

Search[Home](#)[Articles](#)[Training & Events](#)[Consulting](#)[Reports & Books](#)[About NN/g](#)
Search

Between-Subjects vs. Within-Subjects Study Design

by [Raluca Budiu](#) on May 13, 2018

Topics: [Research Methods](#) [User Testing](#)

Summary: In user research, between-groups designs reduce learning effects; repeated-measures designs require fewer participants and minimize the random noise.

When you want to compare several user interfaces in a single study, there are two ways of assigning your test participants to these multiple conditions:

- **Between-subjects** (or **between-groups**) study design: different people test each condition, so that each person is only exposed to a single user interface.
- **Within-subjects** (or **repeated-measures**) study design: the same person tests all the conditions (i.e., all the user interfaces).

(Note that here we use the word “design” to refer to the *design of the experiment*, and not to website design.)

[Settings](#)

For example, if we wanted to compare two car-rental sites A and B by looking at how participants book cars on each site, our study could be designed in two different ways, both perfectly legitimate:

1. Between-subjects: Each participant could test a single car-rental site and book a car only on that site.
2. Within-subjects: Each participant could test both car-rental sites and book a car on each.

Any type of user research that involves more than a single test condition has to determine whether to

Experimental Design in Quantitative Studies

Unlike qualitative studies, quantitative usability studies aim to result in findings that are statistically likely to generalize to the whole user population. How the data from such studies is analyzed depends on the way in which the study was designed (that is, on the study's **experimental design**).

Often, the main goal of quantitative usability studies is to compare — a site with its competitors, two different iterations of a design, or two different groups of users (such as experts vs. novices). Like in any scientific experiment in which we want to detect causal relationships, a quantitative study involves **two types of variables**:

- **Independent variables**, which are directly manipulated by the researcher
- **Dependent variables**, which are measured (and expected to vary as a result of the independent-variable manipulation)

(If the study produces statistically significant results, then we can say that a change in the independent variable *caused* a change in the dependent variable.)

Let's go back to our original car-rental example. If we wanted to measure which of the two sites, A or B, is better for the task of reserving a car, we could choose *Site* (with two possible values or *levels* — A and B) as the independent variable, and the time on task and the accuracy for booking a car could be the dependent variables. The goal of the study would be to see whether the dependent variables (time and accuracy) change when we vary the site or they stay the same. (If they stay the same, then none of the sites is better than the other.)

To plan our study, the next question that we need to answer is whether the study design should be between-subjects or within-subjects — that is, whether a participant in the study should be exposed to all the different conditions for the independent variable in our study (within-subjects) or only to one condition (between-subjects). The choice of experimental design will affect the type of statistical analysis that should be used on your data.



This website uses cookies

We rely on cookies to play videos, display tweets, remember your preferences, and to analyze our website traffic. You consent to our cookies if you click "OK".

OK

Settings

**Within-subjects design**

The same participant tests all conditions corresponding to a variable.

**Between-subjects design**

Different participants are assigned to different conditions corresponding to a variable.

NNGROUP.COM NN/g

It is possible that an experiment design is both within-subjects and between-subjects. For example, assume that, in the case of our car-rental study, we were also interested in knowing how participants younger than 30 perform compared with older participants. In this case we would have two independent variables:

- Age, with 2 levels: under 30, over 30
- Site, with 2 levels: A and B

For the study, we will recruit an equal number of participants in each age group. Let's assume that we decide that each participant, whether under or over 30, will make a car-rental reservation on both site A and on site B. In this case, the study is within-subjects with respect to the independent variable *Site* (because each person sees both levels of this variable — that is, both site A and site B). However, the study is between-subjects with respect to *Age*: one person can only be in a single age group (either under or over 30, not both). (Well, technically, you could pick a group of under-30-year olds and wait until they turn 30 to have them test the sites again, but this setup is highly impractical for most real-world situations.)

This website uses cookies

We rely on cookies to play videos, display tweets, remember your preferences, and to analyze our website traffic. You consent to our cookies if you click "OK"

Some independent variables may impose the choice of design. *Age* is one of them, as seen above.

Others are *Expertise* (if we want to compare experts and novices), *User Type* (if we want to compare different user groups or personas — for example, business traveler vs. leisure traveler), or *Gender* (assuming that a person cannot be of several genders at the same time). Outside usability, drug trials are one common case of between-subject design: participants are exposed to only one treatment:

either the drug being tested or a placebo, not both. And sometimes the manipulation itself changes the state of the participant: for example, if you want to see which of two curricula is more effective for teaching reading, you could not have the same student be exposed to both, because once she's

Which Is Better: Between-Subjects or Within-Subjects?

Unfortunately, there is no easy answer to this question. As seen above, sometimes your independent variables will dictate the experimental design. But in many situations, both designs may be possible.

