



WHITE PAPER



A/B TESTING

Comparing Data

October 25, 2007

Version 4.0



1 What is A/B Testing?

A/B testing is a means by which you can compare data sets from two different pages, banners, products, etc. A/B testing is similar to the eye exams you take at the ophthalmologist. The doctor will have you look through the scope and then ask, “Which one of the following is clearer? A or B?” After choosing between the two, the doctor would make a small adjustment and then ask again. Each time an adjustment is made, the result is a little better to the point that your vision is as perfect as possible.

In another example, imagine that you own a company that specializes in baking cookies. You have a solid customer base that regularly buys your cookies, but you would like to improve upon your cookie recipe and make a few changes. The worst thing you could do is to just replace the old recipe without verifying the new recipe is better. You don’t want to risk losing your customer base if the new cookie doesn’t taste better than the original. In order to refine and prove the new cookie is better, you would create sample groups or taste testing groups to test the new cookies and help you determine whether you should replace them or not.

Internet marketing and business are very much like this cookie company. Many companies make both subtle and major changes to their public web sites and marketing without properly validating that the changes improve upon the original.

As you start to make changes to your campaigns and web site, you can use sample groups, or taste testers) to help you verify that your changes are returning the desired effect. Omniture offers you the opportunity to test the changes to your web site by conducting A/B testing. The following list contains some main steps that are included in A/B testing.

- Determine areas of the site that have opportunity to positively affect conversion
- Hypothesize on possible changes using the current data to validate your conclusion
- Test your hypothesis by running multiple versions of a page
- Use data to tell you which page had the best conversion ratio

Let’s use Omniture as an example. Omniture’s corporate site (www.omniture.com) is constantly running A/B testing on its internal campaigns. Omniture’s site is a lead generation site, so completing a lead form is the success event. On Omniture’s home page is a large area dedicated to internal campaigns. This area houses offers for those visitors that complete the lead form. This offer might be a new white paper on media tracking or another on how to increase conversions. Once visitors complete the lead form, they can download the white paper.

Omniture often conducts A/B tests on this area of the site. Different sample groups will see different versions of the campaign. The visitors that act upon these campaigns are tracked using an eVar. Then, Omniture can run an analysis to see which version of the campaign is resulting in the most leads, or is the most successful. This same strategy can be applied to the pages or elements of your site, and you will be able to see which changes are more or less successful on your site.

2 A/B Testing Techniques

Though the first step to A/B testing is deciding what pages or site elements to include in your A/B test, the actual A/B testing doesn't begin until you decide on the testing technique you want to use – consecutive testing or synchronous testing.

2.1 Consecutive Testing

Consecutive testing is the process of presenting all visitors to your site with one page and after a certain amount of time, you switch out that page with another page for the same amount of time. In other words, you might present the visitors with Version A for one week, and then you switch it out and present the visitors with Version B for one week.

Consecutive testing is very easy to set up and track. If you conduct a consecutive test, all of your SiteCatalyst reports become A/B reports based on only the calendar range of your reports. For example, let's say you place Version A of a landing page on your site for one week and then switch it out for Version B of a landing page for the next week. In this case, performing an A/B test would be as easy as switching your reporting calendar from the one week where you implemented Version A of the landing page to the week where you implemented Version B of the landing page. The biggest advantage to using consecutive testing is that it requires fewer resources because you do not have to set anything up and send different versions to different groups of people.

However, that is not to say that consecutive testing does not have its disadvantages. Though the test results will be valid, they will not be as accurate as the results received from synchronous testing. With consecutive testing, a number of external variables that cannot be controlled might affect your data. Some of these variables might include:

- Stock market
- Politics
- News
- Competitor behavior
- Weather

Of course, these external factors are only examples. In reality, there are any number of external factors that could affect consecutive A/B testing. With this type of testing, the difficulties lie in determining whether any conversion changes were a result of the changes you made or of an external variable. This particular issue varies for different Omniture customers because some web sites are more sensitive to outside influences.

For example, let's say that in week 1 you use landing page A and in week 2 you use landing page B, but week 2 shows a lower number of conversions. In addition, week 2 was also a week in which your competitors launched a major sales campaign. In this case, there's really no way to properly analyze if the lower conversions were a result of the change to your site or a result of the sales campaign.

A second disadvantage to using consecutive testing is that you are showing a page to all visitors to your site. For example, if you using landing page A in week 1, and you are using landing page B in week 2, 100% of your visitors for either week are seeing the page you are showing for the week. As such, you are risking showing 100% of your traffic the new version of the page. If the change is not a positive one, you just gave the bad version of the page to all your traffic.

