness, efficiency, and satisfaction. Through this session, evaluate how participants move through the experience you put in front of them through metrics. In this session, you aren't asking for qualitative feedback, as that would skew the metrics. You're also trying to mimic a real-life experience as much as possible.

For quantitative usability testing, you are trying to:

- Evaluate the effectiveness, efficiency, and satisfaction of the product
- Understand the issues and problems participants encounter in the product
- Determine the overall usability of our product

Metrics for quantitative data

Certain metrics you track allow you to measure this information. The most common examples of metrics include:

Task success

This simple metric tells if a user can complete a task (0=Fail, 1=Pass). You can get fancier with this by assigning more numbers that denote users' difficulty with the task, but you need to determine the levels with your team before the study.

Time on task

This metric measures how long participants can complete or fail a task. This metric can give you a few options to report on, where you can provide the data on average task completion time, average task failure time, or overall average task time (of both completed and failed tasks).

Single Ease Question (SEQ)

The [SEQ](#) is one question (on a seven-point scale) measuring the participant's perceived task ease. Ask the SEQ after each completed (or failed) task.

Usability Metric for User Experience (UMUX or UMUX-Lite)

The UMUX/UMUX-Lite is the “newer” version of the SUS and measures the experience’s perceived usability. The advantage is the UMUX is only four items, and the UMUX-Lite is only two, making it easier for participants to take.

Because you already have these metrics, the steps for analyzing a quantitative test are fewer and more straightforward.

Steps to analyze quantitative usability tests

Step 0: Define your metrics

Before anything, you must define which metrics you will track for the usability test. This step ensures that you can track them across the different tests to have consistency. This step is crucial if you are running moderated tests and will manually be collecting the data.

Step 1: Collect the data

Once you choose your metrics, it’s time to run the tests and collect the data. There are two ways this can happen:

- **Unmoderated studies**, which you configure to run on their own. Many unmoderated platforms automatically collect metrics like task success or time on task, but make sure that you add in any other metrics you need, like the UMUX or SEQ.
- **Moderated studies**, where you collect the data manually as the participant is completing the test. For this option, it’s necessary to have someone helping you so you aren’t attempting to juggle tracking four metrics simultaneously.

Step 2: Create a stoplight chart

You can create a stoplight chart using spreadsheet software like Microsoft Excel or Google Sheets. If you’re having difficulty getting started, [check out this template](#).

Task	P1	P2	P3	P4	P5	P6		Summary
Task 1	60 second	45 second	100 seconds	50 seconds	Fail	113 seconds		Average time
Task 2								
Task 3								
Task 4								

For each task, assign colors to make it clear what the task success was for each participant and the average per task:

- Red = failed task
- Orange = struggled with the task
- Green = succeeded with the task

Step 4: Populate the chart

For each participant and task, record the task completion, time on task, and any other metrics you tracked.

Step 5: Categorize and describe each issue

Review the usability issues you've identified during the usability test, and categorize each into issue types.

Critical

These are issues that severely hinder or prevent users from completing essential tasks or achieving their goals. They can result in high frustration or abandonment of the task or system.

Important

These issues are significant but may not be as severe as critical ones. They can cause user frustration or confusion but are not complete showstoppers.

Minor

These are minor issues or suggestions for improvement that don't significantly impact the user experience. They may be cosmetic or inconsequential.

Step 6: Prioritize and plan action

After creating the stoplight chart, review it to identify patterns and prioritize which issues to address first. Critical issues should be addressed immediately, as they significantly impact usability and user satisfaction. Important issues should also be addressed, while minor problems can be tackled whenever there is space in the roadmap.

