marketingsherpa

Quick Guide to Conversion Rate Optimization

9 tactics to ensure you are running valid tests

CONTENTS

Quick Guide to Conversion Rate Optimization

9 tactics to ensure you are running valid tests

Author
Bobbi Dempsey, Editor, *Quick Guide to Conversion Rate Optimization*

Contributors
Flint McGlaughlin, Managing Director
Daniel Burstein, Director of Editorial Content
Laura Harkness, Research Analyst
Paul Cheney, Editorial Analyst
Diana Sindicich, Senior Research Manager

Editor
David Kirkpatrick, Manager of Editorial Content

Production Editor
Erin Hogg, Copy Editor

Copyright © 2014-2017 by MarketingSherpa LLC

All rights reserved. No part of this report may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, faxing, emailing, posting online or by any information storage and retrieval system, without written permission from the publisher.

TACTICS YOU CAN LEARN OVER LUNCH

Welcome to *Quick Guide to Conversion Rate Optimization* – a resource from MarketingSherpa featuring bite-sized tips for busy marketers.

In this report, we provide tips for ensuring you are running valid conversion rate optimization (CRO) tests.
It may shock you to know that MarketingSherpa research has shown that nearly half of marketers don't even calculate statistical significance when they test. This means they may be making major marketing and business decisions based on incomplete, or even worse, misleading data.

When it comes to doing accurate CRO testing, it's not as simple as just splitting traffic between two versions of a webpage. You must be assured that the results you see represent what would happen when using your entire list and/or audience. This means you need to confirm the validity of your tests.

The good news is you don't have to be a rocket scientist to run good tests. But you do need to know some fundamentals so you can avoid a few common pitfalls that threaten or destroy the validity of CRO tests.

In this report, we'll show you:

- Why valid testing is so important
- 3 common validity threats to be aware of
- Why detective work is sometimes needed
- How personalization can make testing a challenge

Note: This report only provides a very brief glimpse into validity threats related to CRO tests. For a much more comprehensive look at online testing, check out the MECLABS Online Testing Course.

We know you're in a hurry, so let's begin. We're eager to share these tips on how you can make sure your CRO tests are free from validity threats.

Bobbi Dempsey
Editor, *Quick Guide to Conversion Rate Optimization*

About *Quick Guide to Conversion Rate Optimization*

MarketingSherpa's *Quick Guide to Conversion Rate Optimization* is designed with you, the busy marketer, in mind. We provide quick, simple tips you can use right away. For each Quick Guide, we scour the vast MECLABS library of marketing research, from MarketingSherpa case studies and Benchmark Reports to MarketingExperiments optimization tests and analyses. We highlight tips to help improve your marketing performance right now – or at least by the time you're done with lunch.

© Copyright 2014-2017 MECLABS

For permissions: research@meclabs.com

Tactic #1. Understand the impact (and cost) of bad data

Marketers who don't give much (or any) thought to the validity of their testing are likely failing to realize how much this oversight can cost them – literally.

Testing and optimization guides your strategy and greatly influences your decisions. If there is a problem with the validity of your tests, it will have a ripple effect on all of your other efforts, and may send you off into a completely wrong direction causing you to waste valuable time and resources.

In the MarketingExperiments Web clinic, "Bad Data: The 3 validity threats that make your tests look conclusive (when they are deeply flawed)," Flint McGlaughlin, Managing Director, MECLABS, estimates that 75% to 80% of all tests being run on the Internet right now are flawed because they have undetected validity threats.

"There is nothing worse than making a confident decision with poor data," McGlaughlin said. "It would almost be better if you didn't test at all, because you'd be more careful without the false sense of confidence that poor data can give you."

In the MarketingExperiments Blog post, "Artificial Optimization: Why at least 40% of marketers shouldn't test," Paul Cheney, Editorial Analyst, MECLABS, echoed that thought, saying marketers who aren't testing may actually be better off than those who are – if those who are testing are doing it wrong.

