

4

EXPERIMENT

I come across many startups that are struggling to answer the following questions: Which customer opinions should we listen to, if any? How should we prioritize across the many features we could build? Which features are essential to the product's success and which are ancillary? What can be changed safely, and what might anger customers? What might please today's customers at the expense of tomorrow's? What should we work on next?

These are some of the questions teams struggle to answer if they have followed the “let's just ship a product and see what happens” plan. I call this the “just do it” school of entrepreneurship after Nike's famous slogan.¹ Unfortunately, if the plan is to see what happens, a team is guaranteed to succeed—at seeing what happens—but won't necessarily gain validated learning. This is one of the most important lessons of the scientific method: if you cannot fail, you cannot learn.

FROM ALCHEMY TO SCIENCE

The Lean Startup methodology reconceives a startup's efforts as experiments that test its strategy to see which parts are brilliant and which are crazy. A true experiment follows the scientific method. It begins with a clear hypothesis that makes predictions about what is supposed to happen. It then tests those predictions empirically. Just as scientific experimentation is informed by theory, startup experimentation is guided by the startup's vision. The goal of every

startup experiment is to discover how to build a sustainable business around that vision.

Think Big, Start Small

Zappos is the world's largest online shoe store, with annual gross sales in excess of \$1 billion. It is known as one of the most successful, customer-friendly e-commerce businesses in the world, but it did not start that way.

Founder Nick Swinmurn was frustrated because there was no central online site with a great selection of shoes. He envisioned a new and superior retail experience. Swinmurn could have waited a long time, insisting on testing his complete vision complete with warehouses, distribution partners, and the promise of significant sales. Many early e-commerce pioneers did just that, including infamous dot-com failures such as Webvan and [Pets.com](#).

Instead, he started by running an experiment. His hypothesis was that customers were ready and willing to buy shoes online. To test it, he began by asking local shoe stores if he could take pictures of their inventory. In exchange for permission to take the pictures, he would post the pictures online and come back to buy the shoes at full price if a customer bought them online.

Zappos began with a tiny, simple product. It was designed to answer one question above all: is there already sufficient demand for a superior online shopping experience for shoes? However, a well-designed startup experiment like the one Zappos began with does more than test a single aspect of a business plan. In the course of testing this first assumption, many other assumptions were tested as well. To sell the shoes, Zappos had to interact with customers: taking payment, handling returns, and dealing with customer support. This is decidedly different from market research. If Zappos had relied on existing market research or conducted a survey, it could have asked what customers thought they wanted. By building a product instead, albeit a simple one, the company learned much more:

1. It had more accurate data about customer demand because it was observing real customer behavior, not asking hypothetical questions.
2. It put itself in a position to interact with real customers and learn about their needs. For example, the business plan might call for discounted pricing, but how are customer perceptions of the product affected by the discounting strategy?
3. It allowed itself to be surprised when customers behaved in unexpected ways, revealing information Zappos might not have known to ask about. For example, what if customers returned the shoes?

Zappos' initial experiment provided a clear, quantifiable outcome: either a sufficient number of customers would buy the shoes or they would not. It also put the company in a position to observe, interact with, and learn from real customers and partners. This qualitative learning is a necessary companion to quantitative testing. Although the early efforts were decidedly small-scale, that did not prevent the huge Zappos vision from being realized. In fact, in 2009 Zappos was acquired by the e-commerce giant [Amazon.com](https://www.amazon.com) for a reported \$1.2 billion.²

For Long-Term Change, Experiment Immediately

Caroline Barlerin is a director in the global social innovation division at Hewlett-Packard (HP), a multinational company with more than three hundred thousand employees and more than \$100 billion in annual sales. Caroline, who leads global community involvement, is a social entrepreneur working to get more of HP's employees to take advantage of the company's policy on volunteering.

Corporate guidelines encourage every employee to spend up to four hours a month of company time volunteering in his or her community; that volunteer work could take the form of any philanthropic effort, painting fences, building houses, or even using

pro bono or work-based skills outside the company. Encouraging the latter type of volunteering was Caroline's priority. Because of its talent and values, HP's combined workforce has the potential to have a monumental positive impact. A designer could help a nonprofit with a new website design. A team of engineers could wire a school for Internet access.

