



SPLIT TESTING SUCKS AND OTHER HERESIES

Split testing is what marketing is all about; right? Wrong! There is a much better way to get the answer to 95% of the kinds of questions you might consider split testing. Let's step through it.

First of all, you must decide your desired outcome. Is it profitability? Is it search engine ranking? Is it more traffic? Is it more inbound links?

Once you have decided your desired outcome, you must find a way that OTHER PEOPLE'S sites can be measured for that outcome.

That's right; I'm not advocating split testing because there is a much faster and even more accurate way to get the same results by simply looking at the results of OTHER PEOPLE!

It's called statistical analysis. It's what most scientific advancements have been based on! When you go to the doctor, does he split test antibiotics and exercise on you? Of course not! You would fire him immediately. Instead, he looks at the results of studies where other people had the same desired outcome you want (getting rid of that fever, cough, whatever) and prescribes the drug or treatment that was proven safe and effective for a statistically significant number of OTHER PEOPLE!

Will it always work for you because it worked on a statistically significant number of OTHER PEOPLE? No; of course not. But it will work a majority of the time. The fact is that split testing has the same problems. When you find an answer from split testing and choose the "a" version over the "b" version based on 20 actions... There is a percentage chance that you actually chose the wrong version. Increase the number of actions and you increase your chances of picking the right one. The same is true when looking at OTHER PEOPLE'S results instead of your own.

OK; so how can we measure some of these things. Search Engine Ranking is easy. You can compare sites at the #1 ranking with sites at the #100 ranking... or sites in the top 10 with sites in the 101-110 range in ranking. No problem.

More traffic? You can find open logs for a statistically significant number of sites. A less perfect measurement might be Alexa or Jupiter ratings. When dealing with less accurate measurements like that, simply increase your sample size or compare more distant extremes (Alexa rankings in the top 100 vs Alexa ranking in the 99,900-100,000 traffic rankings).

More inbound links? Once again you can use open logs and look at referral entries... or you can trust the link: command at MSN... or something less reliable like the link: command at Google (once again, just increase your sample size and/or compare more distant extremes in number of inbound links).

How about profitability? This is my favorite. This is why most of us are here. I used this measurement for my Glyphius software and all of the copywriting statistical studies posted on my blog and in the Statistical Copywriting Online Home Study Course.

How can you get the profitability figures for other web-sites? Are they just going to turn them over? Actually, public companies do just that. You could use something like that public data. Also, many affiliate networks give indirect profitability figures (the marketplace at Clickbank, the EPC rating on CJ are both less accurate profitability numbers).

Or you could do what marketers have quietly done for decades before ecommerce and the Internet even existed. This is how Glyphius and the Statistical Copywriting course were created. It is really quite simple. Ask yourself these two questions:

1. If I was paying for advertising to a site that was profitable... would I continue to pay for that advertising month after month as long as it was profitable?
2. If I was paying for advertising to a site that wasn't profitable... would I continue to pay for that advertising month after month?

The answer to #1 is clearly a "yes" for a vast majority of people. The answer to #2 is clearly a "no" for a vast majority of people. There are exceptions. Some large companies don't even track the profitability of their ads... so they will add some incorrect data if we use this measurement.

The same thing happens in science all the time. Patients in medical studies lie or exaggerate about their results (in both directions). People aren't completely predictable. They do things against their own self-interest sometimes.

The solution? Same as with the other imperfect measurement techniques... increase the sample size and/or compare more distant extremes. If people were 100% predictable, then we wouldn't need a very large sample size to look at these web-sites.

So, let's take an example to see how this works. Let's say some idiot "guru" marketer is telling you that using the digit "7" in your pricing will increase your conversion ratio and your profitability. OK; fine. You have three choices:

1. Believe him/her. Congratulations; you have just joined the religious society of idiot "guru" marketers. Instead of embracing science for those things that can be easily researched, you have decided to apply a religious belief to such a topic instead. You can feel a little better about yourself because 98% of all people are in the same boat. Unfortunately 98% of all people will retire at 65 years old broke.
2. Split test it. Arrgh! You know why you haven't already done this... it takes forever to get statistically significant results. Then you know that there are still hundreds or thousands of things you still need to split test after that.
3. Statistically analyze OTHER PEOPLE'S pricing and their profitability.

