

By Dave Sonders
and Lauren Braun

Why is successful innovation so rare? How do you predict success for a product that has never been seen before? The odds can be improved.

TRADITIONAL
VALIDATION

Validation in the Wild

THE MICRO PILOT

WHEN IT COMES TO INNOVATION AND GROWTH,

there are two fundamental questions:

- 1) What should we create? and
- 2) How do we know it will succeed?

In large companies, the first question is increasingly answered by design-driven innovation—most often referred to as design thinking. The process boils down to a fairly common set of methods executed in a roughly consistent order:

- Conduct qualitative user research to build empathy.
- Synthesize observations into insights.
- (Re)frame opportunity areas.
- Brainstorm with cross-functional teams.
- Prototype ideas and refine them through user feedback.
- Define vision for a new offering.
- Pitch it to company leadership.

Using this now widely adopted process, large organizations have radically improved their ability to come up with compelling answers to Innovation Question #1 (What should we create?). Although this is a fantastic stride forward, far too many promising ideas fail to make it into the

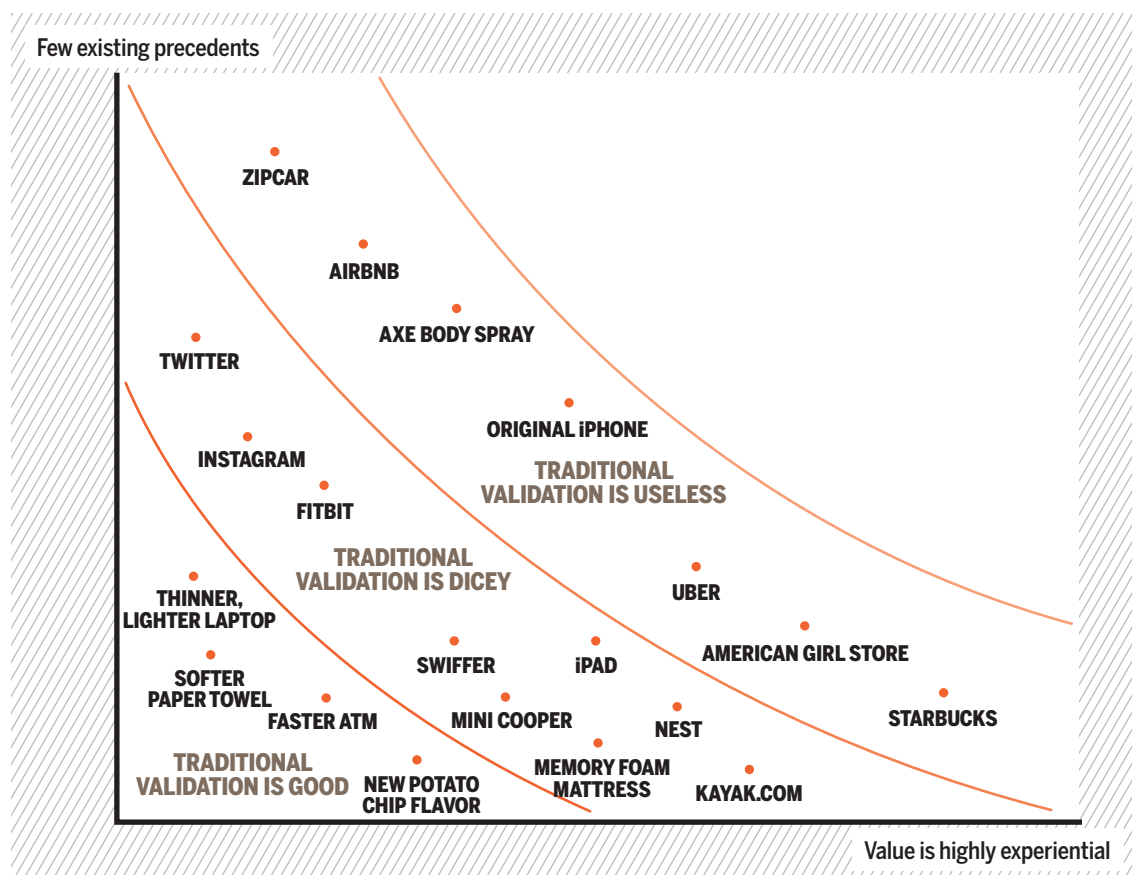
market. And of those that do, the vast majority fall short of expectations. Some estimates place the new product failure rate at 90 percent or more. So why is successful innovation so rare?