Caroline's project is just beginning, and most employees do not know that this volunteering policy exists, and only a tiny fraction take advantage of it. Most of the volunteering has been of the low-impact variety, involving manual labor, even when the volunteers were highly trained experts. Barlerin's vision is to take the hundreds of thousands of employees in the company and transform them into a force for social good.

This is the kind of corporate initiative undertaken every day at companies around the world. It doesn't look like a startup by the conventional definition or what we see in the movies. On the surface it seems to be suited to traditional management and planning. However, I hope the discussion in [Chapter 2](#) has prompted you to be a little suspicious. Here's how we might analyze this project using the Lean Startup framework.

Caroline's project faces extreme uncertainty: there had never been a volunteer campaign of this magnitude at HP before. How confident should she be that she knows the real reasons people aren't volunteering? Most important, how much does she really know about how to change the behavior of hundreds of thousand people in more than 170 countries? Barlerin's goal is to inspire her colleagues to make the world a better place. Looked at that way, her plan seems full of untested assumptions—and a lot of vision.

In accordance with traditional management practices, Barlerin is spending time planning, getting buy-in from various departments and other managers, and preparing a road map of initiatives for the first eighteen months of her project. She also has a strong accountability framework with metrics for the impact her project should have on the company over the next four years. Like many entrepreneurs, she has a business plan that lays out her intentions nicely. Yet despite all that work, she's so far—creating one-off

wins and no closer to knowing if her vision will be able to scale.

One assumption, for example, might be that the company's long-standing values included a commitment to improving the community but that recent economic trouble had resulted in an increased companywide strategic focus on short-term profitability. Perhaps longtime employees would feel a desire to reaffirm their values of giving back to the community by volunteering. A second assumption could be that they would find it more satisfying and therefore more sustainable to use their actual workplace skills in a volunteer capacity, which would have a greater impact on behalf of the organizations to which they donated their time. Also lurking within Caroline's plans are many practical assumptions about employees' willingness to take the time to volunteer, their level of commitment and desire, and the way to best reach them with her message.

The Lean Startup model offers a way to test these hypotheses rigorously, immediately, and thoroughly. Strategic planning takes months to complete; these experiments could begin immediately. By starting small, Caroline could prevent a tremendous amount of waste down the road without compromising her overall vision. Here's what it might look like if Caroline were to treat her project as an experiment.

Break It Down

The first step would be to break down the grand vision into its component parts. The two most important assumptions entrepreneurs make are what I call the value hypothesis and the growth hypothesis.

The value hypothesis tests whether a product or service really delivers value to customers once they are using it. What's a good indicator that employees find donating their time valuable? We could survey them to get their opinion, but that would not be very accurate because most people have a hard time assessing their feelings objectively.

Experiments provide a more accurate gauge. What could we see in real time that would serve as a proxy for the value participants were gaining from volunteering? We could find opportunities for a small number of employees to volunteer and then look at the retention rate of those employees. How many of them sign up to volunteer again? When an employee voluntarily invests their time and attention in this program, that is a strong indicator that they find it valuable.

For the growth hypothesis, which tests how new customers will discover a product or service, we can do a similar analysis. Once the program is up and running, how will it spread among the employees, from initial early adopters to mass adoption throughout the company? A likely way this program could expand is through viral growth. If that is true, the most important thing to measure is behavior: would the early participants actively spread the word to other employees?

In this case, a simple experiment would involve taking a very small number—a dozen, perhaps—of existing long-term employees and providing an exceptional volunteer opportunity for them. Because Caroline's hypothesis was that employees would be motivated by their desire to live up to HP's historical commitment to community service, the experiment would target employees who felt the greatest sense of disconnect between their daily routine and the company's expressed values. The point is not to find the average customer but to find early adopters: the customers who feel the need for the product most acutely. Those customers tend to be more forgiving of mistakes and are especially eager to give feedback.