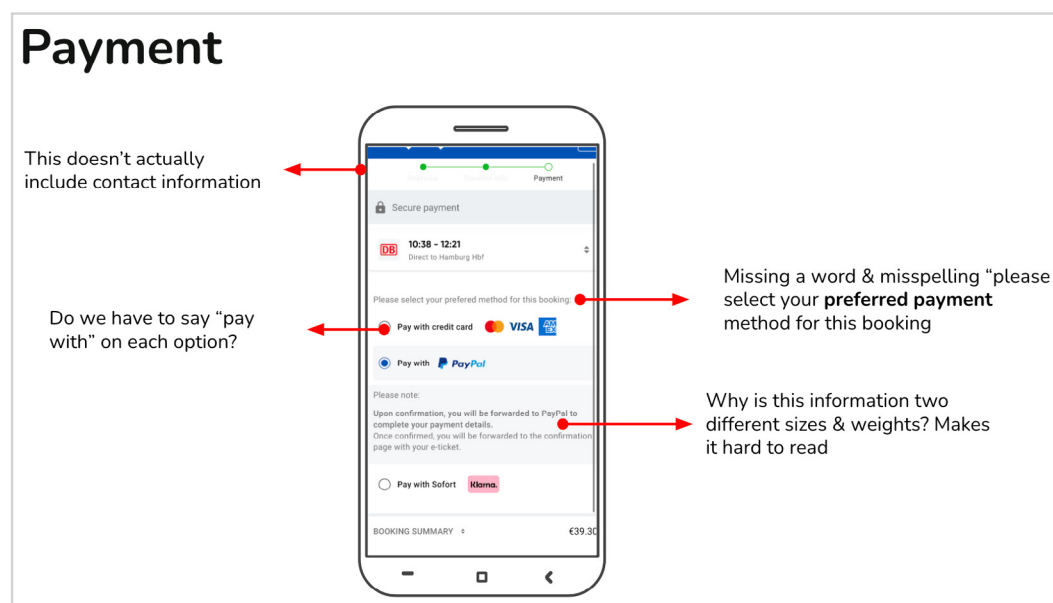
Step 7: Present and activate

Once you have your stoplight chart, you can present your findings in a few different ways, depending on the needs of your team. For many quantitative usability tests, I sit with my team to review the videos of the task failures and most critical issues. With this, brainstorm ideas or solutions to fix the problems.

If you need to, and (sometimes it's necessary!) you can create a usability testing report where you go into more detail with annotations and text.

Analyzing and sharing usability test results is a vital step in the product development process and can bring about significant change. By presenting findings in a clear, accessible manner and emphasizing their impact, you can empower your team to make more informed decisions for users, which is a hugely gratifying experience!

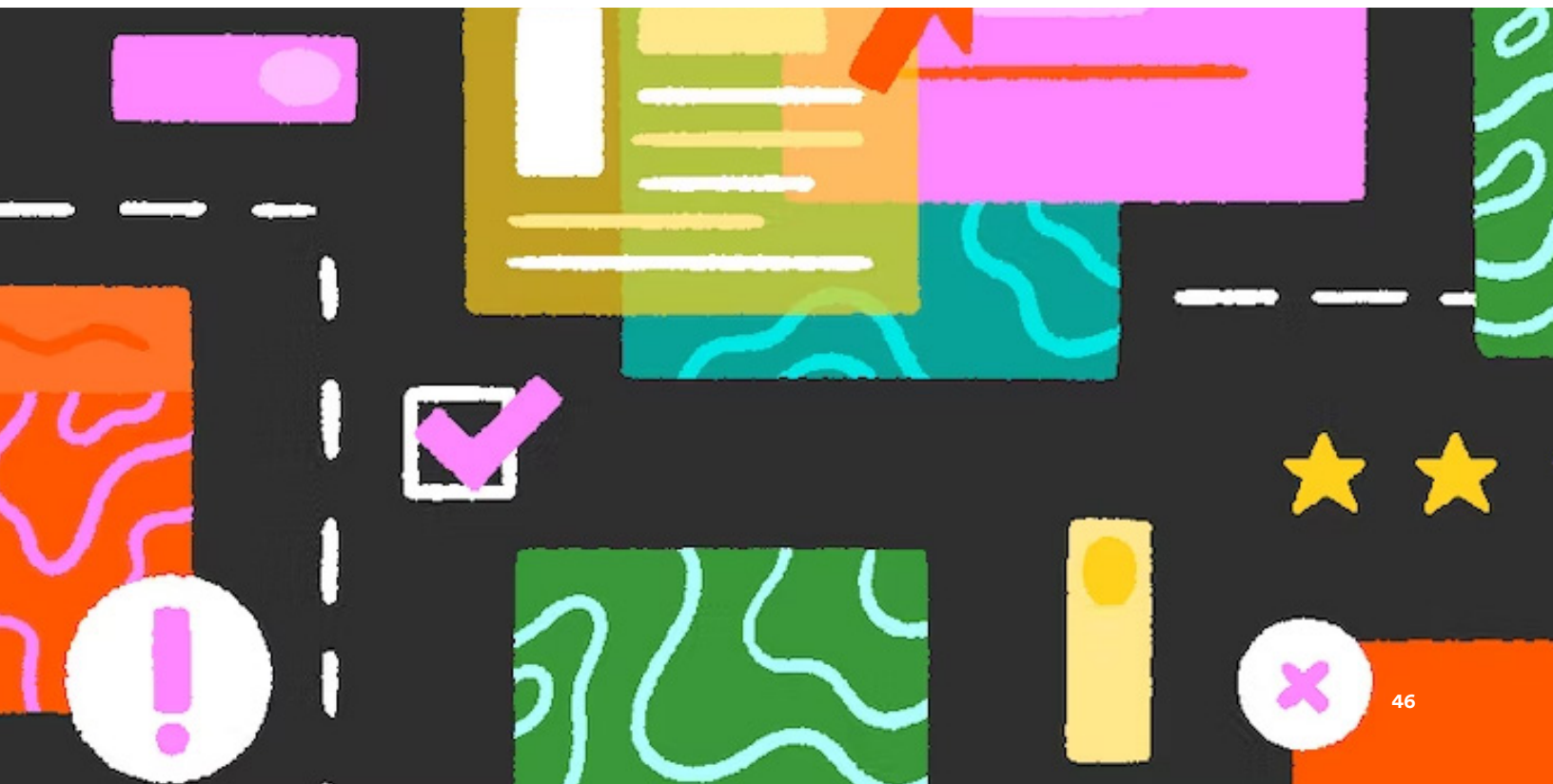
To bring it all together, we'll take a look at a practical case study to illustrate exactly what this process can achieve.



