

Stanford SOCIAL INNOVATION^{Review}

Features

The New Science of Designing for Humans

By Piyush Tania

Stanford Social Innovation Review
Spring 2017

Copyright © 2017 by Leland Stanford Jr. University
All Rights Reserved



The days of privileging creativity over science in design thinking are over. The rise of behavioral science and impact evaluation has created a new way for engineering programs and human interactions—a methodology called behavioral design.

The • New Science • of • Designing for • Humans

BY PIYUSH TANTIA

Illustration by MIKE AUSTIN

Today the design of things that involve human interaction, such as programs, product delivery, and services, is more art than science. Here is how it typically works: We use our creativity to brainstorm a few big ideas, experts decide which one they like, and then investors bet on the winner, often with billions of dollars at stake.

This way of design thinking should be replaced by a superior method that can enable us to innovate with more success and less risk. Specifically, we can use scientific insights to generate new ideas and then systematically test and iterate on them to arrive at one that works.

Advances in two academic fields afford this opportunity. The first is behavioral science, which gives us empirical insights into how people interact with their environment and each other under different conditions. Behavioral science encompasses decades of research from various fields, including psychology, marketing, neuroscience, and, most recently, behavioral economics. For example, studies reveal that shorter deadlines lead to greater responsiveness than longer ones,¹ that too much choice leads people to choose nothing,² and many more observations, often counterintuitive, about how people react to specific elements of their context.

The second academic field is impact evaluation. Economists have used randomized controlled trials

(RCTs) and other experimental methods to measure the impact of programs and policies. Such impact evaluations are becoming more and more common in the social sector and in government. These methods allow us to test whether an innovation actually achieves the outcomes that the designer sought.

Taking a scientific approach also solves another common problem: Sometimes we do not even realize that there is something in need of rigorous, thoughtful design. When we look carefully, the success of most of what we design for people depends as much, if not more, on the human interaction as on the physical product. For example, the first iPhone offered essentially the same functions (phone, calendar, address book, etc.) as a BlackBerry, but it totally changed the experience of using those functions.

In the social and public sectors, programs and services are made up largely of human interactions. And yet anything involving human interaction can be designed more scientifically, and more successfully, when behavioral science and impact evaluation are applied. For instance, a vaccine is a technological product, but how and when parents get their children vaccinated, and how they are reminded to do so, is as much a part of the innovation as the vaccine itself. Poorly designed interactions make products less successful and can also underlie serious social problems.³

By putting behavioral science and impact evaluation together—a methodology we call behavioral design—we can design more like engineers than like artists. We can use behavioral science to develop ideas that are much more likely to work than those relying entirely on intuition. And we can rigorously test those ideas to determine which ones truly work. Following the model of engineering and scientific progress, we can build on prior success to make enormous advances that, under previous approaches, would not be possible.

A BETTER METHODOLOGY

At ideas42, the behavioral science innovation lab I co-lead, we encounter many different approaches to innovation among our partners. I have also spent considerable time comparing notes with experts in design thinking, attending design workshops, and reading about design methodologies. The typical approaches for innovation range from quickly brainstorming some ideas in a boardroom to using some version of human-centered design (HCD). Fundamentally, all of these approaches aim to generate “big ideas” that appeal to the intuition of a few decision makers considered experts in the area where the idea is to be implemented.

HCD appears to be the methodology of choice for a significant, and growing, number of organizations. The most advanced version begins with defining the problem or design mandate, and then conducts qualitative research with potential users and proceeds through a series of structured exercises to promote creative thinking. The design team may also test some crude prototypes to get feedback along the way. This approach is called “human-centered” because it focuses on users’ and other stakeholders’ needs and preferences.

In the qualitative research phase, designers use ethnographic techniques such as qualitative interviewing and observation. They not only interview potential users but also may talk to others, such as program administrators and front-line staff involved in delivering a program or product. In the design phase, HCD employs several techniques to enhance creativity (which remain useful in the next-generation behavioral design methodology as well). Finally, HCD ends with trying a few prototypes with a handful of potential users. Some ethnographic research methods are incorporated into HCD, but on the whole the approach is still much closer to an art than a science.

