



By

Siri Mehus, Ph.D.

---

At Blink, we practice evidence-driven design. That means that the design recommendations and decisions we make are grounded in solid data and sound reasoning.

For the majority of our projects, our evidence comes from qualitative, [observational research](#). Blink researchers are also fluent with [quantitative methods](#), which we use in combination with qualitative methods to deliver insights that meet each client's design and product development needs.

Let's look at the subject of unmoderated usability research: When should a live researcher facilitate research sessions and when is remote unmoderated usability testing a reasonable choice?

## New tools for usability research

In recent years, a number of tools and services for conducting remote usability tests have emerged in the marketplace including UserZoom, UserTesting, Loop11, Validately, and others. These tools make it possible to test products with real users even if you don't have access to a usability lab or a trained researcher to run the sessions, and without spending days or weeks recruiting participants. While this is clearly preferable to the alternative of not testing with users at all, how does it compare to the "gold standard" of in-lab moderated usability testing? Is this relatively inexpensive approach a typical case of "you get what you pay for?" Or are there situations when it can deliver the value we expect from full-service usability testing, but at a bargain price and breakneck speed?

# Standard usability testing

While approaches and projects differ, usability testing as typically performed has some key components:

- Users perform a set of tasks with a product or prototype.
- Sessions are moderated in person by a trained researcher.
- Participants are asked to “think out loud” about what they are seeing and doing.
- Testing takes place in a usability lab.
- Participants’ talk and actions are recorded to facilitate later analysis.

This standard [usability testing](#) model can be — and frequently is — adapted by removing one or more of these elements. For instance, think-aloud protocol may not be employed when, for example, time on task is being measured and participants’ vocalizations could potentially interfere with task performance. Or usability tasks might be performed in a setting other than a usability lab, e.g., as part of a field study. Or tests may be conducted remotely rather than in person when it is important to recruit participants from a wide geographical area. In those cases, a teleconferencing tool (Zoom, WebEx, Skype) can be used to approximate the in-person moderated protocol at a distance.

A more significant divergence from the standard plan is to conduct sessions without moderation from a trained researcher. There are some good reasons for taking this approach, some reasons that sound good but really aren’t, and some definite drawbacks.

## Remote unmoderated usability testing

Remote unmoderated testing differs from the standard model in several respects:

- Participants perform tasks on a website or prototype using their own computers or devices.
- Sessions are not moderated; rather, participants follow directions that have been created beforehand.
- Participants’ clicks and taps, pathways, page views, time on task, or other metrics can be tracked electronically.
- Participants may be asked to provide ratings of satisfaction and ease of use or qualitative feedback in a survey or text entry box.
- Some tools allow testing of apps and websites on mobile devices.
- Participants’ screens and participants themselves may or may not be digitally recorded for subsequent analysis.

This last point is a key distinction between types of unmoderated testing. Eliminating audio, video, or screen recording further simplifies the process, but it also removes the possibility of learning from participants by directly observing their activity and listening to their commentary.

## The promise of remote unmoderated usability testing

### Saving resources

Remote unmoderated testing offers some very attractive benefits. One is obviously and undeniably the cost. Lab costs, recording equipment, and perhaps most significantly, the time and labor of a skilled research moderator, can all be avoided.

## **Saving time**

The speed of remote unmoderated usability testing also contributes to its appeal in the lean tech environment. Recruiting and scheduling qualified participants for an in-person usability test can take two to three weeks. In contrast, if it is a simple recruit, completed tests may start to come in within hours of submitting to a remote usability testing service. The duration of testing can also be shorter, because multiple participants can take the test at the same time. This quick turnaround can be of great value to teams moving and iterating at a fast pace.

Lower costs and faster turnaround open up the possibility for testing with larger numbers of participants than a standard lab usability study. This sounds like a great advantage, but should be considered carefully. Larger numbers do not necessarily provide better answers. They can give you greater confidence in your answers to some questions, but those may not be the questions that product teams need answered.

## **Potential pitfalls with unmoderated testing**

Other reasons are sometimes given for taking this approach that, when considered carefully, fail to deliver as strong motivations for pursuing unmoderated testing.

### **Avoiding effort**

Remote unmoderated usability testing is cheaper and faster, but it may not be easier.

A moderator is not required to be present for each session, asking questions, taking notes, and recording key observations. This saves substantial time, but if sessions are recorded, someone will still need to review them. Watching videos to find key moments, identify insights, and prepare supporting evidence can take a substantial amount of time.

You can reduce reviewing time if you use metrics to identify key moments to observe directly, but this can be more tedious than interviewing and observing people in real time. A streamlined post-session process requires careful study design to ensure that the appropriate metrics are being captured and used to trigger review. More importantly, it introduces the risk of missing valuable user feedback.

Certainly, workload can be reduced significantly by not recording user actions (or not reviewing recordings), but there will be a concomitant reduction in the value to be gained from the usability tests.