Look; some things need to be split tested. Still; wouldn't it be nice if you could come very close to the real result before you even start your split tests? Do you really need to split test a bright purple background on your site when it's obvious that a vast majority of profitable sites use a white background? Of course not! Test only those things that can't be easily found by doing a statistical analysis of OTHER PEOPLE'S sites!

OK; so how do we get the above answer?

First we make us a list of profitable and unprofitable sites. We already know how to do that; right? No? OK; let's break it down.

First we search for a statistically significant number of keywords on Google or any other search engine that shows paid advertising (or we go get some magazines that advertise web-sites... or we go get a nickle ad sheet and find some classifieds that include URLs... it doesn't matter... just go to a source of paid advertising that shows URLs).

Then we make a list showing the exact ad and the URL. Now we wait. How long do we wait? We wait long enough to have some folks with unprofitable ads change them or remove them. If you used magazines, wait 6 months. Some magazines have a policy of only allowing you to run an ad for three months at a time. So, it takes more than 3 months before you see unprofitable ads disappear. On Google, you can probably get away with waiting just a couple of weeks. The longer you wait, the fewer sites you will have to compare (the old comparing more distant extremes rule).

Now you do it again. Go make a list of ads and their URLs that are still paying for advertising.

Now compare your two lists. Wherever you find an ad that is exactly the same months later, you have an ad that is very likely to be profitable. Put that URL on your "profitable" list. Put the other URLs that were on your first list, but are missing on your second list on your "unprofitable" list.

Cool; now you have a list of profitable and unprofitable sites. You can now run all kinds of quick statistical analysis on these sites. Since you can review a hundred sites much, much faster than you can usually get 100 results in a split test... you are way ahead now.

Now, simply go through your "profitable" list and figure out the percentage of prices that have a "7" in them. Now go through your unprofitable list and figure out the percentage of prices that have a "7" in them. Compare the two numbers. Are they very close to each other? Then using a "7" probably doesn't matter

very much in your pricing when it comes to profitability. Are they very distant? Did you find that 83% of unprofitable sites used a / 7 in their price, but only 21% of profitable sites did? Woo hoo! You have found your answer in just an hour instead of waiting for days for a split test.

So how do you know if you had a statistically significant number of sites to test? Yikes; with something this complex, that would take a couple of years of statistical training to figure that out. Let me give you a real world way of finding out that will save you from the hell of majoring in statistics in college (although if you are a single male, I highly recommend doing the major... for some reason statistics chicks are much better looking than chicks taking other majors... statistically speaking).

Predictability. Turn around your study and do a split test. What percentage of the time did the results of the split test match the predicted outcome of the statistical analysis. The closer you get to 100%, the more you can rely on the dataset of profitable and unprofitable sites you built. Of course, 0% isn't the bottom of the scale here... 50% is. So if you are only a bit over 50%, then you really need to look at the assumptions of your study... and the sample size... but if you are well over 75%, then you have a decent dataset that you can rely on for future studies.

You might be thinking... wait! I still had to do a split test! And worse... building that list of profitable and unprofitable sites took me just as long as it would take to do a split test... so what gives?

You only have to build the list of profitable and unprofitable sites and validate it once! Then you can do an unlimited number of studies using the same data!

The same is true for any measurement. My ranking data asks several thousand questions using just one set of data of high ranking and low ranking sites (with a 91% confidence rating)! I've answered dozens of copywriting questions using my one dataset of profitable and unprofitable sites (with a 96% confidence rating in that case).

Yes; it's a lot of work just to get the answer to the "7 question... but then it is very little work to then ask "what about the digit '3"? Or "Are red headlines more profitable or blue headlines?" or "Is long copy really more profitable or not?"

Instead of wasting your time doing a split test for days for each answer (or if you are a newbie still struggling to get traffic... maybe months to get an answer) and during that time not making as much money as you could be making... you can just spend an hour tabulating the results from your profitable and unprofitable lists of sites! Then you can have an answer that you can count on 96% of the time.

Guess what? With 20 actions in a split test (20 sales in this case for both "a" and "b"), you only have a 95% confidence in your results. How long will it take you to get at least 20 sales for both your "a" and your "b" version in a split test?

So there you go. I'm a heretic. I hate split testing my own sites for answers I can easily get by analyzing OTHER PEOPLE'S sites. It isn't the first time I disagree with 98% of marketing "gurus" and it probably won't be the last.

<https://blog.granted.com/>