

[Open in app](#)[Get started](#)

Published in UX Collective



Joe Bernstein

[Follow](#)Dec 20, 2021 · 7 min read · [Listen](#)

Save



Doing RITE wrong

User research is a science — don't overlook the scientific method

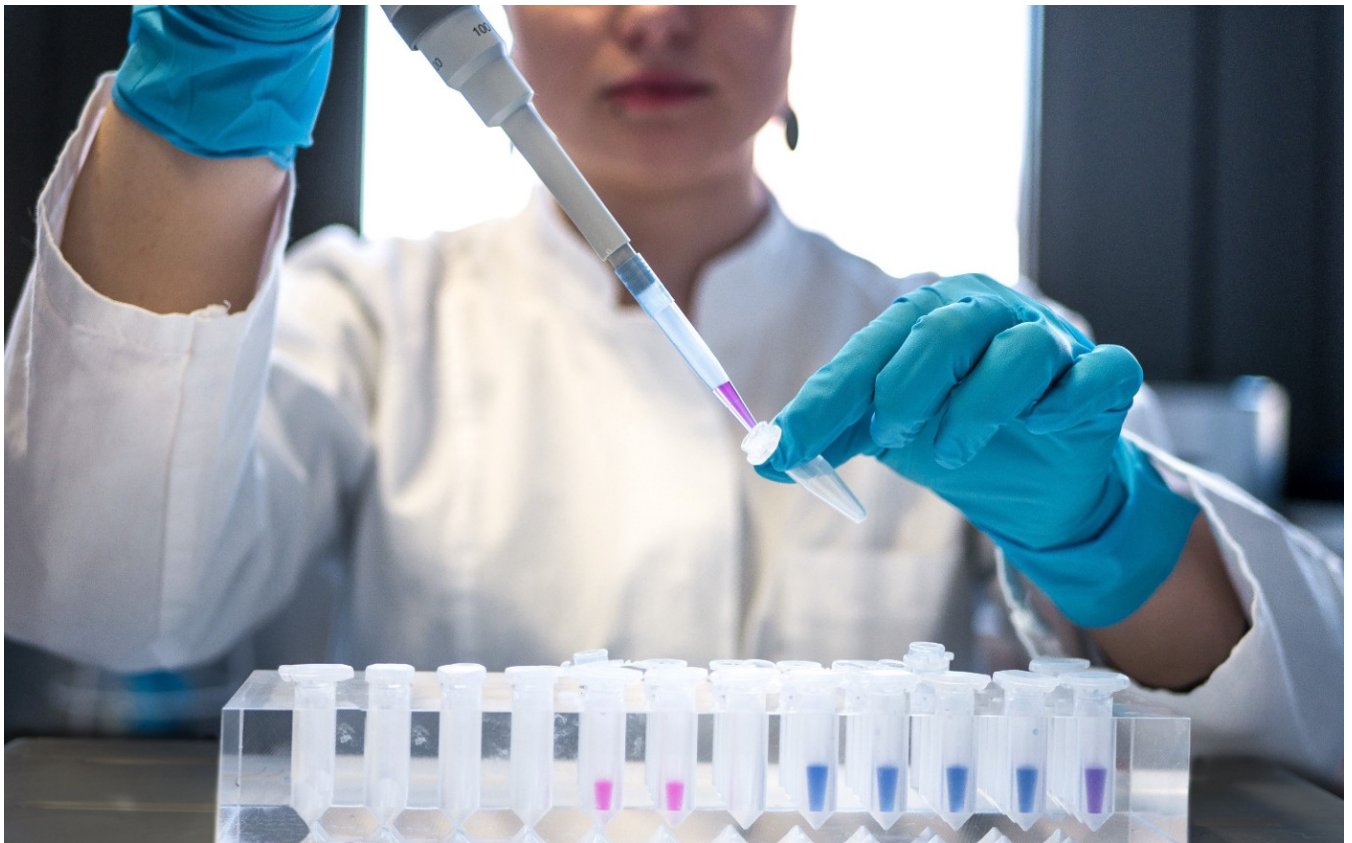


Photo by [Julia Koblitz](#) on [Unsplash](#)

User research comes in many forms, but popular among them is **RITE**, or **Rapid Iterative Testing and Evaluation**. The RITE Method isn't just a popular way to conduct



[Open in app](#)[Get started](#)

I've also participated in RITE studies that were accurately rapid and iterative, but