CHAPTER 4

How DropBox approaches usability

We talked with Design Researcher, Meghan Earley, about how DropBox met user needs by going beyond standard usability testing—leveraging video and diary studies to look longitudinally. As a result, they gained more confidence in their issue reports and saw increased investment from their stakeholders.



The challenge:

People use Dropbox in browser to host and share files—relegating it to more passive, background usage. So when the company began the development of their new desktop app, they strived to create a single workspace for users to organize their content, connect their tools, and bring their teammates together.

But when you build a product for more flexible and frequent usage, you have to be sure that it'll work as intended. And to be confident that it'll work as intended, the insights that you need are often more extensive than what you could glean from a typical usability test.

And so, she turned to dscout.

“I looked at dscout because we wanted to do a longitudinal study,” she says. “It was the first time we had people using the product outside of interviewing and concept testing. So we really wanted to get a sense for their day-to-day: How are they interacting with this app, and what are their attitudes towards it?”

“We were hoping to get some in-context feedback. We needed participants to submit surveys in the moment they were doing things. That made a major impact in our attempts to understand what the real issues were.”

The solution:

The Dropbox study took four weeks total—longer than most. However, the length was necessary for the type of insights Dropbox was looking for, as well as the product they wanted to release.

“We kept it pretty open-ended in the beginning,” Meghan says. “We didn’t want people to feel like they were doing something right or wrong. We really just wanted to be a fly on the wall and understand what was going on.”

The second half of the study was more straightforward. Participants were sent specific survey-like questions asking about different parts of the product. And as the study concluded, Meghan conducted a Live mission—pulling specific users in for 1:1 interviews about their experience.

“The interviews were nice because we already had so much context,” Meghan says. “They were really efficient. You can get straight to the heart of things after having heard from this person on dscout for the past four weeks.”

The impact:

Leaning on dscout’s platform, Dropbox was able to conduct a longer study and expanded the breadth of their insights as a result.

“dscout offered us a complete understanding,” Meghan says. “Usability problems are glaringly important from an evaluative perspective. In a typical usability test, we’ll see someone encounter something once over the course of an interview. But when we’re seeing people encountering things over and over, it’s definitely a signal that they’re more important.”

This allowed Dropbox to address those issues, and address areas in which they could really impact their users’ needs.

“We identified some key problems pertaining to the new functionality that we’re adding. We’re adding features to help people collaborate and work with each other better. And there were some pretty key blockers to people being able to do that.”

“So identifying that has informed our design direction and understanding what we need to do in order to help people collaborate more in Dropbox.”

Benefits of usability testing with dscout

dscout can help you gather the in-the-moment insights you need on your timeline. See how we can support your usability testing needs.

A centralized experience keeps you focused

With built-in operations like recruitment, screeners, incentive processing, and a single-view research activity builder, usability is more accessible and nimble.

A platform approach extends insight impact

Why stop at a single usability test? dscout's suite of moderated and unmoderated tools offer more variety for usability testing: follow-up interviews, pre-session trend spotting, or a rolling iterative approach are all possible from the same product.

A partner to augment your team

dscout's staff of trained researchers can help translate research briefs, advise on analysis approaches, or help with field management. This support goes beyond just tech help: our team becomes an extension of your own, critical in these lean times.

A balanced approach to automation

Thoughtfully integrated automation like response quality checks, expressiveness filtering, full session transcriptions, and usability-specific analysis (e.g., time on task) keep momentum and open time for deeper work where you need it most.

A way to engage stakeholders

Make the insights stick by bringing collaborators and decision-makers into the process, whether that's through easy-access viewer status, one-click data share links, or customized video playlists. Usability with dscout makes research a team sport.

"It was great that I was able to share the results from the analyze page with someone else on my team. Huge time-saver! Otherwise, I spend a lot of time taking the results and re-summarizing it somewhere else." – UXR at a Leading U.S. Airline

Dive into usability testing with confidence (and support)

With dscout you can recruit the right participants, choose from an array of question and task types, analyze with ease, collaborate with your team, and get ideas or advice from one of our designated research advisors.

[See how dscout can support your usability testing goals](#)

You're fascinated by the why. We break down the hows.

Get the People Nerds blog sent to your inbox to hear from top UX practitioners and pick up a few pieces of novel advice for every stage of your research project cycle.

[Subscribe to the People Nerds Newsletter](#)

Join fellow People Nerds in our Slack Community

Discuss all things human insights in a space where folks share ideas and inspiration, discuss industry trends, highlight upcoming events, and so much more.

[Check out the People Nerds Slack community](#)