- **Between-subjects minimizes the learning and transfer across conditions.** After a person has completed a series of tasks on a car-rental site, she is more knowledgeable about the domain than she was before. For example, she may now know that car-rental sites charge an extra fee for drivers under 21, or what a collision-damage waiver is. That knowledge will likely help her become more efficient on a second car-rental site, even though that second site may be very different from the first.
With between-subject design, this transfer of knowledge is not an issue — participants are never exposed to several levels of the same independent variable.
- **Between-subjects studies have shorter sessions than within-subject ones.** A participant who tests a single car-rental site will have a shorter session than one who tests two. Shorter sessions are less tiring (or boring) for users, and can also be more appropriate for remote unmoderated testing (especially since tools like UserZoom usually require a fairly short session length).
- **Between-subject experiments are easier to set up, especially when you have multiple independent variables.** When the study is within-subjects, you will have to use randomization of your stimuli to make sure that there are no order effects. For example, in our car-rental study, we need to make sure that participants don't always start with site A and then move on to site B. The order of the sites needs to be random for each participant. This is easy with just two sites: randomly assign 50% of users to start with each site. But as you increase the number of independent variables and of levels for an independent variable, randomization becomes more difficult to implement within some of the existing platforms for quantitative usability testing.
- **Within-subject designs require fewer participants and are cheaper to run.** To detect a statistically significant difference between two conditions, you'll often need a fair number of a data points (often above 30) in each condition. If you have a within-subject design, each participant will provide a data point for each level of the independent variable. For our car-rental study, 30 participants will provide data points for both sites. But if the study is between-subjects you will need twice as many to get the same number of data points. That means twice the cost.
- **Within-subjects design minimize the random noise.** Perhaps the most important advantage of within-subject designs is that they make it less likely that a real difference that exists between your conditions will stay undetected or be covered by random noise.

Individual participants bring in to the test their own history, background knowledge, and context. One may **IN/g** after a long night of partying, another one may be bored, yet another one may have received a great news just before the study and be happy. If the same participant interacts with all levels of a variable, she will affect them in the same way. The happy person will be happy on both sites, the tired one will be tired on both. But if the study is between-subjects, the happy participant will only interact with one site and may affect the final results. You'll have to make sure you get a similar happy participant in the other group to counteract her effects.

In practice, researchers won't be able to assess such differences between participants — although they may match the gender, the experience, and the age across groups, it will be difficult to predict or detect other factors specific to each participant.

Randomization: Essential for Both Types of Design

Whether your experimental design is within-subjects or between-subjects, you will have to be concerned with randomization, although in slightly different ways.

Above, we discussed why randomization is important in within-subject designs: it counteracts the possible order effects and minimizes transfer and learning across conditions.

For between-subject designs, you must make sure that participants are allotted randomly to conditions, because you want to ensure that your participant assignment does not affect your study results. Thus, if a researcher decides that all the participants that he likes should interact with site A and then he finds that site A performed better than site B, he won't know whether he's discovered a true difference between the sites or whether the result simply reflects his assignment (for example, because people who sense that they are liked tend to return the favor, and may be more patient or have a positive mindset during the test).

Even without such an obvious bias as your personal preferences, it's easy to get randomization wrong. Say that you run a study across four days, Saturday through Tuesday. You might decide to have the first half of the test users start with site A and have the second half of the users start with site B. However, this is not a true randomization, because it's very likely that certain types of people are more likely to agree to a study during the weekend and other types of people are more likely to sign up for your weekday testing slots.

Conclusion

User research can be between-subjects or within-subjects (or both), depending on whether each participant is exposed to only one condition or to all conditions that are varied within a study. Each of these types of experimental design has its own advantages and disadvantages; within-subjects design requires fewer participants and increases the chance of discovering a true difference among your conditions; between-subjects designs minimize the learning effects across conditions, lead to shorter sessions, and may be easier to set up and analyze.



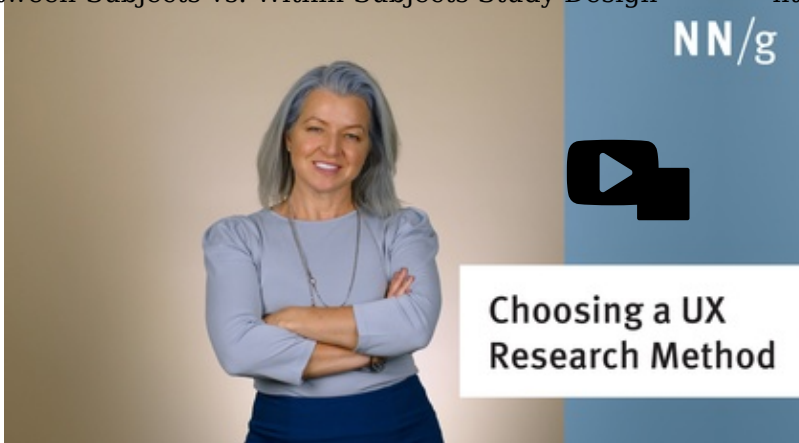
Videos

This website uses cookies

We rely on cookies to play videos, display tweets, remember your preferences, and to analyze our website traffic. You consent to our cookies if you click "OK".

OK

Settings



When to Use Which UX Research Method

5 minute video



5-Second Usability Test

3 minute video



Small vs. Big User Studies — What's Best?

2 minute video

OK

Settings

Copyright © 1998-2019 Nielsen Norman Group, All Rights Reserved.