One of the most serious challenges in marketing is identifying the true impact of a given marketing spend change. In this guide, we show you how controlled marketing experiments can help improve campaign success.
Introduction

Consumers are making purchasing decisions around the clock—at work on their desktops, lounging on the couch with their tablets, and in stores on their smartphones. Today’s customer journey is complex and full of many touchpoints for marketers to connect with consumers across a variety of channels. As a result, campaign budgets—search, video, and more—need to reflect the most efficient marketing mix.

But in this constantly evolving consumer landscape, it’s difficult to know whether those marketing dollars are making an impact. Did the campaign perform as expected, better, or worse? Could the digital investment have been allocated in a more effective way? The pressure to prove value is all the more important because marketers know the budget for one campaign can often determine the investment for their next one.

The ripple effect is that marketers must function as scientists, conducting experiments when it comes to allocating budgets—whether by adapting the media mix, trying out different forms of creative, or exploring new marketing channels altogether. By measuring the full value of digital media investments through controlled marketing experiments, we can prove what is effective and what is not.

In this guide, we’ll help you understand the ideal scenarios for using controlled marketing experiments, why you should use them, and how. The best marketing experimentation follows a clear design process, is easy to interpret, and leads to continued learning and refinement.

And once you’ve aced the science bit, you can get back to the art of connecting with your customers.
Marketing experimentation has the potential to be quite powerful. The very act of testing can shine a spotlight on what works and what doesn’t. For example, if you knew your display ads were driving lots of additional clicks to your website, you might expand your display budget. Or, if you could prove through experimentation that mobile search ads increased in-store sales, you could justify increased investment there.

Of course, the process has a flip side, too. Poorly designed experiments can lead to disastrous misallocations of budgets, especially if too much weight is placed on a single experiment’s results. Say, for example, your experiment shows that generic search ads aren’t effective. As a result, you reduce that budget. Then you find out that the experiment results were wrong, and when you reduced your budget, you ended up with fewer prospects—which, ultimately, led to a decrease in overall conversions.

The key is to approach marketing experimentation with reasonable expectations: even an inconclusive experiment can offer valuable insights—as long as the test is designed and interpreted correctly.

controlled marketing experiments are most useful in determining the performance of a single, specific marketing channel such as video or mobile search. Let’s say, for example, you’re running a video ad campaign for the first time and you want to understand how well it’s performing. Specifically, you want to know if your YouTube campaign is having any impact on the business metrics you care about (such as brand awareness and perception, audience interest, or website visits and conversions). Looking at the initial data, you’re pleased to see that your video has gotten a ton of views, but you don’t know whether it actually drove more interest in your brand or website or was responsible for an increase in sales.

This is an ideal scenario for conducting a controlled marketing experiment. Why? You can ask a focused question based on your specific business goals (for example, “How effective is video at driving website visits?”). Since it’s your first video campaign, you lack benchmarks to help you answer that question—so a well-designed marketing experiment that shows your video to some users but not to others can help you answer this question with confidence.
Here are a few more important marketing questions that would best be answered by a controlled marketing experiment:

How do I …

- prove the impact of one media element in a campaign?
- understand the effectiveness of a risky new strategy before allocating a large budget to run the campaign in all markets?
- measure the impact of digital on in-store traffic and sales?
- distinguish correlational effects from true causal impact?
- confirm the effectiveness of a channel with poor performance before pulling the plug?
- affirm that my attribution metrics are effective so I can move away from last-click measurement?

**Incremental impact: The differentiating metric**

There are innumerable ways to measure a campaign—click-through rates, viewability, conversions, and more. The key difference between experimental results and other forms of measurement is the ability to identify incremental impact. This metric shows you the performance of the particular channel you’re testing in a more meaningful way. For example, rather than showing how many products your target audience bought, this metric shows how many *more* products they bought because of the *particular change in media spend* that you’re testing.
Incremental impact can be measured most reliably with a well-designed, controlled experiment. These experiments rely on groups of clearly defined test subjects (those exposed to the media spend change) and a comparable set of control subjects (those who are not). This test and control framework establishes a baseline so you can clearly identify the impact of your marketing efforts. Without this framework, it's easy for results to be ambiguous or inconclusive. Worse, a poorly designed marketing experiment may lead you to believe that the tested channel is effective when, in fact, the credit should go to one of your other marketing channels.

Keep in mind that many factors can affect the results of your marketing experiment. For example, don't jump to the conclusion that display won't ever work for your business because a particular display campaign performed terribly in your experiment. That specific campaign may have failed for other reasons (say, the messaging was off or the visuals didn't appeal to the target audience). One failure shouldn't stop you from experimenting with display again in the future.