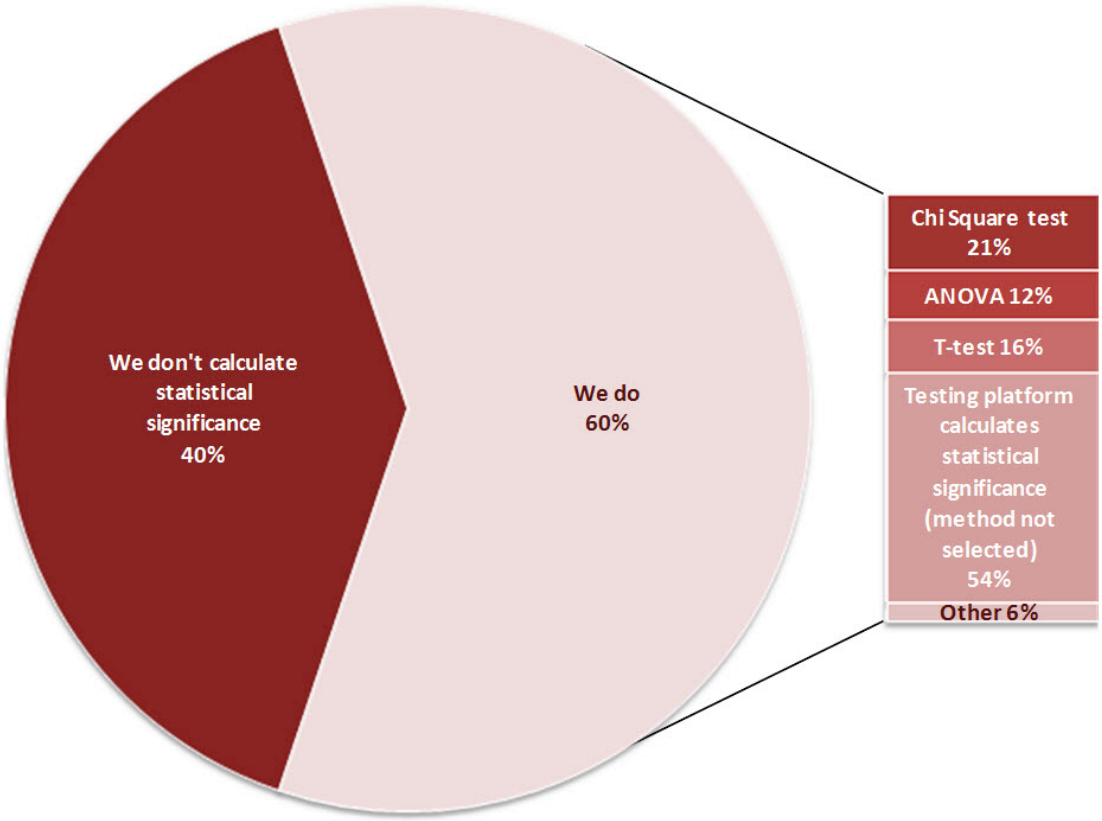
He stated, "If you're not careful, you could be running tests that tell you one thing when, in fact, the situation is completely different. You could be making critical decisions based on bad data. And these are the worst decisions you could make, because you've got the data to confirm that you're right, when you're actually doing things incorrectly."

Unfortunately, as the chart on the next page from the *MarketingSherpa Landing Page Optimization Benchmark Report* shows, at least 40% of marketers who do test fail to consider statistical significance. In other words, they have no way of knowing if the results they see are reliable or relevant. Meaning, they are not only likely wasting their time, but they then may be making bad decisions based on this faulty data.