Next, using a technique I call the concierge minimum viable product (described in detail in [Chapter 6](#)), Caroline could make sure the first few participants had an experience that was as good as she could make it, completely aligned with her vision. Unlike in a focus group, her goal would be to measure what the customers actually did. For example, how many of the first volunteers actually complete their volunteer assignments? How many volunteer a second time? How many are willing to recruit a colleague to

participate in a subsequent volunteer activity?

Additional experiments can expand on this early feedback and learning. For example, if the growth model requires that a certain percentage of participants share their experiences with colleagues and encourage their participation, the degree to which that takes place can be tested even with a very small sample of people. If ten people complete the first experiment, how many do we expect to volunteer again? If they are asked to recruit a colleague, how many do we expect will do so? Remember that these are supposed to be the kinds of early adopters with the most to gain from the program.

Put another way, what if all ten early adopters decline to volunteer again? That would be a highly significant—and very negative—result. If the numbers from such early experiments don't look promising, there is clearly a problem with the strategy. That doesn't mean it's time to give up; on the contrary, it means it's time to get some immediate qualitative feedback about how to improve the program. Here's where this kind of experimentation has an advantage over traditional market research. We don't have to commission a survey or find new people to interview. We already have a cohort of people to talk to as well as knowledge about their actual behavior: the participants in the initial experiment.

This entire experiment could be conducted in a matter of weeks, less than one-tenth the time of the traditional strategic planning process. Also, it can happen in parallel with strategic planning while the plan is still being formulated. Even when experiments produce a negative result, those failures prove instructive and can influence the strategy. For example, what if no volunteers can be found who are experiencing the conflict of values within the organization that was such an important assumption in the business plan? If so, congratulations: it's time to pivot (a concept that is explored in more detail in [Chapter 8](#)).³

AN EXPERIMENT IS A PRODUCT

In the Lean Startup model, an experiment is more than just a

theoretical inquiry; it is also a first product. If this or any other experiment is successful, it allows the manager to get started with his or her campaign: enlisting early adopters, adding employees to each further experiment or iteration, and eventually starting to build a product. By the time that product is ready to be distributed widely, it will already have established customers. It will have solved real problems and offer detailed specifications for what needs to be built. Unlike a traditional strategic planning or market research process, this specification will be rooted in feedback on what is working today rather than in anticipation of what might work tomorrow.

To see this in action, consider an example from Kodak. Kodak's history is bound up with cameras and film, but today it also operates a substantial online business called Kodak Gallery. Mark Cook is Kodak Gallery's vice president of products, and he is working to change Kodak Gallery's culture of development to embrace experimentation.

Mark explained, "Traditionally, the product manager says, 'I just want this.' In response, the engineer says, 'I'm going to build it.' Instead, I try to push my team to first answer four questions:

1. Do consumers recognize that they have the problem you are trying to solve?
2. If there was a solution, would they buy it?
3. Would they buy it from us?
4. Can we build a solution for that problem?"

The common tendency of product development is to skip straight to the fourth question and build a solution before confirming that customers have the problem. For example, Kodak Gallery offered wedding cards with gilded text and graphics on its site. Those designs were popular with customers who were getting married, and so the team redesigned the cards to be used at other special occasions, such as for holidays. The market research and design process indicated that customers would like the new cards, and that

finding justified the significant effort that went into creating them.

Days before the launch, the team realized the cards were too difficult to understand from their depiction on the website; people couldn't see how beautiful they were. They were also hard to produce. Cook realized that they had done the work backward. He explained, "Until we could figure out how to sell and make the product, it wasn't worth spending any engineering time on."

Learning from that experience, Cook took a different approach when he led his team through the development of a new set of features for a product that makes it easier to share photos taken at an event. They believed that an online "event album" would provide a way for people who attended a wedding, a conference, or another gathering to share photos with other attendees. Unlike other online photo sharing services, Kodak Gallery's event album would have strong privacy controls, assuring that the photos would be shared only with people who attended the same event.

In a break with the past, Cook led the group through a process of identifying risks and assumptions before building anything and then testing those assumptions experimentally.

