

ANALYTICS

How to Design (and Analyze) a Business Experiment

by Oliver Hauser and Michael Luca

OCTOBER 29, 2015



JIMMY MUSTO

The rise of experimental evaluations within organizations – or what economists refer to as field experiments – has the potential to transform organizational decision-making, providing fresh insight into areas ranging from product design to human resources to public policy. Companies that invest in randomized evaluations can gain a game-changing advantage.

Yet while there has been a rapid growth in experiments, especially within tech companies, we've seen too many run incorrectly. Even when they're set up properly, avoidable mistakes often happen during implementation. As a result, many organizations fail to receive the real benefits of the scientific method.

This article lays out seven steps to ensure that your experiment delivers. These principles draw on the academic research on field experiments as well as our work with a variety of organizations ranging from Yelp to the UK government.

1. Identify a narrow question. It is tempting to run an experiment around a question such as “Is advertising worth the cost?” or “Should we lower (or increase) our annual bonuses?” Indeed, beginning with a question that is central to your broader goals is a good start. But it's misguided to think that a single experiment will do the trick. The reason is simple: Multiple factors go into answering these types of big questions.

Take the issue of whether advertising is worth the cost. What form of advertising are we talking about, and for which products, in which media, over which time periods? Your question should be testable, which means it must be narrow and clearly defined. A better question might be, “How much does advertising our brand name on Google AdWords increase monthly sales?” This is an empirical question that an experiment can answer – and

that feeds into the question you ultimately hope to resolve. In fact, through just such an experiment, researchers at eBay discovered that a longstanding brand-advertising strategy on Google had no effect on the rate at which paying customers visited eBay.

2. Use a big hammer. Companies experiment when they don't know what will work best. Faced with this uncertainty, it may sound appealing to start small in order to avoid disrupting things. But your goal should be to see whether some version of your intervention – your new change – will make a difference to your customers. This requires a large enough intervention.

For example, suppose a grocery store is considering adding labels to items to show consumers that it sources mainly from local farms. How big should the labels be and where should they be attached? We would suggest starting with large labels on the front of the packages, because if the labels were small or on the backs of the packages, and there were no effect (a common outcome for subtle interventions), the store managers would be left to wonder whether consumers simply didn't notice the tags (the treatment wasn't large enough) or truly didn't care (there was no treatment effect). By starting with a big hammer, the store would learn whether customers care about local sourcing. If there's no effect from large labels on the package fronts, then the store should give up on the idea. If there *is* an effect, the experimenters can later refine the labels to the desired characteristics.

3. Perform a data audit. Once you know what your intervention is, you need to choose what data to look at. Make a list of all of the internal data related to the outcome you would like to influence and when you will need to do the measurements. Include data both about things you hope will change and things you hope *won't* change as a result of the intervention, because you'll need to be alert for unintended consequences. Think, too, about sources of external data that might add perspective.

Say you're launching a new cosmetics product, and you want to know which type of packaging leads to the highest customer loyalty and satisfaction. You decide to run a randomized controlled trial across geographical areas. In addition to measuring recurring orders and help-line customer feedback (internal data), you can track online user reviews on Amazon and look for differences among customers in different states (external data).

4. Choose a study population. Choose a subgroup among your customers that matches the customer profile you are looking to understand. It might be tempting to look for the easiest avenue to get a subgroup (such as online users), but beware: If your subgroup is not a good representation of your target customers, the findings of your experiment may not be applicable. For example, younger online customers who shop exclusively on your e-commerce platform may behave very differently from older in-store customers. You could use the former to generalize to your online platform strategy, but you may be misguided if you try to draw inferences from that group for your physical stores.

5. Randomize. Randomly assign some people to a treatment group and others to a control group. The treatment group receives the change you want to test, while the control group receives what you previously had on offer – and make sure there are no differences other than what you are testing. The first rule of randomization is to not let participants decide which group to be in, or the results will be meaningless. The second is to make sure there really are no differences between treatment and control.

It's not always easy to follow this second rule. For example, we've seen companies experiment by offering a different coupon on Sunday than on Monday. The problem is that Sunday shoppers may be systematically different from Monday shoppers (even if you control for the volume of shoppers on each day).

6. Commit to a plan, and stick to it. Before you run an experiment, lay out your plans in detail. How many observations will you collect? How long will you let the experiment run? What variables will be collected and analyzed? Record these details. This can be as simple as creating a Google spreadsheet or as official as using a public trial registry. Not only will this level of transparency make sure that everyone is on the same page; it will also help you avoid well-known pitfalls in the implementation of experiments.

Once your experiment is running, leave it alone! If you get a result you expected, great; if not, that's fine too. The one thing that's not OK: Running your experiment until your results look as though they fit your hypothesis, rather than until the study has run its planned course. This type of practice has led to a "replication crisis" in psychology research. It can seriously bias your results and reduce the insight you receive. The lesson? Stick to the plan, to the extent possible.

7. Let the data speak. To give a complete picture of your results, report multiple outcomes. Sure, some might be unchanged, unimpressive, or downright inexplicable. But better to be transparent about them than to ignore them. Once you've surveyed the main results, ask yourself whether you've really discovered the underlying mechanism behind your results – the factor that is driving them. If you're not sure, refine your experiment and run another trial to learn more.

Experiments are already a central part of the social sciences; they are quickly becoming central to organizations as well. If your experiments are well designed, they will tell you something valuable. The most successful will puncture your assumptions, change your practices, and put you ahead of competitors. Experimentation is a long-term, richly informative process, with each trial forming the starting point for the next.

Oliver Hauser is a doctoral student at Harvard University.



Michael Luca is an assistant professor at Harvard Business School.

This article is about ANALYTICS

 FOLLOW THIS TOPIC