© Copyright 2014-2017 MECLABS

For permissions: research@meclabs.com

Chart: Marketers' awareness and use of various methodologies to validate test results



marketingsherpa

Source: ©2011 MarketingSherpa
Landing Page Optimization Benchmark Survey
Methodology: Fielded February 2011, N=2,673

Related Resources

Website Analytics: How to use data to determine where to test

Marketing Analytics: 4 tips for productive conversations with your data analyst

A/B Testing: Ecommerce site's 3,000 positive comments show why you can't trust just one test

Tactic #2. Establish a controlled environment

If you don't properly prepare for the test, you may end up wasting your efforts as the test may not function as you intended, or may provide results that cannot be relied upon. Spending some time to correctly set up the test before it runs will greatly increase your chances of a successful test that provides useful data.

In the MarketingExperiments Blog post, "Marketing Experiment: Learn from Our Split Testing Mistakes," Daniel Burstein, Director of Editorial Content, MECLABS, shared some mistakes that MECLABS (intentionally) made when organizing a public A/B test for teaching purposes.

© Copyright 2014-2017 MECLABS

For permissions: research@meclabs.com

In this particular case, the test organized with Copyblogger was to choose the best subject line among entries submitted by readers. The structure of the test – a publicly announced contest – introduced some validity threats you would usually want to avoid under normal circumstances.

For one thing, the test was run in an uncontrolled environment, where a number of variables may have influenced the results – variables such as the timing of the voting period or the quantity of available choices.

Burstein noted, "An A/B test should be conducted in a controlled environment to ensure only the variables that affect the key performance indicators (KPI) are the ones impacted by the experiment."

In other words, if you just run a spontaneous test without establishing any sort of controlled environment, you have no way of knowing whether any number of random outside factors played a role in your test.

The goal of a test is to be able to isolate the one particular element you are testing as much as possible so that other uncontrolled variables don't influence the results.

To reach this goal, you need to establish a controlled environment for your test, eliminating the influence of external factors whenever possible, so that you can ensure the two options being tested can be compared equally without variables that might influence the results one way or the other unfairly.

In addition, as Burstein said, the power of an A/B test, over a focus group or opinion survey, for example, is that your potential customers do not know they are being tested. In this case, that wasn't true, since the contest was announced publicly, therefore the participants may have acted differently than they would have in a true CRO test situation where they were unaware they were being tested.

Related Resources

Validity Threats: 3 tips for online testing during a promotion (if you can't avoid it)

Online Testing Online Certification Course Level 1

Tactic #3. Use the correct sample size

One of the most important initial steps in creating a valid test is choosing a sample size that will be sufficient to provide reliable data.

In the MarketingEXperiments Blog post, "Marketing Optimization: How to determine the proper sample size," Burstein explained that it can be tricky (and nearly impossible) to identify an exact number as your "perfect" sample size. That's because you can't really know the scope of the difference among the results for the tested choices until after your test is complete. If the results are very close, having a larger sample size may have been helpful to get a more clear-cut winner. On the other hand, if there's a landslide winner, you probably could have seen useful results even with a smaller sample.

Often, choosing a sample size comes down to making an educated guess of how large your sample pool must be to provide the results that will satisfy your target level of confidence.

Generally, a larger sample size is considered better, so you will want to lean toward using as big of a sample group as possible.

© Copyright 2014-2017 MECLABS

For permissions: research@meclabs.com

If your sample size is too small, it will be difficult for you to know whether your results would apply to a larger population. With a small sample, the opinions of even just a few participants can greatly sway the results. Individual preferences and biases can play a much bigger role. This is a validity threat known as the sampling distortion effect.

Of course, assembling a large sample group can be a challenge for smaller companies or those whose lists tend to be limited in numbers. If you can't avoid using a relatively small sample, the MarketingExperiments Blog post, "A/B Testing: Working with a very small sample size is difficult, but not impossible," has some tips for getting the best results from your test.