If the disadvantages discussed here are not enough to dissuade you from performing consecutive testing, be sure to watch your results very closely. If a change is showing negative effects, you can undo the change before much damage is caused. In addition, be sure to understand the time frames for all versions. In other words, if you run page A for a month, be sure to run page(s) B, C, D, etc. for a month as well. That way, you are comparing apples to apples. Finally, make sure that your test period is long enough for a valid sample.

2.2 Synchronous Testing

Synchronous testing is different from consecutive testing in that you test multiple versions of a page or web elements at the same time. You basically divide your traffic by sending one group to a page, while you send the other group to a different page.

For example, you want to see which page performs better – Landing Page A or Landing Page B. You make both pages available during the same week, but you send 50% of your traffic to Landing Page A, and you send the other 50% to Landing Page B. The page that has a higher conversion rate obviously is the page that performed better.

Synchronous tests are more accurate than consecutive tests because the page or element variations are being tested at the same time. Any outside variables have no influence on the outcome of the test. In addition, with synchronous testing, you are actually able to select the percent of your visitors that you want to send to the different versions of your pages. You could send 20% of your traffic to the new page, while 80% of the traffic will still be sent to your original page. Most frequently, traffic is split 50/50 between the two pages, but you don't necessarily have to follow that method.

The disadvantage to performing synchronous tests over consecutive tests is that synchronous tests can be much more difficult to perform. First, you have to place the page code on the different pages being tracked. Second, you have to have the appropriate sample groups set up, which means determining if a certain visitor gets the same page every time they visit, or if they get a random version of the page any time they come to your site.. Your development team will have to figure out a way to split your visitors into various testing groups. Third, your organization will have to have the time and resources available to split out your site traffic into two groups.

3 Validating Sample Groups

For consecutive testing, you don't need to validate your sample group because no sample group exists. In other words, consecutive testing has your entire group of visitors viewing one version of a page or element, and then they view a different page or element.

Synchronous testing, on the other hand, requires you to validate your sample groups, especially if you are not splitting the group viewing equally (50/50). To create valid sample groups, Omniture recommends that you follow the best practices shown below.

- Sample users randomly from an existing/constant data set, preferably from across all traffic sources
- Approximately 500 instances of conversion are usually sufficient to identify overall conversion probability
- Define a time period that will give you a holistic representation of your site traffic. This practice "trumps" the bullet point above because if you need to run the test for a week to get a good representation of your site, you may have more than 500 conversions for that week.
- Minimize site changes during the comparison period to ensure accurate results. Don't change the whole site; instead, just focus on one element.

If you are using synchronous testing, Omniture also recommends that you use one of two testing models.

- Benchmark Test
- A/B/A Test

The Benchmark test model requires a two-step process. First, validate your sample group, and second, run the test. To validate your sample group, first segment your traffic into two groups (Group A and Group B) and show both groups the same version of the page (Page A). At this point, you will be able to revise your sample groups until both groups have identical conversion rates so you know that your sample group is valid.



NOTE: You can choose any percentage that you want to risk on the new version with consecutive testing, so you might want to start with a sample group as low as 1% and go from there, if needed.

Once you have matching conversion rates, and you know the sample is valid, you can then switch Group B so that they are interacting with Page B instead of Page A. By watching any changes in the conversion rate, you will be able to see the success rate of the page. Note that the test must be long enough to determine the correct sample size.

Though the Benchmark test requires you to validate your group and then complete the test, the A/B/A test allows you to validate your sample group at the same time that you are running your test. To do this, segment your traffic into three groups. The majority of your traffic will go to Group 1, which will interact with Version A of your page. The remaining traffic will be divided into two other groups (B and C). Group B will interact with Page B, and Group C will interact with Page A. Groups B and C are interchangeable. The main point is that the smaller groups are really identical samples of your traffic.

Next, start to adjust the percentages and sample sources. As you change the sample groups, make sure B and C are always the same. For example, you might give 98% of traffic to Page A, but Groups B and C must be equally split and have 1% each.

4 SiteCatalyst A/B Testing Toolbox

Understanding the theories of A/B testing are useful, but they are not useful if you can't implement them with SiteCatalyst. You can use any of the following Omniture tools to successfully test your web pages or elements.

- SiteCatalyst Reports
- ClickMap
- Commerce Variables
- Traffic Variables
- SiteCatalyst A/B Compare Tool
- ASI

If you are using the consecutive testing method, you can use the SiteCatalyst reports, ClickMap, or the SiteCatalyst A/B Compare tool. If you are using synchronous testing, you can use the Custom Commerce variables, the Custom Traffic variables, or ClickMap.