First, we have to acknowledge that innovation is inherently uncertain and that design-driven methods are not a panacea. But there's another problem lurking in plain view—a fatal but mostly overlooked bug in the operating system of corporate innovation. We believe that companies have no idea how to answer Innovation Question # 2. Today's most widely accepted validation tools simply can't provide good answers when project sponsors inevitably ask, "How do we know this idea will succeed?" The problem is that these tools were not designed for uncertain and ambiguous situations. By definition, this is what you're facing if you're trying to do something innovative. Bold new ideas without clear precedent are inherently unpredictable. If the success of a new idea can be predicted based on what we already know (that is, extrapolated from the past), then traditional validation techniques are great. If not, these tools are inappropriate.

Today's most widely accepted validation tools simply can't provide good answers when project sponsors inevitably ask, "How do we know this idea will succeed?"

FIGURE 1

Traditional validation works best for incremental improvements.



The problem with traditional validation

"You can't put into a spreadsheet how people are going to behave around a new product." - Jeff Bezos

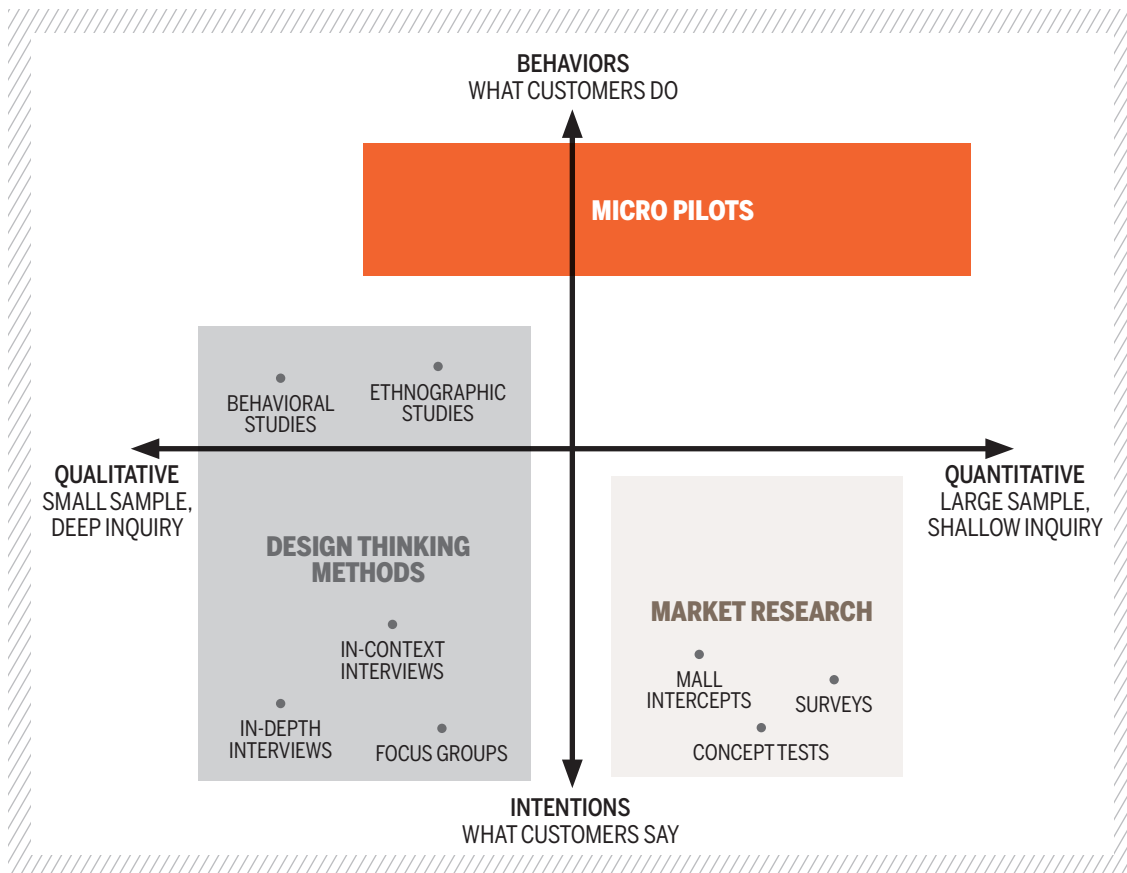
In order to validate new ideas and mitigate risk, large companies employ a number of methods, including sales forecasts, internal rate of return, net present value, and consumer surveys. Often these methods rely on market research data modeling to predict a new product's odds of success. These predictions are fairly reliable, given two specific conditions:

- 1) The new product is clearly related to something consumers already know.
- 2) It delivers tangible benefits that are easy to understand without direct experience.

