

---

CHAPTER

9

# **The A/B Testing Dream Team**

## **Bringing Everyone Together**

---

**A**/B testing is by its nature collaborative and interdisciplinary, straddling traditional departmental lines: marketing, product management, engineering, and design. As a result, the adoption and long-term success of testing requires thinking about how it will fit into your company, not just philosophically and culturally, but *organizationally*. In this chapter we'll explore several different approaches and look at how different organizations have worked testing into their organizations.

## The Centralized Team

At every place Adidas's Scott Zakrajsek has worked, the testing team has started as a team of one or two.

At Staples I was the only person and we basically had to hire up from there, so we hired a front-end web-developer who was just for testing. We hired a creative designer who was just for testing; we hired an additional two analysts.

I think your key players are someone with a good project management background to manage the pipeline. Someone with attention to detail to be the coordinator and make sure you have all the right creative assets. You need a web-analytics person to read the results and do post-test segmentation. That said, smaller companies can do it well with teams of one person. The "beg-borrow-steal" method works well, too.

Zakrajsek currently runs the global analytics and optimization team for Adidas. His team handles all A/B testing responsibilities companywide, including executing tests and delivering results; if anyone has an idea for a test, they tell Zakrajsek, and his team knows what to do with it. The strategy over the longer term, however, is to make this team the centralized keeper of institutional testing knowledge while eventually *decentralizing* the test execution: entrusting product managers with responsibility of running tests for different areas of the site (which we'll discuss ahead).

## The Decentralized Team

The alternative to a centralized structure, where there is a single testing team to whom all other departments come for testing ideas and/or execution, is a *decentralized* structure, where each product manager (PM) has autonomy, the responsibility for a different part of the website, and the authority for testing that set of pages.

CareerBuilder is another company that has moved in this direction. “Previously most of our testing had resided within our development area,” explains Senior Internal Business Systems Analyst David Harris. “If one of the product owners were looking to experiment with something, they had to submit something down into our development group, who would make those changes and test them on their own, so [testing] was kind of siloed out by itself a little.”

The adoption of a no-code-necessary A/B testing platform enabled a change. “Part of what we were eager to do was really to move that into the hands of people who had a more direct stake in that particular page or area, to be able to make those changes themselves and then bring a data-supported request to

development as opposed to just submitting the test and waiting for the result.”

Harris is the point person for all things A/B testing from an administrative standpoint: he trains every new tester on how to use the tool and best practices, and if people have quality assurance questions (like how long to run a test), he takes care of it. However, he is *not* part of the process whereby product owners decide what to test. And he explains why: “We want to give people room to have creativity and freedom to do what they’re doing within those [product] areas.”

A decentralized model allows for greater independence, and reduces the potential for an organizational bottleneck, as each team can be testing its own part of the site in parallel. However, the disadvantage in comparison to a centralized testing team is that these separate testers face the challenge of staying coordinated and in communication about their results and best practices.

At Netflix, A/B testing also happens at the product manager level: each PM operates testing and analytics for a specific streaming platform such as Xbox, web, or tablet.

Bryan Gumm is Manager of Experimentation for the PS3 and Wii consoles, and has worked on various other platforms in the past. “We change that up about every three to six months just so all of the analysts are well-versed in all areas of the product,” he says.

Twice a week, the vice presidents of product and product managers meet for strategy meetings to review and analyze test results and vet ideas for future tests. “Every six to eight weeks, there’s a Customer Science Meeting with our CEO and basically all of the C-levels and we present what we’re testing, what has been rolled out, and what is testing and is not being rolled out because it didn’t work,” Gumm explains.

CareerBuilder has also focused on how to balance the autonomy of their different testing groups with the need for

coordination among them, and part of their solution has involved the creation of internal distribution lists around testing. “We basically encourage anyone who is conducting a test to push out a communication to everyone that’s within that distribution before they set it, or at the time they set it live,” Harris explains.

In that communication, the tester includes screenshots of the original page and variations being tested and a short write-up of the test goal. When the test has reached a statistically significant conclusion, the tester sends a follow-up with results and the key takeaways. Communicating findings and best practices is especially important for the internal team when the test requires code to set up.

With testing teams in 47 different countries, Dell takes a similar approach to transparency into what’s being tested and what’s coming up. “You’ve got to make sure you have a process to bring everybody together,” explains Dell Marketing Director Ed Wu. “We have a meeting every two weeks with all the global and regional stakeholders and we say, ‘Guys, here’s the goal we have over the next cycle, here’s the test idea we have,’ and they can all lock in on the same test idea.”

