Side by side: Comparing two performance marketing tools/agencies



As performance marketers, we're conditioned to want to test everything. From the impact of feed titles to the incrementality of each channel, we want to be sure that we're making the right choice before we commit all our resources to something.

That goes for deciding which tools/agency to use as well. Moving your performance marketing activities from one tool/agency to another (or picking one to start with) is a big commitment and not one you should take lightly.

Most (probably all) tools/ agencies claim to do the same basic thing: improve your campaign performance. The way they do this or the methods they use will differ, but with so much choice out there, how are you to know which one will actually deliver?

To help them make the right decision, many companies will ask for a side-by-side comparison test between two tools/agencies. We were recently asked to participate in a split comparison test against a Philadelphia-based product ad technology for a multinational sports retailer.

Using this latest test as an example, I want to provide you with a few tips and recommendations for conducting an accurate side-by-side test and getting the best results.

Have a clear test purpose

No matter what kind of test you're conducting, you need to make sure you have a clear goal in mind, i.e., a question that can be answered in a definitive and measurable way.

In this case, the retailer wanted to know if replacing their existing tools with our Google Shopping solution would impact their bottom line. To measure the outcome, they set a fixed ROAS target for each tool to meet and measured the resulting revenue figures.

Split test groups evenly

As always, a good test setup is critical if you want the results to be accurate. I've written before about different testing methods, but the simplest way to approach this type of test is to think of it like any other A/B test.

The key is to divide up your campaigns as evenly as possible so there are no extenuating circumstances that might explain one tool's/agency's success or failure. There's no silver bullet here. The best way to split out your campaigns will likely differ depending on your products, the season, shopper demographics and so on.

The important thing is to try to keep as many variables the same as possible. You especially want the current and recent historical figures to match up – e.g., roughly the same traffic, conversion and ROAS numbers.

The two most common ways to split up campaigns are by time or location. For a time slot split, you would duplicate the campaigns you want to use in the test, assign one to each tool/agency, and then activate them on a fixed order of rotation (usually hourly). The benefit of this method is that it does away with any seasonality or geographic concerns by ensuring that each tool/agency is working with exactly the

same product range over the same date range.

For a geo split, you take one of your regions, duplicate the campaigns there and divide into two sections with roughly the same impression and conversion volumes. Then, you assign each tool to one half of the territory and switch them every two weeks. The benefit of this sort of split is that it is technically quite easy to set up and allows you to adjust the scheduling during the test, as well as run promotions. It also allows each tool/agency to work simultaneously, which cuts down on the amount of time you have to run the test.



Technically, you can split your campaigns up any way you like. As I said before, there is no one way to split out your campaigns. It all depends on how your campaigns/territories/customers/products are distributed. You just want to make sure to get the two campaigns as close as possible.

Run your test for long enough

Testing – no matter what kind – costs money. As time goes on, you run the risk of wasting money on a strategy that isn't working or isn't working as well as another strategy. On the flip side, you need to run your test for long enough that you get a clear answer as to which test group performed the best.

A 50/50 testing split means dividing your total traffic in half. You want to make sure the variants receive statistically enough traffic within the time set aside for the test. How long this actually takes will depend on how much traffic you get and how obvious the difference is.

Keep in mind that any tool (or agency) that uses algorithms and machine learning to manage bids will need a certain amount of time to collect data before the automation will really live

up to its potential.



In this case, the retailer ran the test for 10 weeks. You can see that the differences between us and their existing tool were very slight in the first half of the test while the algorithms collected data. But by the second half of the test, we had collected enough data for the automated bidding and the new campaign structure to take effect, creating a much more pronounced difference between our two approaches.

Another important thing to remember is to wait another two weeks after the test finishes before collecting and evaluating the results so that you capture any latent sales.

Evaluating the results

Before the test, make sure you set clear, achievable goals. The goals you want your candidates to reach need to be achievable for the parameters (e.g., budget, time, season, product skew) you set alongside them. You can’t expect the moon just yet.

Now is also not the time to get fancy with your metrics. Set the same type of targets you’ve been using so that you can easily compare the new results to your baseline data. Once your test period is up, you can compare the results from each test/agency against your benchmark KPIs.

In this case, the most important metric for the retailer was who could drive the most revenue within the target ROAS. However, they also looked at other metrics like Cost and ROAS.



You’re not necessarily looking for the hard numbers; what’s more important is the difference between the performance of the two tools/agencies and how that compares to your baseline.

In this case, our competitor was already the established tool, so there was no need to plot the baseline.

If you're evaluating agencies, or if the tool you're looking at has a customer support arm, ask them to give you their take on how the test went. Not only will this provide you with some color commentary for the raw data, but it also will give you an idea of what the working relationship with that company will be like going forward. For example:

- Do they make a lot of excuses?
- Do they have good reporting techniques?
- What is their plan for the future of your accounts?
- What adjustments would they recommend you make to your accounts?

You want to work with a company that will push innovative ideas and new strategies, not one that will sit back and perform the same repetitive tasks over and over.

Recommendations

There are lots of things you can do during testing that may skew the outcome. In addition to setting up a good test, here are a few recommendations for not biasing the results.

First of all, it's generally not a good idea to change much of anything during the testing period, but you'd be surprised how often this happens.

Budgets, and how each agency or tool allocates them, are a key part of what you want to evaluate during the test period. Similar to the budget scenario, changing the attribution model or the target KPI will hugely affect the results of the test thus far.

Check in on progress. You don't want to change things unnecessarily, but you should still be checking in with the results as they're available to make sure your test structure

is working. If something looks radically off in the first couple weeks, it probably is. In that case, you may need to rethink your testing structure and make some changes. Any changes you make will likely extend the end date for the test as well.

Hopefully, this example will provide a guide for your tool/agency evaluations. 2018 is going to be a big year for retail, and it's essential that you cultivate the right tools for the job.

If you're seeking every advantage that will propel your retail business to the top, you can't miss our rich retail-focused content at SMX West. Join us March 13-15, 2018 in San Jose, CA for three full days of SEM and PPC sessions specially designed for e-commerce pros like you. Major players, including Google and Bing, will be there to share firsthand advice that can help you effectively reach your customers and increase overall sales. Together, we will dive deep into the areas of online retail that matter most to you and your bottom line.

Register now for the best rates – early bird pricing expires February 3!

Opinions expressed in this article are those of the guest author and not necessarily Search Engine Land. Staff authors are listed here.

About The Author



Andreas Reiffen is an entrepreneur, marketing technologist and thought leader in data-driven advertising. Between 2006 and 2008 Andreas operated as a PPC super affiliate working for

Zappos and other retailers. In 2008 he founded crealytics, the Berlin-based provider of camato, the leading Google Shopping Tool to automate PLA campaigns and improve performance. Several of the largest international Ecommerce retailers are using camato to manage Google Shopping in more than 100 markets across the globe, generating over $3B in revenues per year. camato has been shortlisted for numerous awards in 2016 including European Performance Marketing and Drum Search.
http://feeds.searchengineland.com/~r/searchengineland/~3/-zoKvPl7jBA/side-side-comparing-two-performance-marketing-toolsagencies-290221