

Usability test quality heuristics

Some signs that a usability test will produce useful results.

This is not an exhaustive list, but can be helpful in planning and assessing usability tests — additions welcome!

Indicator	Good sign	Bad sign
Study purpose	Researcher clearly articulates the purpose of the usability test (for example, as a specific question to answer, an area of inquiry, etc.).	Researcher defines no clear purpose, or the purpose is very broad like “testing the app” or “finding problems.”
Team participation	Multiple members of the product team observe the usability test, but there’s only one moderator. Tests are recorded for sharing with absent stakeholders.	Only one person from the product team participates in (leads) the usability test. Tests are not recorded.
Observer role	Observers are generally quiet, and ask open-ended follow-up questions after each test has concluded.	Observers attempt to sell or explain the product. Observers openly debate the participant’s feedback in their presence.
Number of participants	Tests include a sufficient number of participants necessary to see patterns, and there’s only one participant per session.	Studies includes only one participant, or studies include multiple participants tested simultaneously (like a focus group ).
Sampling bias	Participants are people who actually use the application (or do the task the application is supposed to support).	Participants are experts (vs. average users), friends, family, or close colleagues of the product team.
Task framing	Researcher asks participants to complete	Researcher asks participants to complete

	tasks with the product or service that align with their previously disclosed goals.	tasks which are unrelated to their goals. For example, asking a participant how they might send a fax when their job doesn't call for that.
Researcher provided context	Researcher's wording of the tasks resembles what participants would actually say. (e.g. "I want to travel for the government" instead of "I want to log in to ConcurGov via GSA SSO.")	Researcher's wording of the tasks doesn't resemble what participants would actually say or think. It's too detailed or implies too much knowledge about the system.
Researcher priming	Researcher asks participants to assess usability without indicating the nature or location of functionality. For example, "How would you view the status of your form?"	Researcher guides participants on task accomplishment. For example, "Which of the links in the header would you click to login?"
Task variability	Researcher has a goal and each session tries to answer the same research questions (although the conversations with participants may vary substantially).	The researcher tries to answer different research questions with each session.
Question style	Researcher asks follow up questions that are open-ended and non-leading.	Questions are leading or subtly suggest potential solutions or hypotheses.
Synthesis style	Researcher and observing team jointly records observations immediately following each interview. Researcher looks for patterns across multiple participants and surfaces problems that affected several people.	Researcher reports the most memorable problems without conducting affinity mapping or some other analysis technique.
Presentation of findings	Researcher reports findings to team and stakeholders in an easy to follow, well prioritized way.	Researcher presents team a "basket of issues" or an unprioritized laundry list of potential changes.

Incorporation of findings	Product team translates findings into future user stories or product refinements.	Researcher reports do not affect the product backlog or ongoing development work.
---------------------------	---	---