Related Resources

Analytics & Testing: 3 statistical testing methods for building an advanced customer theory

How to tell if test data is statistically valid…

Tactic #4. Look for sudden changes – and investigate possible causes

In the MarketingExperiments Blog post, "Validity Threats: How we could have missed a 31% increase in conversions," Burstein illustrated an example of a case in which a testing oversight could have meant the team's efforts were totally wasted.

The case study involves a MECLABS Research Partner that wanted to find out which page design would yield the best conversion rate, so the team organized a multifactor split test. They created three new treatment versions to compare against the control page. The team was shocked to discover that the results were inconclusive – none of the treatments outperformed the control with any significant difference.

However, in further analyzing the data, they discovered the treatments had been outperforming the control for most of the testing period – up until things drastically shifted toward the end of the test. The team realized that an email had been sent out around that time, which skewed the sampling distribution by driving traffic to the control. After excluding data collected after that email was sent, the team was able to see that one treatment was a clear winner.

In the Web clinic from Tactic #1, McGlaughlin cited a similar example. In this test, which took place at an Optimization Summit, HubSpot teamed up with MarketingExperiments and the Summit audience to create a treatment email offer landing page for a free chapter of the *MarketingSherpa Landing Page Optimization Benchmark Report*. HubSpot's goal was to find a way to capture more names, so they wanted to see which landing page would help do that.

Three treatment options were created and then the audience voted for their choices of certain variables to test, including the headline, image, layout and the call-to-action.

The test was a much more compacted version of a normal test, since it only ran for 24 hours. Even so, it provided some interesting findings. The results started out trending in one direction and then suddenly made a complete reversal.

© Copyright 2014-2017 MECLABS

For permissions: research@meclabs.com

That's usually a big red flag, as it signals there is likely a validity threat with your test. In this case, the problem turned out to be that attendees at the event were sharing the URL of the treatment on social media, which was driving a lot of traffic to that version, thus skewing the results. You can learn more about the test and how they spotted the problem in the MarketingSherpa article, "Email Optimization and Testing: 800,000 emails sent, 1 major lesson learned."

Tactic #5. Be alert for the history effect

The history effect is when something happens in the outside world to affect the results of the test and causes flawed data. This could be a seasonal aspect, a major news event or a media trend.

In the Web clinic previously mentioned, McLaughlin cited a MarketingSherpa case study involving an online sex offender registry service for parents concerned about their surrounding area. The site wanted to increase the clickthrough rate of a PPC advertisement, so it decided to test the headline.

The team created four ads and ran a seven-day test that received 55,000 impressions. Three of the test headlines included the word "predator." By coincidence, at one point during the test period, the TV news show "Dateline" ran a special entitled "To Catch a Predator" that attracted a lot of viewers and received lots of attention. Not surprisingly, the headline choices with the word "predator" outperformed the other options.

In the contest example mentioned in Tactic #2, the test happened to be conducted the week before Christmas. The two top performing subject lines stressed urgency, which likely was a stronger incentive due to the timing, as many people feel a sense of deadline pressure right before the holidays.

It's important to realize that outside factors, such as what may be happening in the media, or on YouTube or elsewhere on the Internet, can have an effect on your test results.

Some tips for avoiding (or detecting) the history effect:

- Track day-to-day data so you can spot any sudden, unexpected changes.
- Use media tracking tools such as Google Alerts to track any specific terms prominent in your testing, so you can find out about any media events that may affect your testing.
- Make sure everyone in the company knows you are testing, so they can alert you to anything external that might be relevant. (If it's a large company, this may not be possible, so at least make sure the right people know.)

Tactic #6. Avoid the instrumentation effect

The instrumentation effect is when something happens with the testing tools that causes flawed data. This is perhaps the most common type of error causing poor testing data.

An example would be if something is causing your treatment page(s) to have significantly slower loading times than your control page. This would mean many people aren't even reading the treatment page, let alone clicking

© Copyright 2014-2017 MECLABS

For permissions: research@meclabs.com

on it. Also, depending on the metrics tool you're using, you may not even record those visitors who abandon the page before it fully loads.

