

One of the topics that create much confusion in the software world is the difference between usability testing (UT) and user acceptance testing (UAT). User experience (UX) professionals only worry about usability testing, while developers and Q&A people focus on user acceptance testing. Other stakeholders use both terms interchangeably believing in the sentiment of, “It doesn’t matter what you call it — as long as we test our software with users.”

In fact, both tests involve end users utilizing software with the goal of finding shortcomings. However, the flaws being identified differ, and with that the purpose of why you want to execute each test.

Further, the point in time when the tests are conducted may also differ depending on the development process that is being used. So let us try to untangle all of this.

In both usability testing and user acceptance testing end users engage with a product and go through certain test scenarios. Observing test users succeed or fail with test tasks and hearing their comments about the product itself provides insights into the quality of the product.

UT is concerned with understanding the user experience that manifests itself when a user engages with a product or concept. UX is a comprehensive concept describing and measuring the objective and subjective effectiveness and efficiency of the interaction — that is to what degree users can achieve objectives and how much effort they must put forth.

UX also includes the psychological impact of the user-product interaction. Is it perceived as comfortable, joyful, stressful, confusing or straight-forward? During usability tests qualitative and quantitative aspects are being assessed. Qualitative aspects encompass the sentiment of test users; their comments and reactions about the product.

An example is a finding “Test user is surprised by the ‘Close’ button saving their settings in the modal window.” Quantitative aspects are measurable findings, like task success rates, task completion times, number of user errors committed during task execution, etc. For instance, a finding may be: “85% of test users completed test task #1 in under 1 minute.”

UAT is concerned with understanding if the product does what it is supposed to be doing and that its target audience in fact can achieve their objectives. In other words: UAT checks whether defined business requirements are being met by the product. This means that the right functions are built in, and that the code is faultless.

Usability testing

Does the product produce a good user experience?

- Does it allow users to effectively, efficiently and comfortably achieve certain objectives?

User acceptance testing

Is the product fit for purpose?

- Does it enable users to achieve certain objectives?



In most cases usability testing is used formatively — it entails a rework of the product to mitigate identified shortcomings. The testing is part of the iterative design and development process and ensures that usability defects are uncovered as early as possible and fixed before the product is released to the market.

The role of user acceptance testing on the other hand is to verify that the product to be released serves its stated purpose, and that the code is faultless, thus allowing target users to successfully work with it. It provides the company that builds the product a level of assurance that what they are about to release is delivering the capabilities that it was built to provide.

All usability-testing studies involve a participant performing some assigned tasks on one or more designs. There are, however, two types of data that can be collected in a user-testing study:

**Qualitative (qual) data**, consisting of observational findings that identify design features easy or hard to use

**Quantitative (quant) data**, in form of one or more metrics (such as task completion rates or task times) that reflect whether the tasks were easy to perform

**Qualitative data** offer a direct assessment of the usability of a system: researchers will observe participants struggle with specific UI elements and infer which aspects of the design are problematic and which work well. They can always ask participants followup questions and change the course of the study to get insights into the specific issue that the participant experiences. Then, based on their own UX knowledge and possibly on observing other participants encounter (or not) the same difficulty, researchers will determine whether the respective UI element is indeed poorly designed.

**Quantitative data** offer an indirect assessment of the usability of a design. They can be based on users' **performance** on a given task (e.g., task-completion times, success rates, number of errors) or can reflect participants' **perception of usability** (e.g., satisfaction ratings). Quantitative metrics are simply numbers, and as such, they can be hard to interpret in the absence of a reference point. For example, if 60% of the participants in a study were able to complete a task, is that good or bad? It's hard to say in the absolute. That is why many quant studies usually aim not so much to describe the usability of a site, but rather to compare it with a known standard or with the usability of a competitor or a previous design.

	Qual Research	Quant Research
Questions answered	Why?	How many and how much?
Goals	<p>Both formative and summative:</p> <ul style="list-style-type: none"> <li>• inform design decisions</li> <li>• identify usability issues and find solutions for them</li> </ul>	<p>Mostly summative:</p> <ul style="list-style-type: none"> <li>• evaluate the usability of an existing site</li> <li>• track usability over time</li> <li>• compare site with competitors</li> <li>• compute ROI</li> </ul>
When it is used	Anytime: during redesign, or when you have a final working product	When you have a working product (either at the beginning or end of a design cycle)
Outcome	Findings based on the researcher's impressions, interpretations, and prior knowledge	Statistically meaningful results that are likely to be replicated in a different study
Methodology	<ul style="list-style-type: none"> <li>• Few participants</li> <li>• Flexible study conditions that can be adjusted according to the team's needs</li> <li>• Think-aloud protocol</li> </ul>	<ul style="list-style-type: none"> <li>• Many participants</li> <li>• Well-defined, strictly controlled study conditions</li> <li>• Usually no think-aloud</li> </ul>







