



Exploring the Six Factors of Relevance: Testing and Measurement

By Jen Renna, *Account Director*

Okay, so you are applying to your e-mail programs one or more of the relevance factors we've discussed in the past five editions – segmentation, personalization, lifecycle management, triggers, and interactivity – and you think things are going well. But how do you really know whether this factor(s) is performing the way it was intended? You don't, unless you test it!

This may seem an obvious concept, but we're not simply talking about calculating how many people opened an e-mail or running the revenue numbers, but rather "looking under the hood" to determine how well your e-mail is resonating with recipients based on their behavior and responsiveness. Systematic testing, the sixth and final relevance factor, is critical to gathering more data that can help you to further craft campaigns and fine-tune messages that will drive even greater relevance.

Test who, test what?

With testing there are essentially two primary types: control group testing and variable testing. Control group testing is conducted by taking a small portion of your database, also called a holdout group (somewhere between 3% and 10%), and doing nothing new, just proceeding with business as usual. Then, by taking the rest your audience as a second test group and monitoring over time the effects of a new treatment, such as a welcome message or new navigation; the key here is to track effects for an extended period of time. This will allow you to measure the incremental impact of various programs, mailings, and messages and test, for example, how future behaviors are affected by past behavior. Will someone getting a relevant welcome message purchase more than someone who does not get a welcome message?

One particularly important thing to keep in mind with control group testing is the halo effect. The halo effect occurs when customers or prospects respond to a promotion outside of the testing process, but are not counted as having responded. This can deter some companies from using controls because they are afraid that they could be missing out on revenues by testing new ideas on only a portion of their list. But, if you roll a new treatment out to your entire list, how do you know whether it's having any impact and furthermore, how do you measure success?

Through variable and multivariate testing, you are measuring and testing individual elements such as audience, subject line, offer, content, frequency, and creative design, in order to again determine what's working and what's not. It's critical to keep this simple. Don't test too many (more than two or three) elements at the same time because results can quickly become confusing, difficult to "untangle," and audiences may become too small to represent statistically significant results.

The fact of the matter is that neither type of testing should be conducted in isolation; you should be doing both control group and variable testing. You can also choose to use control groups within variable testing to measure the specific lift resulting from the specific variable change. And with all testing, proper planning is as important as execution. Often, people don't think about how they are going to analyze test data, and that plays an important role in how to actually perform a test. For starters, make sure you have a broad enough test bed so that you can be sure the results are statistically significant and ensure that what you're testing is in line with the key performance indicators or other metrics you use to measure overall marketing campaign performance.

How relevant are your relevance factors?

Although testing is a relevance factor in and of itself, it's the litmus for how well all the others are performing. So, let's consider testing in the context of each factor, beginning with segmentation.

Segmentation

Primarily, what you want to know is whether your current segmentations are effective and what interests each segment most. One very simple evaluation is to test your subject lines across the various audience segments to determine what resonates more with each segment. You may find, for instance, that bargain shoppers respond better to a subject line "sale" message, while designer name brands in a subject line resonate with the fashion forward segment. (This brings up the fact that the e-mail channel provides an easy way to test price points. For example, you can deploy two messages with different price points to evaluate whether a 10% reduction in price will drive more purchases and greater overall ROI.)

Personalization

Without a doubt this factor is one of the easiest to test. Simply by using customer and/or prospect names in e-mail going out to one group and not including them in messages to your control group, you will be able to determine whether using a name drives an incremental lift. Another test would be to determine if there is a higher rate of response when loyalty club information is readily available in an e-mail. For example, are people more likely to book a vacation when travel points are published in an

e-mail? Additionally, when many of us think of personalization programs, we often think of preference centers; you can test to which single element or preference a customer better responds in order to drive greater relevance in future communications. For example, if you ask customers about when and where they are planning for the next vacation, and then split the audience and send an e-mail with “when” as the primary message, and another with “where” to those respective groups you can determine whether “when” or “where” is the principal motivator.

Lifecycle management

The objective of testing lifecycle management is to determine at each stage what message or offer not only best resonates with your audience, but also drives them to “convert” to the next stage. What’s nice about lifecycle testing is that by evaluating the responsiveness of a customer over time, you can clearly see if you are effectively moving him through each stage, and you can use this information to plan the content and timing of future campaigns.

Triggers

This factor is all about timing, so test your e-mails to determine the optimal time in which a recipient should receive a triggered message to elicit a behavior; for example, immediate versus 24 hours, 3 days, or 5 days. You should be testing one message versus a sequence to evaluate how many e-mails, how often will make the biggest impact on your audience and overall, determine how triggers affect baseline strategy.

A great example is abandoned shopping cart programs, which by their nature generate a lot of testing variables such as: how many days after abandonment you send a reminder; whether you present a discount or offer to drive recipients to purchase the item(s) in their carts; or, simply, do nothing. Some companies have found that there may be only a very small incremental difference in how much revenue they generate from a simple reminder versus a special discount. But again, without testing, companies can’t put numbers to those programs, and could lose dollars with those unnecessary discounts.

Interactivity

This big question here is whether interactivity is taking away from your call to action or enhancing the user experience. Test to determine if by incorporating things into your e-mail messages such as video links and user-generated content (i.e. survey and poll results, product review ratings, etc.) a customer is more likely to make a purchase. Even testing something as simple as creative layouts, such as short “postcard” style versus lengthy copy, can fine-tune your e-mail program tremendously.



Don't waste time, don't lose money – test everything

Unless you are systematically testing your e-mail programs, you could be wasting time and money either approaching your audiences with messages that aren't really relevant to them or doing things, such as offering special discounts or free shipping, which you don't have to. Simply consider the opportunity cost (the dollars you are losing because you could be doing something that would drive more revenue) of sending e-mails that don't resonate with your audience! Also think about the fact that you can create e-mail fatigue by sending useless e-mail that wastes the number of non-relevant "touches" a customer will accept.

There is one last, extremely important thing to remember all of this is moot if testing isn't completely transparent to the customer. A key premise behind driving e-mail relevance is to make the customer feel more like an individual; nothing makes someone feel less like an individual than unknowingly becoming a guinea pig. Take extra care to make sure a customer isn't on two different audience lists which you are using to test. Not only will they get two different offers, but you will create a lot of confusion and potentially turn them off.

In the end, only through testing can you be sure that the other factors of relevance are having a positive influence on your e-mail marketing initiatives. Testing is also the most effective vehicle for gathering information that can help you to design and/or adjust campaigns to improve future e-mail performance, and thus, your company's bottom line.