Now that we've gone over when a controlled marketing experiment may be useful to your business, let's examine how you can evaluate your digital investments effectively. Here are the key steps to take:

1. **Identify your business goals and performance metrics**

Typically, marketers decide to run an experiment after they've already made plans to change their media spend. But you can introduce experimentation into the mix much earlier in the process. Ideally, experimentation would be a natural part of budget planning. This way, new media opportunities could be determined, in part, by their testability.

Even if you have an existing media plan, it's still important to set clear objectives for what you're testing. Ask yourself what you hope to achieve with this media spend. This means identifying the business goals and then laying down a solid measurement foundation with the right key performance indicators (KPIs) to justify the change. Goals for direct response campaigns may be easier to define (number of purchases, for example), but brand campaigns can also be tested by using brand metrics (awareness, perception, and audience interest).
2. Ask a focused question

Once you’ve thought through your business goals, you should begin your experimentation process by asking that focused, unambiguous question we talked about earlier. If you’re testing an upper-funnel activity, such as a digital branding campaign designed to generate awareness of your product, it may be difficult to accurately tie that campaign to in-store sales. Instead, you could focus on something that’s more realistic to measure, such as whether mobile searches lead to in-store purchases. And if your true goal is to prove the impact of digital media on offline sales, make sure your media buy reflects that goal. For example, you could test location-based mobile ads that include promotions for local stores.

3. Develop a marketing action plan

Build a concrete media plan—one that you expect will drive the desired results. It should be in response to that focused question you’ve identified. Clearly define the types of media you plan on using and testing to achieve your objectives.

4. Design the experiment

Before drawing conclusions, you’ll need to ensure that your experiment includes all the right details (and nothing extraneous) so that you can confidently make decisions based on your results. Your parameters—overall budget, the scope of the media change, and KPIs—can affect experimental outcomes. (Check out the “Nuts and Bolts” table for more details on good experimental design.)

5. Try, test, and test again

It can take more than one try to evaluate the efficiency of a given channel. Don’t stop with a single test. Take what you’ve learned from your first experiment, make some adjustments, and then test again.
HomeAway is a marketplace for homeowners and property managers to list their vacation homes for rent. When the company wanted to increase its number of rental properties, it used display advertising campaigns. But HomeAway needed to determine whether display was really helping to bring in incremental rental listings.

HomeAway first developed a media plan to increase the number of homeowner registrations, using ads on the Google Display Network. It used a carefully designed marketing experiment—with a test group who saw the display ads and a control group who didn’t—to find out if display was driving growth.

Based on the experiment, the team concluded that display was much more valuable than the company’s previous measurement methods (last-click valuation) had indicated: The cost per incremental acquisition (CPiA) for display was 51% lower than previously thought.

“Having a robust test backing up the solid performance of display media, beyond last-click opportunities, is key for HomeAway to grow its business,” said Brittany Heisler, digital analyst at HomeAway.

Following the experiment, HomeAway applied the newly measured CPiAs to develop its future budgeting and bidding strategy. Read the full case study.
Good Experimental Design: The Nuts and Bolts

<table>
<thead>
<tr>
<th>Set Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>A good experiment can be hard to define. That’s because there are many possible scenarios to choose from: How large should my test fraction be? How long do I want to run the test? How much should I spend to achieve a good result?</td>
</tr>
<tr>
<td>These parameters can influence the quality of a test and subsequent analysis dramatically. To get a reliable result, utilize historical data, spend behavior, and other factors unique to the business or marketing channel.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Confidence Intervals and Measurement Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental measurement is based on statistics, and there is always some level of uncertainty in any result. The questions become: How uncertain is the outcome? Did the observed effects happen by chance or because of the particular media tested? How reliable are the results?</td>
</tr>
<tr>
<td>The confidence interval addresses these questions. Simply put, a confidence interval tells you how likely it is that the measured value falls within certain boundaries. A well-designed experiment will aim for a result with a 90% to 95% confidence interval. Say the experimental result shows that your return on ad spend (ROAS) is “3 plus or minus 1.” The confidence interval provides guidance about how the result should be interpreted and applied: You know there is a 95% probability that ROAS is between 2 and 4.</td>
</tr>
<tr>
<td>You should be able to estimate the expected confidence interval before you begin the test. By identifying the confidence interval in advance, you can determine whether the test is likely to generate conclusive insights.</td>
</tr>
</tbody>
</table>
Control and Test Groups

You need to define a clear test group that’s exposed to the variable you’re testing and a comparable control group that isn’t exposed to your marketing efforts. This is a crucial step in setting up your controlled experiment.