Most new product and service concepts fit these criteria nicely—a new flavor of potato chip,

a softer paper towel, a faster way to withdraw cash at the ATM (Figure 1). For these types of ideas, consumer surveys are great tools for answering the How do we know? question. But when a new concept is substantially different from offerings already on the market and/or a big part of its value is tied up in the experience of using it, traditional validation tools are woefully inadequate.

Procter & Gamble's original Swiffer, for example, tested poorly on consumer surveys. BASES tests predicted low repeat rate that would never meet P&G's hurdle rate of \$100 million. However, the brand team's experience and positive user feedback suggested a much rosier picture. When Swiffer launched in 1999, it was the team, not the test, which was proved right. Swiffer generated \$100 million in sales in just four months. Some tools use data about existing

**FIGURE 2**

Micro pilots measure what consumers do, not what they say.

products to predict consumer response to new products. But Swiffer was not just an incremental improvement on the mop and bucket—it offered an entirely new way to clean.

The Swiffers of the world can't be accurately simulated with quantitative models and other traditional validation tools because the tools measure what people say, not what they do. To many in the corporate innovation space, this fact is obvious but irrelevant: "Of course we use imperfect tools—the product doesn't exist yet!" And therein lies the root of the problem: The new product does not yet exist. It's just an idea, and asking people what they think about an idea is a lot different from seeing how they react to a new item on the shelf or a new service experience.

It's a catch 22. We can't know how people will react to a new innovation until it's real, and we

don't want to invest millions of dollars to make it real until we know how people will react. This is why the How do we know? question has remained such a thorny issue for innovators.

The Micro Pilot

"Make a little. Sell a little. Learn a lot, and fail cheap." - P&G Chairman-CEO Durk Jager

There is a better way to approach the How do we know? question. Rather than rely on surveys and data extrapolation to guide decisions, we can invest small amounts of time and money to test limited versions of new offerings. These limited versions allow us to see how consumers behave in real life, not just what they say in a focus group or a survey. This approach has gained momentum in the startup world—spreading through the vernacular of Lean Startup, minimum viable

TYPES OF MICRO PILOTS

Several forms of experiments are starting to coalesce into a preliminary toolkit for startups and progressive corporate innovation teams. The following micro pilot examples are not about efficiency, scale, or profitability. They are targeted experiments that test the underlying business model hypotheses about a new product or service concept.

CROWDFUNDING CAMPAIGNS

Platforms like Kickstarter have revolutionized the way inventors and entrepreneurs secure funding to pursue their vision. Crowdfunding has also become a fantastic way for startup teams to test the waters to see if there really is demand for a new product. With a few images, a description of your idea, and maybe a video, you can ask consumers to back your project with real money. These campaigns are a great proxy for consumer demand because you're asking people to vote with their wallets.

Chicago-based startup Scout Alarm is disrupting the home security market with a simple, customizable, and design-forward security system. Founders Dan Roberts and Dave Shapiro ran a crowdfunding campaign to see if they could get customers to express interest in buying their system. With a rough prototype and an amateur product video, the team secured more than 1,500 pre-orders. This was enough to give the team (and their would-be investors) the confidence they needed to move into final development and production.

FALSE DOORS

You've probably been part of a "false door" test without knowing it. A false door is typically a simple web page that describes a new product or service and involves some kind of call to action—often a button that asks us to buy now or sign up for the free beta. When we click that button, we get a message that politely thanks us for our interest but informs us that the product is not yet available. Behind the scenes, the people running the test are tracking how many people are finding their way to the page and how many of them are clicking

that button. Many false doors also collect email addresses to build a prospect list for the time the product goes on to development.

One early example of a false door was Redfin, the innovative online real estate brokerage. When the company was in its infancy, most (possibly all) people who clicked the "I'm interested" button on Redfin's landing page were told that the service was not yet available in their area. Whether it was available in any area was irrelevant. The team was able to measure interest across the entire country before investing millions to build and scale the business.

CONCIERGE MVP

The term *conciierge MVP* (minimum viable product) comes straight from the Lean Startup movement. The basic idea is that you deliver, by hand, whatever the ultimate product or service would do automatically. Simulating the experience allows you to test the central value proposition with customers without having to build something automated and scalable.