There were two main hypotheses underlying the proposed event album:

1. The team assumed that customers would want to create the albums in the first place.
2. It assumed that event participants would upload photos to event albums created by friends or colleagues.

The Kodak Gallery team built a simple prototype of the event album. It lacked many features—so many, in fact, that the team was reluctant to show it to customers. However, even at that early stage, allowing customers to use the prototype helped the team refute their hypotheses. First, creating an album was not as easy as the team had predicted; none of the early customers were able to create one. Further, customers complained that the early product version lacked essential features.

Those negative results demoralized the team. The usability

them to dry, which takes two to seven days. The result? Clothes are returned in about ten days and are probably not that clean.

Akshay Mehra had been working at Procter & Gamble Singapore for eight years when he sensed an opportunity. As the brand manager of the Tide and Pantene brands for India and ASEAN countries, he thought he could make laundry services available to people who previously could not afford them. Returning to India, Akshay joined the Village Laundry Services (VLS), created by Innosight Ventures. VLS began a series of experiments to test its business assumptions.

For their first experiment, VLS mounted a consumer-grade laundry machine on the back of a pickup truck parked on a street corner in Bangalore. The experiment cost less than \$8,000 and had the simple goal of proving that people would hand over their laundry and pay to have it cleaned. The entrepreneurs did not clean the laundry on the truck, which was more for marketing and show, but took it off-site to be cleaned and brought it back to their customers by the end of the day.

The VLS team continued the experiment for a week, parking the truck on different street corners, digging deeper to discover all they could about their potential customers. They wanted to know how they could encourage people to come to the truck. Did cleaning speed matter? Was cleanliness a concern? What were people asking for when they left their laundry with them? They discovered that customers were happy to give them their laundry to clean. However, those customers were suspicious of the washing machine mounted on the back of the truck, concerned that VLS would take their laundry and run. To address that concern, VLS created a slightly more substantial mobile cart that looked more like a kiosk.

VLS also experimented with parking the carts in front of a local minimarket chain. Further iterations helped VLS figure out which services people were most interested in and what price they were willing to pay. They discovered that customers often wanted their clothes ironed and were willing to pay double the price to get their laundry back in four hours rather than twenty-four hours.

As a result of those early experiments, VLS created an end

product that was a three-foot by four-foot mobile kiosk that included an energy-efficient, consumer-grade washing machine, a dryer, and an extra-long extension cord. The kiosk used Western detergents and was supplied daily with fresh clean water delivered by VLS.

Since then, the Village Laundry Service has grown substantially, with fourteen locations operational in Bangalore, Mysore, and Mumbai. As CEO Akshay Mehra shared with me, “We have serviced 116,000 kgs. in 2010 (vs. 30,600 kg. in 2009). And almost 60 percent of the business is coming from repeat customers. We have serviced more than 10,000 customers in the past year alone across all the outlets.”⁵

A LEAN STARTUP IN GOVERNMENT?

On July 21, 2010, President Obama signed the Dodd–Frank Wall Street Reform and Consumer Protection Act into law. One of its landmark provisions created a new federal agency, the Consumer Financial Protection Bureau (CFPB). This agency is tasked with protecting American citizens from predatory lending by financial services companies such as credit card companies, student lenders, and payday loan offices. The plan calls for it to accomplish this by setting up a call center where trained case workers will field calls directly from the public.

Left to its own devices, a new government agency would probably hire a large staff with a large budget to develop a plan that is expensive and time-consuming. However, the CFPB is considering doing things differently. Despite its \$500 million budget and high-profile origins, the CFPB is really a startup.

President Obama tasked his chief technology officer, Aneesh Chopra, with collecting ideas for how to set up the new startup agency, and that is how I came to be involved. On one of Chopra’s visits to Silicon Valley, he invited a number of entrepreneurs to make suggestions for ways to cultivate a startup mentality in the new agency. In particular, his focus was on leveraging technology

and innovation to make the agency more efficient, cost-effective, and thorough.