If you opt not to record sessions, other means must be used to determine what users do and whether they are successful. Asking users to report their own success or failure and comment on their experience is unreliable, at best. Event tracking may be limited. In a previous remote study of a mobile website, we found that we could only track participant actions that corresponded to a unique URL. We were able to overcome that limitation and detect participant actions with drop-down menus and carousels by adding code to the prototype, but that required more time and front-end development.

Whether or not recordings are made, remote unmoderated tests must be carefully designed and written because it is impossible to clarify intent or make adjustments during a session. An ambiguous question or misleading instruction can easily result in useless data if it leads participants in the wrong direction. Studies need to be written with the greatest attention to possible misinterpretation and should be piloted or pre-tested with several participants before release to a larger group.

In short, remote unmoderated usability testing can be easier and faster than in-lab testing, but

doing it well requires a concerted effort by someone with significant expertise.

## Removing bias

Some advocate unmoderated testing because moderators can inadvertently bias participants through the way they word their questions or respond to participant statements or actions. However, poorly worded session instructions in an unmoderated test can introduce as much bias as a poorly trained moderator.

Moderation actually offers opportunities to understand and thereby reduce the effect of participant bias. A competent moderator will probe participant responses in order to uncover biases that may be informing their responses.

For example, we sometimes hear participants express preferences for particular companies or brands. As researchers, we want to know why. For instance, if a participant says they would be more likely to trust Amazon, we will probe. If a Seattle resident says it is because Amazon is “local,” we recognize that the user is displaying a bias that is not representative of this company’s global customers.

Moderated tests allow us the opportunity to start with open-ended questions and then probe to make sure that we address particular issues of interest. On the contrary, asking open-ended questions in an unmoderated test is riskier because the participant may take it in a direction that is not of much value to the researcher. The test-creator may thus choose to use more closed or directive questions, which can potentially limit possible answers.

In short, moderation is not inherently more biased than unmoderated testing and offers many advantages for better understanding participant motivations for acting or answering in a particular way (in other words, their biases).

## Improving validity

Conventional wisdom says that testing larger numbers of participants provides more valid results. While you can conceivably test with larger numbers of participants, unmoderated testing simply does not provide more information about why users are having that problem, or how to solve it. It can be useful to know how prevalent an issue is, but most product teams are more concerned with what causes the problem and how to prevent it. That’s where moderated testing delivers the best insights, whether the testing is remote or not.

# Where unmoderated usability testing falls short

## Understanding why

Unmoderated testing is useful for learning whether or not users are able to navigate a site or use a product. It can identify where things go wrong and, if a large enough sample is recruited, predict how frequently particular problems will occur. However, this is not where designers and product teams tend to get the most value from usability testing. Understanding why users are having a particular problem provides direct guidance for making design changes.

### A very simple example:

Unmoderated usability testing reveals that no one is clicking on a call to action.

The product team decides they need to make it more prominent by changing the call to action from a text link on the side of the screen to a big green button at the top.

## ***Versus***

Moderated usability testing uncovers that no one is clicking on the call to action.

Participants indicate that they do not think that is the right place to complete the action they want to perform.

Product team understands that they need to change the labeling in order to convey the purpose of this link.

Making a call to action more prominent will not help if users do not believe it is the right place to go to perform the action they want to accomplish. In the first scenario, research and design resources are wasted because, while testing uncovered a problem, it did not uncover the reason behind the problem.

## **Adaptability**

Moderated testing provides the opportunity to find out why a user makes a particular choice, understand their interpretation of what they see, and follow a user down another path, which can lead to unexpected insights.

When we notice that a participant has fallen silent, understandably focused on the challenging task we have asked them to complete, we can gently remind them to verbalize their impressions and intended actions.

In an artificial scenario, it is a challenge to ensure that users act in the way they would naturally. A lab may not be a natural scenario, but a co-present moderator can remind and encourage participants to take tasks seriously and truly imagine themselves in the envisioned situation. Participants engaging in these tasks on their own in an unmoderated environment may be more likely to simply “go through the motions” in order to finish a test and collect their payment.

## **The right tool for the job**

Each year, we run thousands of user research sessions at Blink. New tools make it easier than ever to moderate sessions from almost anywhere. We have dedicated in-studio labs for moderating remote research sessions but often run sessions from home.

We rely on tools like Zoom, InVision, AdobeXD, OptimalSort, and Mural to conduct remote UX research studies. We have successfully used remote unmoderated testing in our work for clients, and will do so again.

More than anything, we believe in using the right tool for the right job. Understanding “why” is always our first priority, but there are times when gathering data (about “what” or “how often”) can offer clients more comprehensive insights. Remote unmoderated usability testing is one tool we can use to do this, but we know it isn’t always the best way to answer the questions our clients care most about.

*Siri is a seasoned member of the Blink UX research team and has also [contributed to UX Magazine](#) on the topic of conversational repair.*

[Contact us today](#)