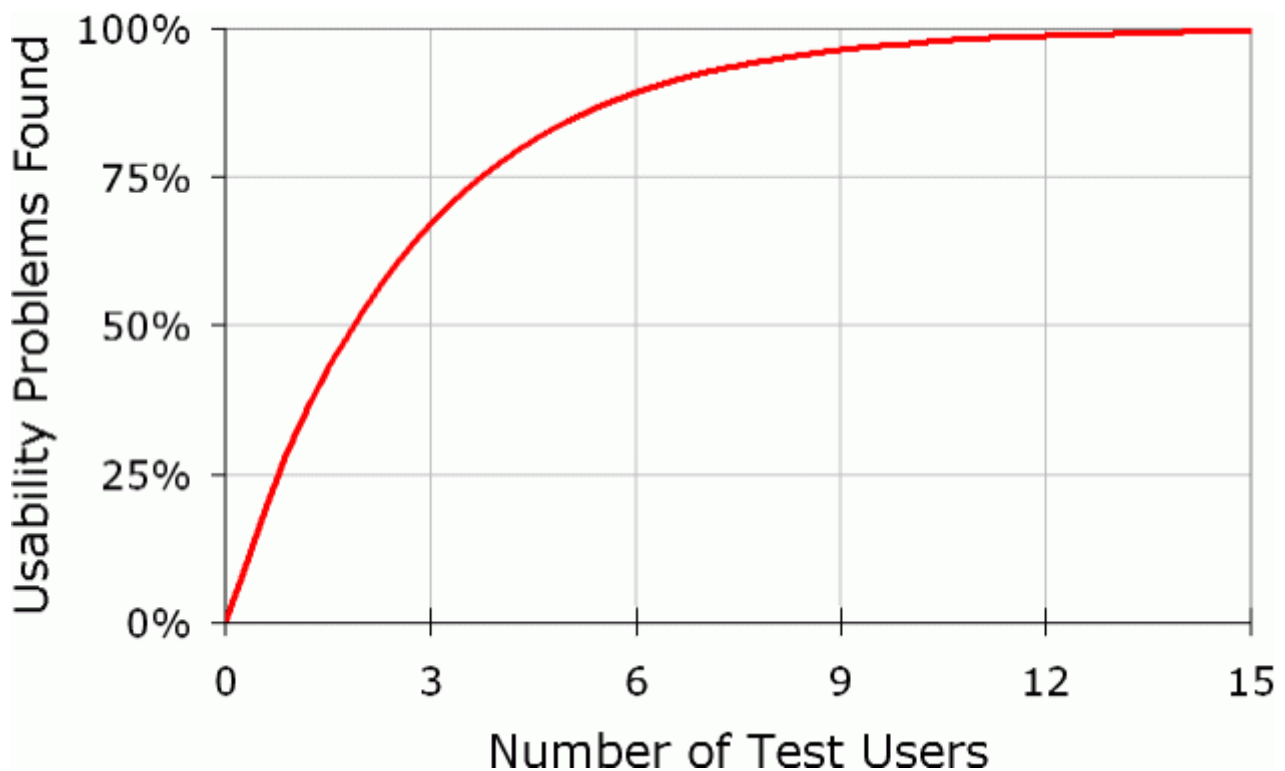
whose methods were poorly applied, damaging the quality of data obtained from the tests and extending the length of the design process exponentially. Before I can explain what went wrong and what we should have done better, I need to address the reason for the RITE Method's enduring popularity: when performed correctly, it's not only more efficient but it simply makes sense.

The RITE Method

The RITE Method was first outlined in 2002 by Michael Medlock and his team at Microsoft Games Studios while developing *Age of Empires II*. As the abstract of the team's white paper asserts, rapid iterative testing “leads to a high ratio of problems found to fixes made and then empirically verifies the efficacy of the fixes.” This is because the central idea of the RITE Method is that you test on small numbers of users at a time, identify usability issues from the first studies, devote some time to fix them, and then test again on a few more users with your higher quality product. Traditional user research would suggest testing the same prototype on a large number of users, on the order of 100 or more, so that usability issues can be validated by a statistical analysis of the large sample size. By comparison, RITE studies are lean, inexpensive, and provide almost the same quality of feedback as the more inert traditional methods.