4.1 Consecutive Testing

If you are using the consecutive method for A/B testing, virtually any report in SiteCatalyst becomes an A/B testing tool for you. Of those reports, you need to determine which one provide your Key Performance Indicators (KPI). For example, if you have a lead generation site, you can look at the leads conversion percentage that was achieved during the time that you had the new version displayed vs. the old version. You could also look at the pathing for the two separate times to see if it was better or worse during those times.

If you want to use ClickMap, to help you with your testing, it can be a good indicator of whether the page performed better with the new version as compared with the old one. With ClickMap, not only can you check the number of clicks, but you can have conversion metrics enabled (via Live Support) so you can see how much conversion was driven by links on the new page.

Finally, when you run item-based reports (or reports that have a list of items and are measured by a metric) in SiteCatalyst, you will typically have the option of running the A/B compare tool from the toolbar in the report. This tool has two options for comparison type: Sites/Segments or Date Ranges. The Date Range option has obvious application here because it compares any two time periods, which is exactly what is needed.

4.2 Synchronous Testing

For synchronous testing, you can use eVars, which by nature, are perfect for A/B testing because they allow you to set a variable on a page and then wait to see what conversion occurs. If you are testing two versions of a page, you can simply set the eVar to a value that describes the page, in the regular code, when the visitors load the page. Then, as mentioned, you can wait to see if they convert, after which, you can run the SiteCatalyst report to show you the conversion. The solution described here is the most common type of A/B synchronous test. Consider the following example.

Similarly, you can use Custom Insight Variables, called s.props, in A/B testing for traffic and pathing-related data. If you are primarily interested in knowing how different versions affect traffic and pathing on your site, then you can do the same thing you did with an eVar, except to put it to a s.prop instead. In this case, Omniture advises that you mirror the page name into an s.prop, adding version distinctions as well. Next, have your supported user contact Omniture Live Support, and have them enable pathing for that s.prop, which will give you an entire set of Paths reports for your A/B testing. You will be able to see if visitors path differently based on the versions of the page you have given them to view.

Consider the following example.

Let's say you have two versions of your home page. You would like to see which one performs better. This is an instance in which you are comparing just two versions of a single page. In this case, assign a value to the same s.prop and eVar on both pages. For example, on Home Page A, give the value as shown below.

```
s.pageName="HomePage"  
s.prop1="A"  
s.prop2="A:HomePage"  
eVar1="A"
```

The value A stands for "Home Page A," but you can use any value that will help you remember s.prop1 and eVar1 when you see the value displayed in the reports. Similarly, on Home Page B, you would enter a similar value with s.prop1 and eVar1 as shown below. Make sure s.prop1 and eVar1 have the same value for each page.

```
pageName="HomePageB"  
s.prop1="B"  
s.prop2="B:HomePage"  
eVar1="B"
```

Similarly, if you have two completely different web sites, and you want to compare their performance, you would assign a value to s.prop1 and eVar1 (or another selected s.prop and eVar) of every page of the site.

Finally, ClickMap is a helpful A/B testing tool in a synchronous test as well as in a consecutive test. The beauty of using it in a synchronous test is that you don't have to archive ClickMap to run the test. You can display both versions of the page at the same time and run the test on either one.

For example, you want to see if changing the location of the internal search field on your Home Page affects your conversion rate. Let's say that version A has the search field on the right side of the home page. You can track this by setting a commerce variable (eVar2) equal to "HP Search Right." Similarly, for version B, you move the search box to the top, center of the home page to test the results. You can track this by setting a commerce variable (eVar2) equal to "HP Search Top." When you run the eVar2 report (renamed "Home Page Variations) in SiteCatalyst and create an order conversion rate calculated metric, you can determine which version of the page or element is more successful.

4.3 Segment Testing

The previous solutions discussed have referred to A and B versions of pages, meaning you would be making changes to your site to see which version of the pages help conversion or pathing, etc. However, one other type of A/B testing is noteworthy. This type of testing refers to two or more segments instead of pages in your web site.

For example, if you have a membership club on your site, and that club has gold, silver, and bronze members. Perhaps you would like to know if your gold members convert better than your silver members or your bronze members. This is usually as simple as running an eVar report, but if these segments are important enough that they have their own report suites, you can use the A/B Compare tool in most of the SiteCatalyst reports.

In the A/B Compare tool, simply select the "Site/Segment" Comparison Type instead of the Date Range Comparison Type. Then, you can compare the report from two different report suites at the same time, side by side, including comparison percentages.

In addition, ASI lends itself to this method very well because you can compare an ASI suite to the whole source report suite, or you can compare two ASI segments to see the differences in any reports, including Products, Pages, etc.



CALL 1.877.722.7088
1.801.722.0139

www.omniture.com
info@omniture.com

550 East Timpanogos Circle
Orem, Utah 84097

