



World Leaders in Research-Based User Experience

Formative vs. Summative Evaluations

Summary: Formative evaluations are used in an iterative process to make improvements before production. Summative evaluations are used to evaluate a shipped product in comparison to a benchmark.

By [Alita Joyce](#) on July 28, 2019

Topics: [Research Methods](#)

In the user-experience profession, we preach [iteration](#) and evaluation. There are two types of evaluation, formative and summative, and where you are in the design process determines what type of evaluation you should conduct.

Formative evaluations focus on determining which aspects of the design work well or not, and why. These evaluations occur throughout a redesign and provide information to incrementally improve the interface.

Let's say we're designing the onboarding experience for a new, completely redesigned version of our mobile app. In the design process, we [prototype](#) a solution and then test it with ([usually a few](#)) users to see how usable it is. The study identifies several issues with our prototype, which are then fixed by a new design. This test is an example of formative evaluation — it helps designers identify what needs to be changed to improve the interface.

Formative evaluations of interfaces involve testing and changing the product, usually multiple times, and therefore are well-suited for the redesign process or

while creating a new product.

In both cases, you iterate through the prototyping and testing steps until you are as ready for production as you'll get (even more iterations would form an even better design, but you have to ship at some point). Thus, formative evaluations are meant to steer the design on the right path.

Summative evaluations describe how well a design performs, often compared to a benchmark such as a prior version of the design or a competitor. Unlike formative evaluations, whose goal is to inform the design process, summative evaluations involve getting the big picture and assessing the overall experience of a finished product. Summative evaluations occur less frequently than formative evaluations, usually right before or right after a redesign.

Let's go back to our mobile-app example. Now that we've shipped the new mobile app, it is time to run a study and see how our app stands in comparison to the previous version of the app. We can gather the time on task and the success rates for the core app functionalities. Then we can compare these metrics against those obtained with the previous version of the app to see if there was any improvement. We will also save the results of this study to evaluate subsequent major versions of the app. This type of study is a summative evaluation since it assesses the shipped product with the goal of tracking performance over time and ultimately [calculating our return on investment](#). However, during this study, we might uncover some usability issues. We should make note of those issues and address them during our next design iteration.

Alternatively, another type of summative evaluations could compare our results with those obtained with one or more competitor apps or with known industry-wide data.

All summative evaluations paint an overview picture of the usability of a system. They are intended to serve as reference points so that you can determine whether you're improving your own designs over time or beating out a competitor.

The ultimate summative evaluation is the go/no-go decision of whether to release a product. After all is said and done, is your design good enough to be inflicted on the

public, or do we think that it will harm our brand so badly that it should never see the light of day? It’s actually rare for companies to have a formal process to kill off bad design, which may be why we encounter many releases that do more harm than good for a brand. If you truly embrace our proposition that [brand is experience](#) in the digital age, then consider a final summative evaluation before release.

Origin of the Terms

The terms ‘formative’ and ‘summative’ evaluation were coined by Michael Scriven in 1967. These terms were presented in the context of instructional design and education theory, but are just as valuable for any sort of evaluation-based industry.

In the educational context, formative evaluations are ongoing and occur throughout the development of the course, while summative evaluations occur less frequently and are used to determine whether the program met its intended goals. The formative evaluations are used to steer the teaching, by testing whether content was understood or needs to be revisited, while summative evaluations assess the student’s mastery of the material.

When Each Type of Evaluation Is Used

Recall that formative and summative evaluations align with your place in the design process. Formative evaluations go with prototype and testing iterations throughout a redesign project, while summative evaluations are best for right before or right after a major redesign.

Great researchers begin their study by determining what question they’re trying to answer. Essentially, your research question is the same as the type of evaluation. Below is a list of possible research questions you might have and the corresponding evaluation. For that reason, this table is descriptive, not prescriptive.

Questions you might ask	Type of evaluation

How is our interface performing compared to our competitors?	Summative
What usability issues exist in our interface?	Formative, Summative
How does our interface compare to the industry benchmark?	Summative
Do users understand our navigation?	Formative
How has our overall experience changed over time?	Summative
Does our interface comply with recognized usability principles?	Formative
Is this product good enough to launch? (Go/no-go decision)	Summative

Research Methods for Formative vs. Summative Evaluations

After it is clear which type of evaluation you will conduct, you have to determine which [research method](#) you should use. There is a common misconception that summative equals quantitative and formative equals qualitative — this is not the case.

Summative evaluations can be either qualitative or quantitative. The same is true for formative evaluations.

Although summative evaluations are often quantitative, they can be qualitative studies, too. For example, you might like to know where your product stands compared with your competition. You could hire a UX expert to do an [expert review](#) of your interface and a competitor's. The expert review would use the [10 usability heuristics](#) as well as the reviewer's knowledge of UI and human behavior to produce a list of strength and weaknesses for both your interface and your competitor's. The

study is summative because the overall interface is being evaluated with the goal of understanding whether the UX of your product stands up to the competition and whether a major redesign is warranted.

Additionally, formative evaluations aren't always qualitative, although that is often the case. (Since it's recommended to run an extended series of formative evaluations, it makes financial sense to use a cheaper qualitative study for each of them.) But sometimes big companies with large UX budgets and high level of [UX maturity](#) might use quantitative studies for formative purposes in order to ensure that a change to one of their essential features will perform satisfactorily. For instance, before launching a new homepage design, a large company may want to run a quantitative test on the prototype to make sure that the number of people who will [scroll below the fold](#) is high enough.

Conclusion

Formative and summative evaluations correspond to different research goals. Formative evaluations are meant to steer the design on the correct path so that the final product has satisfactory user experience. They are a natural part of any iterative user-centered design process. Summative evaluations assess the overall usability of a product and are instrumental in tracking its usability over time and in comparing it with competitors.

For more on formative and summative evaluations, check out our courses like [Usability Testing](#) or [Measuring UX and ROI](#).

References

Greenstein, Laura. *What Teachers Really Need to Know About Formative Assessment*. ASCD, 2010.

Share this article: [Twitter](#) | [LinkedIn](#) | [Email](#)

Copyright © 1998-2021 Nielsen Norman Group, All Rights Reserved.