## **The Three Key Ingredients of a Scalable Testing Strategy**

At Optimizely we’ve worked with companies of every size who are establishing testing teams, and over the years we’ve noticed that the most successful teams included four key elements:

1. A point person

What we have heard again and again from the companies we’ve worked with is that no matter how large or small

the team of people executing A/B tests is, at least one person has to live, breathe, and evangelize optimization. The first step is to give someone ownership of testing. This individual might have other jobs besides A/B testing: Lizzie Allen, for instance, held this role alongside her other data-analyst duties at IGN.

If no one at your company is willing to take on the responsibilities, we encourage you to invest in hiring someone who can dedicate themselves to A/B testing. While you know by this point that A/B testing is tremendously valuable, it's easy to neglect it in favor of more "urgent" tasks. There's usually a blog post, a product release, or a code review that can take precedence. In reality, in terms of the bang for the buck, A/B testing is one of the best possible investments of an organization's resources because it makes every dollar spent on other marketing activities more effective. Without clear direction on who will *own* A/B testing, however, you risk its becoming another item on the to-do list. To ensure that it gets done, make it part of a job description.

## 2. Advocates across your organization

Your A/B testing point person (or team) won't be effective in isolation. They'll need allies across your marketing department and your engineering and product teams to be successful. This is not only because of the help they can provide in generating hypotheses and implementing experiments, but also because they're key in ensuring that testing is a regular part of the product/marketing/design planning process.

As Wu explains, allies are pivotal to growing a testing culture:

At Dell we have global stakeholders, the teams responsible for longer testing programs, and we also have

regional stakeholders that are responsible for day-to-day, weekly, monthly revenue generation, lead generation, etc. Most of [the regional stakeholders] have become believers of A/B testing and analytics. They have become champions for us. The testing teams ourselves, we are basically the carpenters—we founded our program and over time [the regional stakeholders] have become strong believers of ours so we can continue to fund the growth of our program.

### 3. Enabled testers

Empowering and enabling the people in charge of running the tests is one of the most important components of building a testing culture. The most successful customers we work with are the teams who allow people to test with creativity and resources.

What do we mean by this? Testing cannot work without *creativity*, that is, thinking outside of what exists on your site today and having the willingness to test it. Successful testing also requires certain *resources*: training, design/engineering work, educational tools, best practice information, and institutional encouragement. It's vital to create an environment where the people doing testing feel enabled and have the resources they need for the team to reach its full potential.

### 4. A track record

As the volume and impact of your tests increases, so does the need to keep track of what you're testing. First, come up with a *naming convention* for your tests. We recommend using *initials* when naming tests: a lone tester today may be part of a bigger team in a year, and so it's important to know who ran the test. You need to establish accountability and transparency early on.

Once you have a naming convention for your tests, start a log of tests run with their results: what worked, what didn't work, and why. As your company (or testing team) grows, it becomes crucial to educate new people on the past lessons. It's also extremely wise to have a record that will persist even if the tester leaves the company or role. It's dangerous to trust one person within an entire organization to keep this testing history in his or her head: if that person leaves, you risk losing all of those pieces of wisdom. Keep a shared document or internal wiki of your tests and include *screenshots*. It will be a treasure trove to you.

Finally, if you run a test that requires a particularly complex or nuanced technical integration, use the log to record how you did it. Teams down the road may want to run a similar test, and you'll save them from having to reinvent the wheel.



## TL;DR

- A/B testing is by nature **interdisciplinary** and **cross-departmental**. Collaboration is key.
- Some companies have a **centralized testing team** responsible for coordinating and executing all A/B tests companywide.
- Other organizations adopt a **decentralized** model, where each product owner/manager is responsible for testing within his or her product domain.
- Regardless of which structure you adopt, make sure there is at least one **point person** within the organization whom people can come to about all things testing.
- Ensure your point person or team maintains **allies across your organization** to ensure that testing is part of your planning process.
- Make sure that your testers are enabled and **empowered to be creative**, and that the friction of running and iterating new tests is low.
- **Maintain records** about who tested what and when, how the test was set up, and what the result was. This will enable your organization to work collaboratively to build up a set of collective wisdom and best practices.