This is because the RITE Method builds upon one of the foundational recommendations from the Nielsen Norman Group: you only need to test with 5 users. According to the NNG report, published by Jakob Nielsen in 2000, there are diminishing returns as you scale the number of participants in a research study. Straightforwardly, you won't find any usability problems if you test zero users. Your first test participant will find a certain number of usability problems and each subsequent participant will find even more, but many of those problems will overlap with the participants that came before them. Plotting out these diminishing returns indicates that you can identify more than 80% of the usability problems in a product with just the first five test users. And if that's the case, then it's not worthwhile to spend time and money testing even more users when you could instead take those 80% of your usability issues, *fix them*, then test the product on another small group of users. Whereas testing 15 users can identify nearly all of the usability issues of your v1



[Open in app](#)[Get started](#)

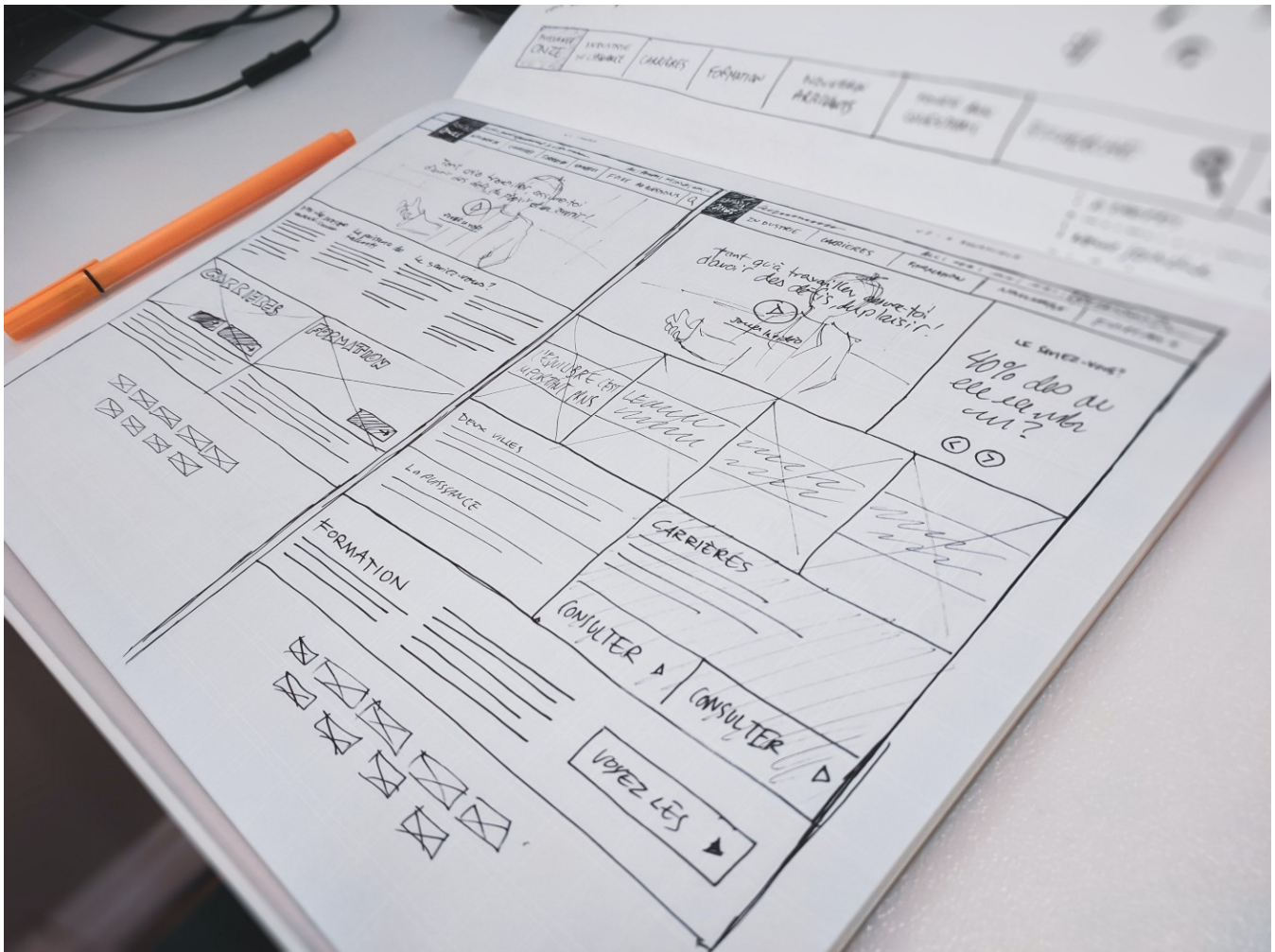
Plot of usability problems found after each successive user test, per a formula developed by Jakob Nielsen and Tom Landauer. [Source](#).