It is time to build on HCD with a better method. Let us begin our investigation by comparing how engineers invent new technology. Two features stand out. First, engineers rely on a rich set of insights from science to develop new ideas. Every invention builds on countless previous attempts. For example, the Wright brothers are credited with inventing the airplane, but the key parts of their design leaned on previous inventions. The wing was based on science that went back to 1738, when Daniel Bernoulli discovered his principle about the relationship between pressure and the speed with which a fluid is moving. The engine design was borrowed from automotive engines invented more than 25 years earlier. They were able to test model wings in a wind tunnel thanks to Frank H. Wenham, who had invented that critical apparatus 30 years before that, in 1871.⁴

Second, contrary to popular belief, inventions do not come simply from a single flash of insight, but rather from painstaking refinement in small steps. Sir James Dyson, the famous vacuum cleaner tycoon, went through 5,126 failed iterations of his new wind tunnel design to separate dirt from air before he landed on the right one.⁵

PIYUSH TANTIA is co-executive director of ideas42, a social enterprise that uses insights from behavioral economics to invent fresh solutions to tough social problems. Since joining ideas42 in 2009, he has worked closely with leading academics from Harvard University, MIT, and Princeton University to design and implement solutions in various areas

including household finance, education, international development, poverty, criminal justice, and health care. Along with ideas42’s cofounders, he transitioned the organization from a research initiative at Harvard University to an independent 501(c)(3) nonprofit with 80 employees working on more than 80 projects across more than 30 countries.

Inventors sometimes iterate only on particular components before working on the complete invention. For example, the Wright brothers tested some 200 wing designs in a wind tunnel before settling on the right one.

Why do engineers work so differently from those of us who are designing for human interactions? Until recently, we did not have a sufficiently large body of scientific insights that describes how humans interact with their environment, and each other, under different conditions. True, the field of user-experience design offers some insights, but it is very new and is still restricted to certain elements of digital interactions such as Web-page layout and font size. Direct marketers within for-profit businesses have experimented with letters and phone scripts for years, but those findings also cover a very narrow set of interactions and are often not public.

The second engineering feature—experimenting and iterating—is also hard to replicate, because measuring whether something “works” in this case is more complex than simply turning on a piece of technology and playing with it. We must first clearly define what outcomes we want from the design, devise a way to measure them, and finally run a test that reliably tells us whether our design is achieving them.

MORE RIGOROUS TESTING OF IDEAS

The problem with HCD and similar approaches to innovation is that they depend too much on intuition. Research has repeatedly shown that our intuitions about human beings are often wrong. Take the commonsensical idea that penalties always help prevent people from engaging in bad behaviors; this notion may have intuitive appeal, but it has proven false. For example, in a study of Israeli day-care centers that sanctioned parents for being late to pick up their children, researchers found that penalties made parents *even more likely* to be late.⁶ This is because they viewed the penalty as a cheap price for the option to be late, versus feeling bound by a social obligation to be timely.

Not only do the social and behavioral sciences give us better starting points, but it also enables us to prototype and test ideas more readily, because we can measure if they are working using impact evaluation methods as well as lab testing procedures from experimental psychology. We can then iterate and improve on the idea until we have a solution ready for implementation.

The behavioral design methodology incorporates HCD’s fundamental approach of being human centered and thoughtful, but adds scientific insights and iterative testing to advance HCD in three significant ways. First, it applies observations about people from experimental academic research. HCD’s reliance solely on self-reported and intuitive insights presents a risk, since so much human behavior is unconscious and not transparent. Also, psychology research shows that people’s self-perception is biased in several ways.⁷ When we do supplement academic insights with qualitative research, we can use behavioral science to make the latter less vulnerable to bias.