My suggestion was drawn straight from the principles of this chapter: treat the CFPB as an experiment, identify the elements of the plan that are assumptions rather than facts, and figure out ways to test them. Using these insights, we could build a minimum viable product and have the agency up and running—on a micro scale—long before the official plan was set in motion.

The number one assumption underlying the current plan is that once Americans know they can call the CFPB for help with financial fraud and abuse, there will be a significant volume of citizens who do that. This sounds reasonable, as it is based on market research about the amount of fraud that affects Americans each year. However, despite all that research, it is still an assumption. If the actual call volume differs markedly from that in the plan, it will require significant revision. What if Americans who are subjected to financial abuse don't view themselves as victims and therefore don't seek help? What if they have very different notions of what problems are important? What if they call the agency seeking help for problems that are outside its purview?

Once the agency is up and running with a \$500 million budget and a correspondingly large staff, altering the plan will be expensive and time-consuming, but why wait to get feedback? To start experimenting immediately, the agency could start with the creation of a simple hotline number, using one of the new breed of low-cost and fast setup platforms such as Twilio. With a few hours' work, they could add simple voice prompts, offering callers a menu of financial problems to choose from. In the first version, the prompts could be drawn straight from the existing research. Instead of a caseworker on the line, each prompt could offer the caller useful information about how to solve her or his problem.

Instead of marketing this hotline to the whole country, the agency could run the experiment in a much more limited way: start with a small geographic area, perhaps as small as a few city blocks, and instead of paying for expensive television or radio advertising to let people know about the service, use highly targeted advertising.

Flyers on billboards, newspaper advertisements to those blocks, or specially targeted online ads would be a good start. Since the target area is so small, they could afford to pay a premium to create a high level of awareness in the target zone. The total cost would remain quite small.

As a comprehensive solution to the problem of financial abuse, this minimum viable product is not very good compared with what a \$500 million agency could accomplish. But it is also not very expensive. This product could be built in a matter of days or weeks, and the whole experiment probably would cost only a few thousand dollars.

What we would learn from this experiment would be invaluable. On the basis of the selections of those first callers, the agency could immediately start to get a sense of what kinds of problems Americans believe they have, not just what they “should” have. The agency could begin to test marketing messages: What motivates people to call? It could start to extrapolate real-world trends: What percentage of people in the target area actually call? The extrapolation would not be perfect, but it would establish a baseline behavior that would be far more accurate than market research.

Most important, this product would serve as a seed that could germinate into a much more elaborate service. With this beginning, the agency could engage in a continuous process of improvement, slowly but surely adding more and better solutions. Eventually, it would staff the hotline with caseworkers, perhaps at first addressing only one category of problems, to give the caseworkers the best chance of success. By the time the official plan was ready for implementation, this early service could serve as a real-world template.

The CFPB is just getting started, but already they are showing signs of following an experimental approach. For example, instead of doing a geographically limited rollout, they are segmenting their first products by use case. They have established a preliminary order of financial products to provide consumer services for, with credit cards coming first. As their first experiment unfolds, they will

have the opportunity to closely monitor all of the other complaints and consumer feedback they receive. This data will influence the depth, breadth, and sequence of future offerings.

As David Forrest, the CFPB's chief technology officer, told me, "Our goal is to give American citizens an easy way to tell us about the problems they see out there in the consumer financial marketplace. We have an opportunity to closely monitor what the public is telling us and react to new information. Markets change all the time and our job is to change with them."⁶



The entrepreneurs and managers profiled in this book are smart, capable, and extremely results-oriented. In many cases, they are in the midst of building an organization in a way consistent with the best practices of current management thinking. They face the same challenges in both the public and private sectors, regardless of industry. As we've seen, even the seasoned managers and executives at the world's best-run companies struggle to consistently develop and launch innovative new products.

Their challenge is to overcome the prevailing management thinking that puts its faith in well-researched plans. Remember, planning is a tool that only works in the presence of a long and stable operating history. And yet, do any of us feel that the world around us is getting more and more stable every day? Changing such a mind-set is hard but critical to startup success. My hope is that this book will help managers and entrepreneurs make this change.