Thus, the RITE Method recommends confining tests to 4–5 users at a time, giving the design or development team time to fix the first issues found, and then test again through several iterations.

What went RONG

Earlier in my career, I was pulled into a project where user testing was already underway. The user tests were being performed by PMs, and while they weren't necessarily conducting a RITE study by the book, they had arranged the studies such that they tested the same group of five customers once a week for three consecutive weeks. The idea was to increment improvements based on each week's feedback and then test the same users again with an improved prototype. The PM facilitating the interviews created the Round 1 prototype, a few hand-drawn sketches with very little detail, and proceeded to conduct the first round of tests.



[Open in app](#)[Get started](#)

I can't share the actual hand sketches we used, but they were similar to this. Photo by [Sigmund](#) on [Unsplash](#)

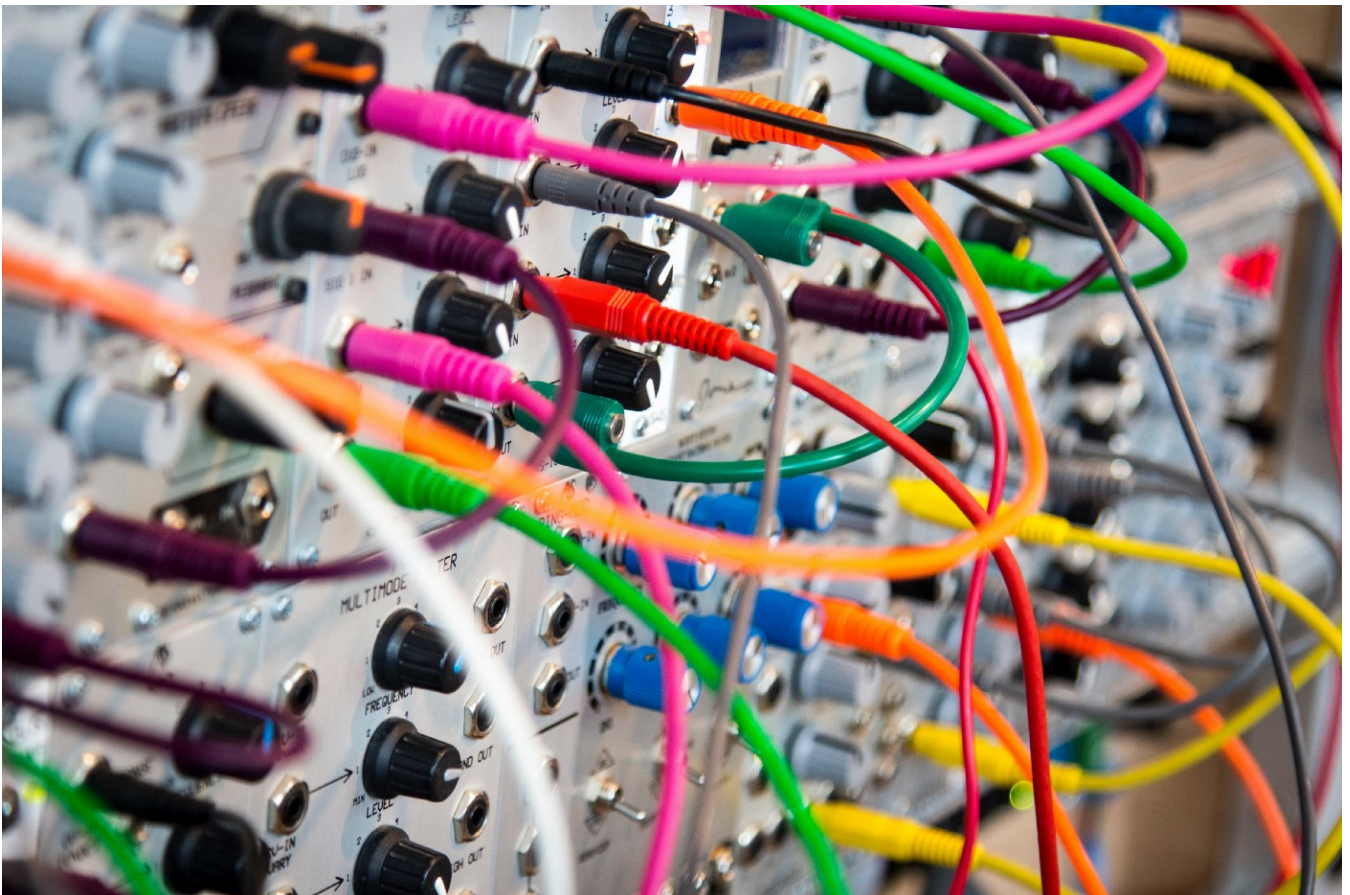
This is when they brought me in to the project. With the Round 1 tests mostly complete, they wanted my help to take the hand-drawn sketches, recreate them in Figma, and at the same time improve some of the early flaws detected from the first round of walk-throughs. In theory, this was simple enough; the challenge was staging the recreated screens before the Round 1 debrief Friday morning and working through the list of improvements in time to deploy Round 2 by Monday morning.

