

**What is  
Multivariate  
Testing?**

Multivariate testing includes the simultaneous comparison of combinations of variables within an e-mail, and can detect their combined effect.

# Testing for Impact:

## How Multivariate E-mail Testing Can Make a Lasting Impression

If you have no idea why e-mail testing works, you are not alone; the majority of e-mail marketers fail to embrace testing as a regular business practice. Jupiter Research reports that only 40 percent of e-mail marketers employ testing strategies on a regular basis. However, when they do, Jupiter finds they are almost twice as likely to attain conversion rates that exceed industry averages, compared with marketers that do not.

If the industry knows that testing is a highly-effective tool, why aren't more e-mail marketers adopting it as a key component of their programs? The answer is simple and complex at the same time. While e-mail testing can greatly improve campaign effectiveness and increase conversion rates, testing is time consuming and requires resource allocation to be done effectively. But it doesn't have to be painful. If you start simply and focus on adding complexity correctly, you can turn testing into, dare I say, fun.

### Where to Start

Begin your e-mail testing with a business goal to achieve. Without a business goal your test may examine a lot of "interesting things," but in the end you will likely wind up with many "interesting but not useful things." State your business goal as a behavior you would like to change, and how you would like it to change (more of, less of, fewer of, higher level of, etc.).

For example, you may want to learn how to drive more return visits from one-time buyers. You may want to know how to get a group of non-openers to open and ultimately buy.



You may want to get a group of prospects from stage two of your sales funnel to stage three. All of these goals examine a business issue and don't just focus on one element within one e-mail. If you start with a business goal, you can utilize your results for a longer period of time and make your investment worthwhile. At the same time, you should consider how you will measure the impact from your test: open rates, click rates, conversion rates, sale-cycle shifts, satisfaction ratings, and the like.

With your goals and metrics defined, the next step is determining the factors to test that will best support your objective. Factors, also called variables, are the things you are changing in an attempt to change behavior. Choosing your variables and implementing them correctly is where most mistakes are made. Just testing "things" is a waste of time: those tests tend to reveal something about one e-mail, but leave you with nothing about what to do in other mails. Instead, you should be able to describe what the user will see: those concepts, or general depictions, are your factors. Don't make your test about subject lines that read "Buy Now" vs. "Sale Until Wednesday." Test an immediate call-to-action vs. a mention of sale.

### **A/B vs. Multivariate**

In the e-mail marketing world there are two kinds of tests: A/B tests and multivariate tests. A/B tests examine one change at a time, while multivariate tests simultaneously compare combinations of variables within an e-mail, and can detect their combined effect. If you are currently not doing any e-mail testing, A/B are a great way to get your feet wet.

The correct way to A/B test is to alter only one element of an e-mail at a time to see how simple changes can influence the outcome of your campaign. For example, you could test different subject lines (or better, different variables around your subject lines), frequency of messages, or shorter vs. longer e-mails. As mentioned previously, structure your test by describing what your variables are. For instance, the variable could be amount of copy and the different options would include offering lengthy copy (defined as three paragraphs) vs. short copy (one paragraph). Changing more than one element at a time during an A/B test can be done, but should be avoided, since such tests will not reveal exactly what caused the difference in responses.



While simple A/B tests provide short-term answers if they are structured correctly, they take considerable time to see how combinations of variables work together. On the other hand, multivariate tests allow you to more quickly analyze combinations of changes to determine precisely what blend of offer, benefit, etc. will most effectively meet your goal.

### **Multivariate Testing—Getting Started**

The reason for e-mail testing is simple: to determine what works best with your customers and prospects. It will take time to lay out an effective multivariate test and track the data, but that will be a small price to pay in exchange for discovering what really works.

Before you start, you need to understand the costs associated with testing. Costs to consider include: the effort, time, and potential expense of a more complex production schedule, the content assets needed to test, creative (do you have in-house staff or do you have to outsource?), and the database you are mailing to (is it owned or do you have to pay each time you use it?).

The testing process becomes easier if you understand what success means to you. Does it mean more purchases; more spend per purchase; reactivation of non-active users; or more prospects into the sales funnel? Once your test is complete you will need to examine how the combinations of factors impacted these metrics. The measure of success is not whether the test variables had a positive effect, but whether you learned something. If none of your variables have any impact, you've learned some very important things. If all your manipulations depress response, you have learned what not to do again.

