

## 20.4: Designing tests

To design tests successfully, you need to know what you can test, how you can test, and what sort of time periods you are looking at for testing. If it's your first time doing conversion optimisation, you should start with simple and quick tests, to get a feel for the process before embarking on more complicated tests.

### Types of tests

When we talk about conversion rate optimisation, we are usually referring to running split tests. A split test is one where we show different versions of a web page to groups of users and determine which one has performed better.

We can run A/B tests. Here each version being tested is different from all the others. A/B tests always involve just two versions of what is being tested: the original and an alternative.

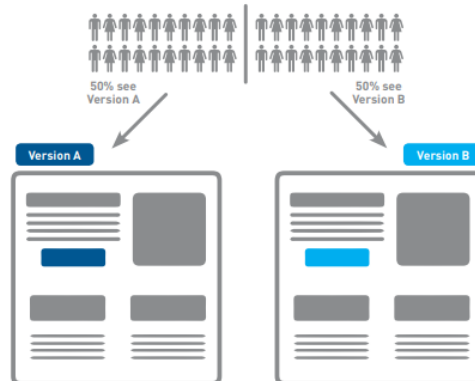


Figure 20.4.1: A/B testing explained visually *Adapted From Explore Web Solutions, 2013*

A/B tests are ideal for an initial foray into conversion optimisation, as they can be easy to set up. If you are running just one alternate and the original, it can also mean that you get a quicker result. When conducting A/B testing, you should only change one element at a time so that you can easily isolate what impact each factor has on your conversion rate.

We can also run multivariate tests, sometimes referred to as MVTs. Here, a number of elements on a page are tested to determine which combination gives the best results. For example, we may test alternative headlines, alternative copy and alternative call to action buttons. Two versions of three elements mean that we are testing eight combinations!

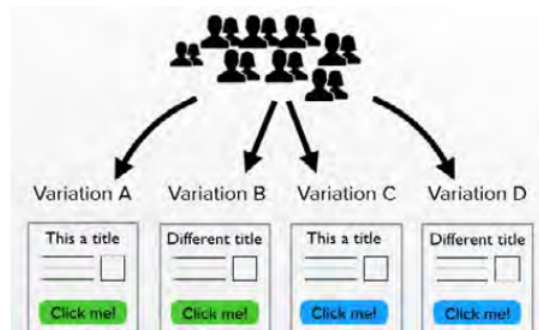


Figure 20.4.2: Multivariate testing combines a variety of elements *Adapted From Working Homeguide, 2014*

Multivariate tests can be more complicated to set up, but allow you to test more elements at once. Multivariate tests are ideal when you have large traffic volumes. If traffic volumes are not very high, it can take a very long time to reach a statistically significant result, especially if there are many combinations being tested.

### Length of tests and sample size

Several factors determine which tests you can run. Relatively simple calculations help you to determine how long a test is likely to take, which is based on the number of participants as well as the improvement in conversion rate. We've included some sample size calculators in [Tools of the trade](#), in section 20.6 of this chapter.

## Number of participants

The number of participants in the test is determined by how many users actually see the page being tested, as well as what percentage of your potential customers you want to include in your test.

The number of users who see the page being tested may not be the same as the number of visitors to your website. You'll need to use your data analytics to determine the number of users viewing that specific page. Of course, if you are running advertising campaigns to direct traffic to the page being tested, you can always spend a bit more money to increase the number of users coming to that page.

You also want to determine what percentage of users will be involved in the test. In a simple A/B test, if you include 100% of your visitors in the test, 50% will see version A and 50% will see version B. If you include only 50% of your visitors in the test, this means that 25% of your overall visitors will see version A, and 25% will see version B. Including 100% of your visitors will give you results more quickly. However, you may be concerned that your alternative version could perform worse, and you don't want to compromise your performance too much.

## Change in conversion rate

While this is not something you will know upfront, the percentage change in conversion rate also affects the length of a test. The greater the change, the more quickly a statistically significant decision can be made.

## Number of variations

The more variations you have, the longer it will take to determine which combination performs the best.

These factors can then be used to calculate the suggested length of time for a test to run. There are several online calculators that do this for you. A good one to try is this one, offered by Visual Website Optimizer: [visualwebsiteoptimizer.com/ab-splittest-duration](http://visualwebsiteoptimizer.com/ab-splittest-duration).

Estimate Test Duration	Scenario 1	Scenario 2	Scenario 3	Scenario 4
# Test combinations	2	2	2	3
# Page views per day	200	200	300	350
% Visitors in experiment	100	100	50	50
% Current conversion rate	30	30	30	30
% Expected improvement	20	40	20	30
Days Duration	4.01	1.04	5.35	4.22

Figure 20.4.3: Small changes can affect your online testing *Adapted From Stokes, 2013*

It is usually preferable to test bigger changes or variations, rather than very small changes, unless you have a very large audience.

## Designing for analysis

The purpose of running tests is to improve performance. To do this, you analyse your results against what you expected to find and then choose the option that performed better. This sounds simple, but how do you know what counts as better enough to warrant a change? Is it one more click than the other option, three more clicks, or should one perform 25% better than the other? You also need to think about chance: how certain are you that the differences in your results were not just coincidental? These can be tricky questions.

To determine which option in your split test did better, set parameters and assess the statistical significance of your results. In statistics, we create a null hypothesis. For split tests, the null hypothesis is that there is no difference between the performance of the two options and any difference recorded is due to chance. You then use statistics to calculate the p value, which shows whether the difference was likely due to chance (or not). If the difference is significant, it is probably not due to chance. Generally, to be significant, the p value should be less than or equal to 0,05, indicating a less than 5% probability that the difference in performance between the two options was due to chance.

You do not have to be able to perform the complex statistical calculations. Handy tools like VWO's split test significance calculator will do this for you. All you have to do is enter the number of users that visited your control version and the number of conversions, as well as the number of users that visited the variation and the number that converted. The calculator then provides your p value and states whether your test is significant enough to change to the variation. You can find the calculator here:

<https://vwo.com/ab-split-test-signif...ce-calculator/>

Are your results significant?

	Control	Variation
Number of Visitors	250	250
Number of conversions	25	42

CALCULATE SIGNIFICANCE

P-Value  
(range from 0 to 1)

0.012

Significance

Yes!

Figure 20.4.4: The results from VWO's significance calculator *Adapted From VWO, 2017*

When designing your tests, it is important to consider the null hypothesis and set the parameters for significance. By keeping these aspects in mind, you will develop tests that allow for clearer and easier analysis, which will make the whole testing process that much more effective.

This page titled [20.4: Designing tests](#) is shared under a [CC BY-NC-SA 3.0](#) license and was authored, remixed, and/or curated by [Rob Stokes](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.