CHAPTER

11

How A/B Tests Can Go Awry

Potential Mistakes and Pitfalls to Avoid

At this point in the book the hazards of *not* A/B testing should be clear; however, it's worth mentioning that A/B testing has its own set of pitfalls and "worst-practices."

Testing without Traffic

The good news is that you need only two things to conduct an A/B test: a website with some content on it, and visitors. The more traffic you have the faster you will see statistically significant results about how each variation performed.

What A/B testing your site *can't* do, however, is generate that traffic in the first place. A blogger who's just getting off the ground and has only 100 visitors per month would be better off focusing primarily on *content* and building a following of users (bolstered perhaps by SEO or paid ads) who provide traffic to the site *before* delving into the statistics of optimizing that traffic. After all, you have to generate the traffic in the first place before you do anything with it. (In addition, in a site's fledgling period, a handful of conversations with real users will offer more feedback than you will get from an A/B test on a sparsely trafficked site.) While optimization can help even the smallest enterprise, it's also true that testing becomes faster, more precise, and more profitable the more user traffic you have to work with.

The Beginning of the Funnel versus the End: UserVoice

Testing will occasionally reveal a change that increases one metric but decreases another. You'll recall from Chapter 2 that defining quantifiable success metrics is the first step in any test, and so it's worth reflecting on some nuances and complications around how to evaluate whether a test variation is in fact a "win."

The testing team at online help-desk software provider UserVoice hypothesized that removing fields from their free-trial signup form would increase the trial signup rate (Figure 11.1).

After the UserVoice team A/B tested the page, they found that, indeed, the fewer fields they required, the more people signed up for a free trial. However, they began to feel that perhaps
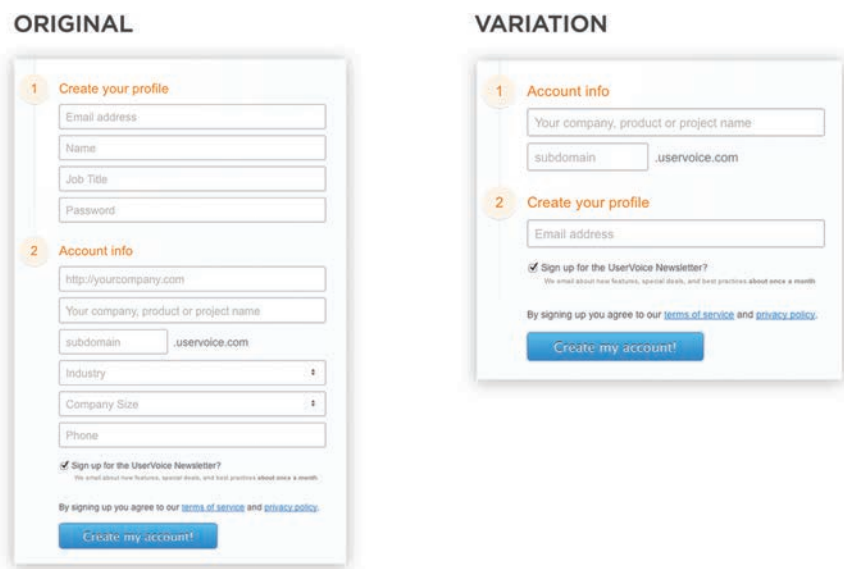
FIGURE 11.1 Original UserVoice signup form versus variation form with fields removed.

they'd made it *too* easy to run a trial. A plethora of people had begun trials, but the sales team was finding it increasingly difficult to determine who the best leads were.

The first thing they needed to do was decide what kind of leads they were going after. Were they seeking individuals who were well educated and informed about what UserVoice was and what it could offer them? Or did UserVoice plan to educate these people via email marketing *after* they became leads? Both scenarios are plausible and reasonable.

Here's where the *key performance indicators (KPIs)* began to clash. The marketing team was motivated to reach the broadest audience possible, while the sales team was motivated to turn the highest percentage of the most promising leads into customers. With the question of the free trial, it wasn't immediately clear which goal was best served in which way.