From the users' perspective, there should be no noticeable difference in the process involved in accessing or interacting with your control page and the treatment(s). If one page is more difficult or inconvenient for the user, that will influence the results of your test. Slow load times, long URLs, server problems or site redirection and/or forwarding are examples of other issues that can cause the instrumentation effect.

In addition, any change in the way the metrics are measured or evaluated can also be a concern here. For example, if you use one type of analytics program for one page and a different one for another, that would affect the validity of your results. Likewise, if there were any changes or problems with your analytics program that occurred as the test was in progress, this is also something you'd want to track.

One way to help identify this problem is to use a good backup metrics tool and compare that data to the analytics you see on your primary metrics tool. But be sure to use the same metric tool, or tools, on all versions of the page you are testing so you'll be able to compare equivalent metrics. With both, you want to try and spot anything of concern. In this case, it would be anything that would indicate many visitors are abandoning the page before it loads completely.

Again, you should watch for noticeable abnormalities.

Tactic #7. Beware of the selection effect

The selection effect can occur when we wrongly assume a portion of traffic represents the total traffic.

An example of this is when you run a test using your best list (your house list). These people are likely more highly motivated than an average test group, which will affect your results. When you put the "winning" version of the test live on your site, you may see much different results from the general public, who will not have the same motivation level as your house list.

Nick Usborne, MECLABS contributor, gave another simplified example in the MarketingExperiments Blog post, "What is a 'Selection Effect' in online testing? And why does it matter?"

He said, "If you wanted to survey the New York City population about their support for tax reform and took your sample exclusively in the lobby of the Tiffany jewelry store, the results would be skewed due to selection effect."

Another more subtle example comes from a MarketingExperiments test involving a major news publisher. We had radically redesigned its subscription offer process for the electronic version and were in the middle of testing when the company launched a new text link ad campaign from its main website to the electronic product.

This changed the mix of traffic arriving at the subscription offer process from one where virtually all traffic was coming from paid search engines to one where much of the traffic was arriving from a link internal to its website (highly prequalified traffic).

© Copyright 2014-2017 MECLABS

For permissions: research@meclabs.com

The average conversion rate increased overnight from 0.26% to more than 2%. Had the team involved not been monitoring closely, they might have concluded that the new process had achieved a 600%+ conversion rate increase.

Note: Another list-related issue you need to consider is the list pollution effect. This means you cannot run another test or treatment on a group that has already participated in a previous version of the test. In other words, you cannot recycle the same list. You must use a fresh list for each new reiteration of the test.

Tactic #8. Consider the special challenge of sequential testing

Sequential tests can have a much larger risk of validity threat than a split test. Each step of the test process may be subject to its own variable conditions and its own specific threats.

In the MarketingExperiments Blog post, "Marketing Optimization: How to design split tests and multifactorial tests," Diana Sindicich, Senior Research Manager, MECLABS, said sequential tests are extremely prone to history effects (see Tactic #5).

She explained, "For example, an email sent out to the mailing list will increase traffic to whatever homepage treatment is currently online, distorting the actual effect of the design changes. This effect is usually noticeable as a sudden rise on an analytics traffic or conversion chart. Although it is not an optimal research design, this type of study can distinguish between a control and a treatment page. Results should only be interpreted if the possibility of history effect has been considered and found insignificant."

If you do use sequential testing, you should try and time the process so as to eliminate, or at least reduce, the history effect and any other validity threats that may be a concern.

Tactic #9. Eliminate problems caused by personalization

As Laura Harkness, Research Analyst, MECLABS, noted in the MarketingSherpa Blog post, "Web Optimization: 3 considerations for testing personalized webpage content," content personalization is perhaps one of the fastest-growing optimization tools, allowing marketers to segment visitors and deliver a more personalized message to optimize conversion.