Round 2 brought us more feedback, but the feedback was distinctly different than that from Round 1. Because we pivoted from sketches to Figma, and for expedience sake the Figma mockups were made using high-fidelity library buttons and controls, **customers expected a higher fidelity of functionality to match these screens.** Rather than test the narrow features that had been highlighted in the hand sketches,



[Open in app](#)[Get started](#)

Since I was acting in a guest capacity on this project, I didn't feel comfortable pushing back as we pivoted to Round 3. The product team, hearing feedback for more realistic product flows, requested I round out the variety of clickpaths the customers could take during their explorations. With a sigh and an eyeroll that I hid to myself, I worked my magic, producing a 20+ frame prototype that anticipated a variety of clicks a participant could ask about.



Before I realized it, our focus had shifted from testing a few simple concepts to making sure we closed every possible loop. Photo by [John Barkiple](#) on [Unsplash](#)

We proceeded to Round 3, where customers were quick to point out how much more professional the screens looked. But after the first two customers failed to find their way through the tasks, it became apparent that the actual purpose of this product wasn't as clear as our team had hoped. Before the third test, I made a series of tweaks (that now needed to be replicated across 20+ screens) and we proceeded to test again. We had now changed the wording of some features as suggested by the earlier tests,



[Open in app](#)[Get started](#)

What's the damage?

The direction of the product — and of these tests — seemed to spiral outside of my control. Because of the heavy changes we opted to make between tests, later tests ended up getting postponed, turning this three-week venture into five. And because we got unsatisfactory feedback from Round 3, we extended a few of the participants into a Round 4 and 5.

These were the only user tests we would end up performing for this product, but we ultimately spent six more months iterating on the designs. Engineering feasibility dictated a number of directional changes, as well as a revolving door of features added and removed by the product team, and ultimately my Figma file contained 14 versions of incremental prototypes. The product would eventually be deployed, but it looked very different than what we had placed in front of our interviewees.

What did we learn?

This RITE study started out with so much promise, but we missed some opportunities to stay aligned to the RITE methodology and didn't reap the efficiency that the method promises. I conducted a post-mortem and determined that a successful RITE study needs to:

- **Develop a hypothesis to test.** Like all scientific method experiments, it's okay if the hypothesis is wrong, but if you don't have something specific to prove, you won't prove anything.
- **Narrow the scope of the test.** Rapid testing works best to test a high-level concept. It takes a long time to design robust UIs where every button has a function and a flow, but most of those functions aren't important to test. We let users explore around with open-ended questions and ultimately identified the half-baked peripheral elements on the page rather than validate the product itself.
- **Commit to a single design for all interviewees.** By vastly changing the UI halfway between our test participants, we had less consistent, less valid data as some participant feedback was for the older design and some was for the new



[Open in app](#)[Get started](#)

• • •

User research is a key element in the UX process, but it is often conducted by a variety of professionals in a variety of methods. It can be easy to forget that user research is a science and it benefits from the scientific method just as much as anything else. This means that in order to be successful, **a research study needs structure**. It needs a clear hypothesis that can be proven or disproven and it needs to manipulate only one variable at a time. After all, our goal is to let data drive our decision-making; we can only do that by committing to the validity of the data we collect.

Sign up for The UX Collective Newsletter

By UX Collective

A weekly, ad-free newsletter that helps designers stay in the know, be productive, and think more critically about their work. [Take a look.](#)

[Get this newsletter](#)