The UserVoice team began to notice that homing in on immediate metrics like the percentage of users that went from one page of their signup funnel to the next made it easy to get the wrong result from their A/B tests. "It's what I call *pushing failure down the funnel*," says co-founder and CEO Richard White. "We want quality leads, not quantity leads," he explains.

> Getting somebody to a page and getting somebody to use a product day-in and day-out are two different things. You end up with this insidious kind of race to the bottom, where you just want to remove all of the fields and everything and just say, "Go get your trial," and then they get into the trial and they have no idea why they're there.
>
> The goal will be a marketing site which should be educating them to get to the next step, so I think the tricky thing about A/B testing is the right variation may be the one where 80 percent of the people drop out on one page because they're

thinking this is not for them, but the other 20 percent of the people love it and sign up and try it.

Off-Brand Testing: Dell

There's a joke among A/B testing veterans that almost any variation of a button loses to a button that says, "Free Beer." But that doesn't mean that every company should liquidate its assets and go into the brewing business. More broadly, it's important that every company put its testing process into the broader context of its identity as a company.

Eric Ries, author of *The Lean Startup*, has witnessed many a cautionary tale, where a company's A/B test results have run it off the rails. "Abdicating product vision is a very common danger with A/B testing," says Ries. He recalls talking with an ex-employee of a company that went under. Ries asked the former employee what went wrong, and the employee explained, "Yeah, if you're on a social platform, and if all you do is A/B test all the time then you immediately get yourself into sexting, porn, and nasty stuff because that's what converts."

Not every company finds itself falling into quite such a dramatic identity crisis, but the broader point is one that almost any organization can relate to: how to balance customer feedback and brand identity. For example, an e-commerce site may find that displaying sale prices bigger and with big red strike-through lines improves their conversions. But that increase in conversions could cost them something potentially more valuable: their brand. The site may begin to look like a discounter when it's really a boutique. "It comes down to understanding who you are," says Chrome Industries e-commerce Director Kyle Duford. "Just because you can doesn't mean you should."

Many companies are extremely particular about their brand image to the point that they are not willing to test things that they wouldn't put live on their site.

It's sound advice. A question worth asking about every test is, "Would you be happy showing the winning variation to all of your traffic?" If not, then what exactly are you hoping to gain from the test?

There are several valid answers to this question, and it's worth pointing out that many companies do in fact experiment with things they wouldn't necessarily be committed to rolling out at full scale. One such company is Dell, where Marketing Director Ed Wu argues that there *can* be a place for testing variants of a page that the company isn't interested in implementing in the near-term.

Using A/B testing to learn about how visitors engage with certain types of content or color or messaging can be illuminating and can feed back into later conversations with Dell's brand team or the global site design team. "Our global design team has very consistent and very stringent guidelines in terms of what color you can use on dell.com," says Wu. The global guidelines for dell.com call for a blue banner of a specific hue at the bottom of the page, and Dell's testing team was interested in pushing back on that guideline to see whether that blue is in fact the best color to have on the site. So Wu's testing team worked with Dell's branding team and global design team to come up with some alternatives, including a bottom banner that was red. "[We] know that it's violating our brand standard, we would not be able to implement that, but eventually we'll all agree, let's go ahead and test it and understand how it works," says Wu. Testing for learning and understanding, not just conversions, allows Dell to "push the thinking" on what is optimal for the site.

TL;DR

- A/B testing can help even the smallest of organizations and sites, but it becomes more powerful the **more traffic** a site has. If your site is starved for traffic, A/B testing probably shouldn't be your first priority.

- Testing will occasionally reveal a change that increases one metric but decreases another. **Don't push failure down the funnel**.

- Consider whether you are willing to test a page variant that is in some way **off-brand** or one that you wouldn't necessarily be quick to roll out to all your users should it "win" the test.