### **Who to Test**

Pick a random group, not just from the list, but from within each segment. If you do not pick random groups, but instead test by showing version one to one pre-selected group, and show version two to another pre-selected group, you have not accurately tested and have potentially wasted a mailing. Keep in mind that every random selection, on average, is truly equal to every other random selection, and this is the only way to understand the impact of your variables.



You may not want to give a specific level of a variable to every group. For example, you may feel that low-value customers should not receive the high-level 15 percent discount offer. In that case, it is okay to withhold that level from that group. A common solution is to describe the offer as “high” or “low,” and have the specifics vary by value-group (5 percent vs. 10 percent for the low-value group, 10 percent vs. 15 percent for the high-value group). Note that you are still selecting random samples to see each version, but they are randomly selected from within each segment. In certain cases it won’t be possible to have a truly random sample, such as in the case of an up-sell or cross-sell opportunity where the purchasers “self-select” themselves. In that instance, you can only test people who made recent purchases. However, within that group, do the best you can to randomly select which person within that group sees each combination of variables.

It is also important that your test has a control group that provides a baseline for comparison. Variables, such as seasonality, can have a huge effect on performance. In addition, the impact of your test variables may be hidden by the effect of offline media, news events or other factors. The only way to be sure is to have a control group exposed at the same time as your test groups. Your control group will not see the variations you are changing; they either see nothing at all, or the best performing past offer that you have.

### **What Can’t We Test?**

E-mail tests can be run on anything that is part of the campaign, including timing, subject line, length, graphics, etc. Typical e-mail tests examine things like promotions, benefit statements or aspects of subject lines. In addition, you can test variations of things outside the e-mail to understand how they interact with the overall e-mail experience. For example, you can test whether a product linked to a specific product page or to a category page results in a better response. Many designers will split a Web page up into component parts (a header, a sidebar, a primary content area, a secondary content area, and a footer). Each of these can be a place to manipulate creative, tone, or approach (hard or soft sell). In addition, one can test obvious vs. subtle relevance, i.e., “You left X in your shopping cart” vs. “You might enjoy these savings on category Y (which product X happens to be in).” You might also test the occurrence of mails, to better understand what triggers allow for frequency, when it’s not warranted, what groups respond best to what rate, and to tell how



### **Laying Out a Test Step 1:**

The first thing you should do when laying out the test is to decide which factors (concepts), or variables, you want to test. For example, a typical test would include things like offers, benefits and/or landing pages, each with various options or levels.

often is too often. Or, you might examine whether linking mails into sequences works better than treating each mail as its own unique unit.

Basically, anything that could impact the user experience can be tested, but be sure to consider the real impact and value of variables. It's fine to test anything of interest, but will it really change things? For example, finding the right day of the week to mail sounds like a silver bullet and is easy to test. However, its impact will be minimal and most research proves that, in most cases, there is no best day to mail. Similarly, time of the day to mail is usually not worth testing since research shows that people open mail when they want to, not when you send. So rather than waste time with variables which may have little impact or are already well understood, consider changes that will be worth the time and effort.

### **Layout of the Test**

The first thing you should do when laying out the test is to decide which factors (concepts), or variables, you want to test. For example, a typical test would include things like offers, benefits and/or landing pages, each with various options or levels. The promotion could have three levels: 5 percent, 10 percent and 15 percent off. Your product of interest could have three different benefits to feature, and you may have e-mails with links to a category-landing page and a product-specific page on your site. When we present every different combination of levels (e.g., Benefit 1 + 5 percent off + Category page, Benefit 2 + 5 percent off + Category page, Benefit 3 + 10 percent off + Product page, etc.), we would have 18 (3x3x3) different test combinations (or cells) to display.

For some mailers, 18 different cells may seem too complex to create, mail and track. Even the best companies will need to build in extra time to make sure that all the content fits well together and creates the desired effect. How many cells are too many? That depends on your capabilities, including asset creation, dynamic content mailing capability, tracking capability and analytic skills.

### **What to Measure**

For most direct marketers, lifetime value is the ultimate metric reflecting success of a business. However, depending on your needs, the metrics you choose may hinge on more



near-term tactical requirements. Looking at who opens your e-mail is great for testing subject line text, however you should also measure clicks (and conversions, if possible) from those openers to make sure that the right people are choosing to open based on your manipulations. Clicks are a popular choice for direct response, as are estimated change in lifetime value of the customer). In addition, a test may also look at “unsubs,” to see if some manipulations are reducing list churn. Beyond measuring just one mail, it’s important to examine customer behavior over time. For example, testing presence of more service-oriented content may result not in more clicks on any particular mail, but instead on longer term loyalty (more mails clicked at least once over time, less “unsubs” among that group).