The founders of dress-sharing service Rent the Runway developed three different *conciierge MVPs*

to test the central hypothesis of their business model—that women would rent a dress without trying it on. First, they purchased several dresses at retail and offered them in-person to Harvard undergrads. Women could try them on and, if they wanted, rent a dress for the night. This test helped the team gauge acceptance of the rental model, as well as get a sense for preferences around color, cut, brand, and price point. The founders ran a second test as a trunk show, but eliminated the try-on option. For their third test, they took orders from a PDF email showing dress options.

WIZARD OF OZ

Like a *conciierge MVP*, a Wizard of Oz experiment simulates the experience of a new offering without automating it. But in this case, the customer has no idea that the service is not automated. In fact, there are people behind the curtain pulling levers and pushing buttons to make the service work.

Before Zappos became the internet's most dominant shoe store, the founding team ran an

ingenious experiment to find out if people would order shoes from a website. They put up a simple online retail page and took orders for shoes. But rather than building a fully automated e-commerce business with warehouses and sophisticated order fulfillment, the founders simply ran out to a store, purchased shoes on a credit card, and shipped them to customers.

A micro pilot is a quick, inexpensive experiment that allows us to validate ideas and business models with real consumer behavior.

product, and customer development—but is still relatively unknown in most large companies. At Gravitytank, we refer to these real-world product experiments as *micro pilots* (Figure 2, p.23).

A micro pilot is a quick, inexpensive experiment that allows us to validate ideas and business models with real consumer behavior. Every micro pilot is custom-built to test a specific hypothesis. It is a fast and efficient way to pressure-test crucial elements of a new offering—typically, the most risky and uncertain elements—before committing to a bigger investment. At Gravitytank we still use traditional concept validation techniques when appropriate, but we have found that micro pilots are an extremely valuable addition to our toolset, and a better way to de-risk innovation in many situations.

CONSIDER THE STORY OF ZIPCAR

In 2000, the term *sharing economy* did not exist, but Zipcar founders Antje Danielson and Robin Chase had a bold vision—a vast network of shared cars, available any time and nearly everywhere. The car-sharing idea was potentially disruptive to automakers and car rental companies, but it was also extremely uncertain. No one knew whether consumers would embrace the service because there was no precedent for it in the US (the most car-obsessed market in the world). The Zipcar team didn't conduct surveys or focus groups to test their idea. Instead, they bought a green VW Beetle and parked it at Chase's house. They recruited about 20 neighbors to share the car, using an online calendar for scheduling and a notepad in the glove box to record time and mileage. When someone wanted to take it out for a spin, they would sneak into her backyard and find the car key hiding beneath a cushion on the porch. This is how people used Zipcar for the first three months.

This experiment proved to the founders that they could find customers who were interested in trying a car sharing service, and that users could be trusted to contribute to a car's ongoing operating needs (refilling the gas, keeping reasonably accurate records, not leaving trash,

KEY PRINCIPLES FOR CONDUCTING MICRO PILOTS

FACTS OVER OPINIONS

Don't get hung up debating the merits of an idea. Replace opinions with facts by moving quickly from concept to experiment.

SURGICAL STRIKES

Focus on the most uncertain and riskiest parts of the business model. For instance, if your model assumes a specific user acquisition cost and conversion rate, conduct a small marketing campaign to see if you can hit the required numbers.

GET TANGIBLE

Potential users need something tangible to react to. Build lightweight versions of the offering to test with real consumers in the wild either by making it (build something that you intend to use later) or faking it (use a smoke-and-mirrors approach to simulate part of the experience).

BUILD AS LITTLE AS POSSIBLE

Don't build new assets just because you can. Prototypes are wonderful, but only if they help you learn.

SKIP THE ACCOUNTING

Focus solely on the amount of learning your test will produce for the investment of time and money. There's no reason to optimize for cost, scalability, or profit margin before you know if people want what you are selling.

PUT IT OUT IN THE WORLD

Behavioral economics research has demonstrated that people are really bad at predicting their own behavior—so don't ask them to. Put your offering out into the world and see what happens, gathering empirical evidence along the way. Someone choosing to buy your product is the ultimate feedback.

MEASURE RESULTS

One of the most important parts of your experiment design is defining what metrics will validate your hypotheses (for example, trial rate, viral coefficient, user acquisition cost), and also how to capture those metrics (for example, web clicks, in-store interactions, behavior over a period of time).