While this is generally a good thing, of course, it can create a big challenge during the testing process. That's because in many cases a single user may be a match for several personas, causing some of the segments to overlap. The solution is to design the persona characteristics so they are mutually exclusive, meaning each user would only fall into a single category. The drawback, though, is that this will result in less traffic for each individual persona.

Note: For a handy spreadsheet that allows you to record and evaluate your data while also considering any validity threats, download the free MarketingExperiments Validity Tool.

© Copyright 2014-2017 MECLABS

For permissions: research@meclabs.com

Useful Links and Resources:

MarketingExperiments Blog – Optimization 101: How to get real results from A/B testing

MarketingSherpa Article – Website Optimization: Testing program leads to 638% increase in new accounts

MarketingExperiments Blog – Marketing Optimization: You can't find the true answer without the right question

MarketingSherpa Video Archive – A/B Testing: Optimizing calls-to-action for maximum conversions

MarketingExperiments Blog – Online marketing tests: How could you be so sure?

2015 Year in Testing: Lessons on value copy and friction

3 A/B Testing Case Studies from Smart Brand-Side Marketers

Page Templates Tested: How a few UX tweaks to 45 template pages generated a 52% increase in leads

Testing and Optimization: How to get that "ultimate lift"

© Copyright 2014-2017 MECLABS

For permissions: research@meclabs.com

About MarketingSherpa LLC

MarketingSherpa LLC is a research firm publishing case studies, benchmark data and how-to information read by hundreds of thousands of advertising, marketing and PR professionals every week.

Praised by *The Economist*, Harvard Business School's Working Knowledge Site and Entrepreneur.com, MarketingSherpa offers practical, results-based marketing information researched and written by a dedicated staff of in-house reporters.

MarketingSherpa's publications, available at www.marketingsherpa.com, include:

- More than 1,100 case studies, searchable by company or topic
- Annual Benchmark Reports featuring primary research and collected "best-of" secondary research, on statistics related to search marketing, email marketing, online advertising, e-commerce and business technology marketing

MarketingSherpa Newsletters

Visitors to MarketingSherpa.com may sign up for their choice of seven newsletters, including specific case studies for B2B and B2C marketers, email-focused studies and more.

MarketingSherpa's newsletters include:

- Best of the Week
- B2B Marketing
- Inbound Marketing
- Marketing Research Chart of the Week
- Email Marketing
- Consumer Marketing

Sign up for newsletters at www.marketingsherpa.com/newsletters.

MarketingSherpa Summits and Training Workshops

Register for Summits and Workshops at www.marketingsherpa.com, or contact:

MarketingSherpa (Customer Service available M-F, 9 a.m. to 5 p.m., ET)
service@marketingsherpa.com
1-877-895-1717 (outside the U.S., please call 401-383-3131)

© Copyright 2014-2017 MECLABS

For permissions: research@meclabs.com

MECLABS
INSTITUTE

4315 Pablo Oaks Court
Jacksonville, FL 32224

meclabs.com

MECLABS SERVICES

To support their ever-expanding research program, MECLABS has developed a series of science-based agency services, including:

- **Customer Research & Conversion**
- **Website Development & Design**
- **Product Creation & Launch**
- **Value Proposition & Brand Work**

MECLABS has conducted these science-based agency services in countries around the world spanning from Australia to Japan. They have worked with small startups and with many of the Fortune 50 leaders.

Now, with the world's largest independent research program in the field, their library includes:

20,000 **EXPERIMENTAL TREATMENTS**

2,500 **BRAND-SIDE CASE STUDIES**

500,000 **EXECUTIVE INTERVIEWS**

37,000 **BENCHMARKED COMPANIES**

Bank of America. Expedia Google H&R BLOCK Schlumberger McGraw Hill Education verizon

Learn how the **MECLABS Services Team** can help you get inside your customer's thinking & drive breakthrough results

CONTACT US TODAY:

EMAIL Services@MECLABS.com

PHONE **1.904.834.9803**