The value of multivariate testing is not in the number or type of metrics you use, but in understanding the combined impact of the variables. Each variable on its own has a main effect. The combined effect is referred to as an interaction effect. Both of these should be examined. It’s all about interactions of variables and finding the right combination that costs the least amount and provides the most revenue. You can also weigh the importance of each metric based on your goal to determine which is the best combination. For example, if it is conversions you are after, give them 50 percent weight while other metrics get 25 percent.

### **Test Parameters**

How long should you run a test? Run the test for a long enough period of time to make sure each audience segment has seen the test, usually at least 48 hours, and preferably including a weekend day for consumer marketing. One of the biggest mistakes novice testers make is to test in the morning and mail based on those results in the afternoon. The biggest problem with this process is that your decision on what to mail everyone is based on the response of early responders, a unique group. What works for them may not work for people who read mail in the evening after the kids go to bed.

How large should your test be? The ballpark rule is to use 10,000 for a significant sample size. You may need more than 10,000 for unsubscribe and/or conversion tests, since these tend to have smaller effect sizes (that is, 30 percent of your mailing may open, but only .05 percent of net mailed may convert). You can get away with less (as small as 2,000), if you have a very responsive audience — or if you have a small list and have no choice.



Documenting your test also is extremely important. Make sure that with every test you record your goal, what you are testing, who you are including in the test, the methodology used and the measurements and expected findings. Use those findings for future mailings; hard work will pay off in better results. And remember to pass results along to the right division in the organization; your shareholders may thank you.

### **Results**

If you've laid out your e-mail test correctly, your results will tell you exactly what to do next. You will be able to recommend the action to take that will have the most effective impact on each group of interest. If you are looking for the "winner" of the test, however, you may be disappointed. There is no real winner, since different things work for different groups. Instead, you may have variables of impact which may be different for each group. It's also possible that your tests will not show a change in behavior. If so, think of that result as a great learning experience. It may show that you need to change your plan of attack. Perhaps your offer-strategy and content approach is working well, and you need to examine other factors in the e-mail.

Testing is the only way to really learn what works. Instead of playing the "what if" game, try to approach e-mail testing with a "test to see what happens" attitude. Try to throw in one unexpected category or variable for every five logical and rational selections. Embrace your unexpected results, because that's where the fun is. Happy testing!

### **About e-Dialog**

Established in 1997, e-Dialog is a proven provider of precision e-marketing solutions with deep-rooted strengths in e-mail and database marketing. Through a unique combination of marketing intelligence and precise relevance technology, e-Dialog enables some of the world's most recognized brands, such as American Eagle Outfitters, Avis, BMG Music Service, British Airways, the NFL, Reuters, The TJX Companies, and Tesco to maximize long-term customer value with contextually targeted communications. The company's service offerings empower large, multifaceted companies like these to enhance permission-based e-mail marketing efforts through fully integrated, cross-channel communications, including dynamically printed direct mail, RSS and mobile messaging.



JupiterResearch ranked e-Dialog the leading e-mail marketing provider among service-oriented ESPs in 2005 based on its value and market suitability, highlighting its account servicing, strategic and creative input, and campaign management and analytics applications. e-Dialog is a privately held company with offices in Boston, London, Seattle and New York. Investors include Flagship Ventures and Commonwealth Capital.

For more information, visit [www.e-Dialog.com](http://www.e-Dialog.com) or contact Arthur Sweetser at 781-372-3353 or Peter Duffy at +44 (0) 20 3219 6220.

---

BOSTON | 131 Hartwell Avenue | Lexington, MA 02421 | 888-256-7687 | [maxroi@e-dialog.com](mailto:maxroi@e-dialog.com)

NEW YORK | 21 Rhode Island Avenue | Massapequa, NY 11758 | 888-256-7687 | [maxroi@e-dialog.com](mailto:maxroi@e-dialog.com)

SEATTLE | 320-108<sup>th</sup> Ave NE Suite 300 | Bellevue, WA 98004 | 206-682-7777 | [maxroi@e-dialog.com](mailto:maxroi@e-dialog.com)

LONDON | 247 Tottenham Court Road | London W1T 7QX | 020 3219 6200 | [eurosales@e-dialog.com](mailto:eurosales@e-dialog.com)