This may mean identifying certain characteristics of your subjects such as demographics, buying behaviors, or geography. And make sure that your test and control groups are not systematically different from one another. In less rigorous approaches, those who are exposed to ads could be more or less likely to convert than those who are not exposed.

Random Selection

A well-designed controlled test employs randomized assignment of the test units (users, for example) to the test and control groups.

For example, this can be done by identifying website visitors individually (through a cookie-based test) or through the visitor’s geographic region (using a geo-based test). Picking a random population ensures that you don’t select a test group prone to showing more optimistic results.
DefShop is the biggest online shop for streetwear and hip hop clothing in Europe. The company had implemented a dynamic remarketing campaign on the Google Display Network, enabling the brand to reach previous site visitors with customized ads and bids. It wanted to assess the true value of this audience targeting: “It was about really understanding remarketing’s impact on our business, going from having only a feeling of what our ads might be driving to actually knowing their direct impact. And knowing is a much better feeling!” says Matthias Spangenberg, the company’s marketing manager.

To determine this, DefShop set up a randomized controlled experiment on product and cart abandoners. Users were assigned to test and control groups.

The experiment revealed that dynamic remarketing drove an incremental 12% increase in purchases, 23% more site visits, and 38% more brand search queries among users targeted with the remarketing campaign.
Panel Method

How it Works
Panel testing uses a preselected sample of volunteers who allow in-depth tracking of their behavior after exposure to an ad. Generally, only a subset of a given panel will be exposed as a test/control group.

Best Use Scenarios
Panels are best for longer-term studies that look at the impact of changing media spend on the target audience. Using existing channels, allowing participants to opt in to the study, results in a more organic experience—and better analysis.

When It Fails
Panels need to be very large to measure activities such as sales. Since only a small fraction of the panelists will convert, a very large panel is needed to generate definitive results. Selecting a diverse sample can be difficult, and unknown selection bias can affect the analysis. Also, it’s sometimes costly to provide incentives for individuals to participate while collecting the necessary data.
User-Based Method

How it Works

Test and control groups are based on known users or cookie IDs. For example, companies with a logged-in user base can separate users into segmented test and control groups. For companies without a readily identifiable user base, cookie IDs are applied instead. Read more about Google's unique user-based technology for measuring ad effectiveness.

Best Use Scenarios

When you need to test media impact on a specific customer segment, such as returning customers, cookie/user-based testing is ideal. Experiments with logged-in users can also detect cross-device effects.

When It Fails

Cookies may have a limited shelf life or may be deleted completely by the user. Cross-device measurement may also be unreliable if logged-in behavior changes across devices. Testing different media on logged-in users may generate privacy concerns, especially when your goal is to observe offline behavior (for example, in-store sales). To add to the complexity, if a user is exposed to multiple media channels, you run the risk of counting that user twice in an experiment. This also makes it harder to properly randomize the test and control groups.
How it Works

Comparable geographic regions are assigned to control and test groups. Everyone within a given region will be exposed to the same media. Geolocation for digital media is highly accurate.

Best Use Scenarios

These tests work for companies that may not have a logged-in user base. The ability to use aggregated data (by geo, by day) enables offline-tracked metrics such as in-store sales. In addition, they capture cross-device impact (since geolocation is typically device independent) and can determine the impact of many types of media in a campaign.

When It Fails

Geo-experiments won't work if a business is confined to a very small geographical region or country, or if marketing campaigns are very geographically specific. Even in larger experiments, the media being tested must be significant enough to generate detectable differences in user behavior such as offline purchase behavior. A geo-experiment may not be feasible, depending on the marketing scenario and the objectives of the advertiser.
Conclusion

Digital marketing is constantly evolving. It seems there’s a new media format, campaign strategy, or consumer behavior to adapt to almost every day. You’ll be better equipped for these changes if you include controlled marketing experiments in your toolkit.

Controlled marketing experiments provide concrete business value. Understanding the incremental impact of a given campaign is always useful. But beyond that, the very act of conducting a marketing experiment can be beneficial to your business. Testing encourages systematic thinking and forces marketing leaders in your organization to question what works, why it works, and what might be done differently. Simply going through the process of designing a marketing experiment can positively affect how you work together. And then you can all move on to the business at hand—implementing the strategies that allow you to connect with your customers.