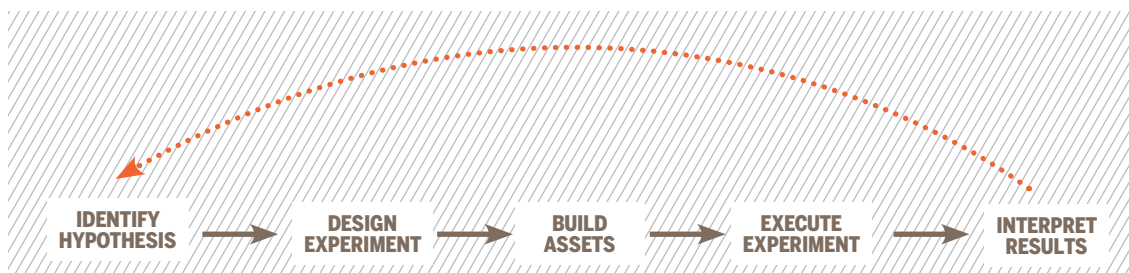
RINSE AND REPEAT

Don't expect to answer all your questions with a single micro pilot. Use what you learn from each micro pilot to refine your hypotheses and run new experiments.

and so on). Think about the investment in time and money it took to run this simple but effective test. The only real capital expense was the car (about \$16,000 at the time). Other than that, the financial cost was basically zero. And the team didn't write a single line of code or even put up a landing page. Now consider what it would have cost the founders to launch a full pilot test, even in just one city—dozens of cars, corporate insurance, car access systems (for example, RFID cards), website, marketing campaign, customer service

FIGURE 3

Micro pilots are based on the scientific method.



staff... Instead, Zipcar ran a small-scale test with a handful of customers and a single car, allowing them to validate their value proposition at a fraction of the time and expense. All it took was a bit of hustle, some imagination, and a few willing neighbors. This is the essence of micro pilots.

A micro pilot is essentially the scientific method applied to innovation (Figure 3). And like scientists pursuing new discoveries, innovation teams must be prepared for failed experiments. More precisely, they must be prepared for unexpected results. Much of the time, our experiments will not pan out the way we anticipate because our hypotheses and assumptions are wrong. Finding the flaws in our business model before scaling up is incredibly valuable. Each failed experiment allows us to tweak our hypotheses and try again. In this way, we steadily replace uncertainty with understanding, fog with clarity. And the faster we go, the better our chances of creating a sustainable new business before we run out of funding, time, or passion.

Rewiring the corporate innovation process

“We’re a collection of dozens of internal startups. This is now the standard practice.... How many weeks after having the idea can you get a version into users’ hands that tests key hypotheses? We call it leadership by experiment.” – Scott Cook, Intuit founder

A handful of large companies have started to transition toward a culture of experimentation. But the vast majority of corporate innovation teams are shackled to techniques that snuff out the most ambitious and potentially groundbreaking

ideas. In the typical corporate innovation process, the final pitch marks the end of prototyping and iteration—a handoff from the idea people to the development people. According to this philosophy, idea validation is all or nothing (pass = invest; fail = kill it). But the notion that you can simply run one test to see if something is a good idea is deeply flawed when it comes to innovation. Just because an idea fails its first test doesn’t mean it’s not a good idea. Many of the biggest innovations of our time went through major iterations before finding just the right mix of features, benefits, positioning, and pricing. The first incarnation of Starbucks had no chairs, baristas in bowties, menus written mostly in Italian, and nonstop opera music. Twitter started out as a platform for creating and sharing podcasts. YouTube was originally intended to be a video-dating site!

Successful innovations are most often the result of continued evolution—multiple iterations that gradually nudge a promising idea toward success. Micro piloting allows innovation teams to systematically refine and validate their hypotheses until they are confident that they have a sustainable, scalable business model. With micro pilots, there is no final pitch, no politicking, no committee that decides the fate of an idea on a whim. For teams empowered to use micro pilots to test their ideas “in the wild,” the only objective is to learn how consumers actually behave, iterating and shaping their way to success over a series of small experiments. And when sponsors ask Innovation Question #2, How do we know it will succeed? there’s no need to speculate. We know because we’ve seen it work, and we have the data to prove it. ■



Dave Sanders is a strategist and associate partner at Gravitytank, where he spearheads efforts to blend entrepreneurial methods with corporate innovation efforts. He holds a master of design methods degree from IIT’s Institute of Design and a BA in political science and economics from the University of Michigan. He is also an adjunct professor at the Institute of Design, where he teaches a capstone course on new product definition.



Lauren Braun is a strategist at Gravitytank. She helps clients define, shape, and test new offerings and craft stories compelling enough to move ideas from the slide onto the shelf. Prior to Gravitytank she worked in public affairs and international development. She holds a masters of design from IIT’s Institute of Design, as well as a BA in art history and a BBA, both from the University of Michigan.