For example, we can get more unvarnished answers by asking subjects what their peers typically do rather than what they themselves do. When asked about themselves, subjects may be embarrassed to admit to certain behaviors or may feel compelled to give what they assume the interviewer thinks is the “right” answer.

Second, behavioral design can enhance HCD in the design phase. The behavioral science literature can contribute ideas for solutions based on previously tested interventions. As behavioral design becomes more widely used, more and more data will become available on what designs work and under what conditions. In filtering ideas, we can use behavioral science to anticipate which solutions are likely to suffer from behavioral problems such as low adoption by participants or misperception of choices.

Third, this new approach improves upon HCD by adding more rigorous testing. Many HCD practitioners do test their ideas in prototype with users. While helpful, and part of behavioral design as well, quick user testing cannot tell us whether a solution works. Behavioral design leverages experimental methods to go much further without necessarily adding considerable cost or delay.

Using this approach, we test whether something works—whether it triggers a desired behavioral result—rather than whether the subject thinks something works. We can also test a single component of more complex designs, such as whether a particular piece of information included on a Web page makes a difference, in a lab setting with subjects from our target audience. This is analogous to aeronautical engineers testing wing designs in wind tunnels. By testing and iterating in the field, we do not need to bet on an untested big idea but instead can systematically develop one that we know works. Testing is also what makes it possible, in the design phase, to build on previous successful ideas.

ideas42's work includes many examples of using behavioral design to invent solutions to tough social problems. For example, we recently worked with Arizona State University (ASU) to encourage more eligible students to apply for a special federal work-study program called SEED. In fall 2014, before we started working with ASU, only 11 percent of eligible students were applying for SEED jobs, leaving nearly \$700,000 in financial aid funds unused. ASU wanted our help to increase this proportion.

Diagnosing the problem through a behavioral lens, and interviewing students and staff, we learned that students mistakenly believed that SEED jobs were menial and low-wage. Some thought that a work-study job would interfere with their education rather than complement it. Others intended to apply but missed the deadline or failed even to open the e-mail announcing the program. We designed a series of 12 e-mails to attempt to mitigate all of these barriers. The e-mails dispelled the misperceptions about work-study jobs by stating the correct facts. They made the deadline more salient by reminding students how many dollars of aid they stood to lose. Behavioral research shows that losses loom larger than gains, so the loss framing promised to be more impactful than telling students how much they stood to gain. The e-mails asked students to make a specific plan for when they would complete the work-study job application to reduce the chance that they would forget or procrastinate past the deadline. These behaviorally informed e-mails were compared against a control group of 12 e-mails that contained only basic information about how to apply to the SEED program.

With the redesigned e-mails, which ASU has now adopted, 28 percent more students applied for jobs, and the number of total applications increased by 56 percent. As we were sending 12 e-mails, we used the opportunity to test 12 different subject lines to try to maximize the number of students who opened the e-mail. In five out of the 12 cases, the rate of opening increased by 50 percent or more, relative to a typical subject line. A subject line that increased the open rate from 37 percent to 64 percent made students feel special: “You have something other freshmen don’t.” The control in this case was commonly used language to remind the recipient of impending deadlines: “Apply now! SEED jobs close Thursday.”

THE BEHAVIORAL DESIGN METHODOLOGY

Efforts like this one may sound like nothing more than trial and error, but a systematic and scientific process underlies them that tracks the success of engineering or medicine more closely than HCD. It begins with *defining* a clear problem, *diagnosing* it, *designing* solutions, *testing* and refining the effectiveness of those ideas, and then *scaling* the solutions.⁸ It also starts from a body of knowledge from behavioral science, rather than intuition and guesswork, so that the solutions tried are more likely to succeed.

Let us take a closer look at these steps:

1. Define. The first step is to define the problem carefully to ensure that no assumptions for causes or solutions are implied and that the desired outcome is clear. For example, organizations we serve commonly ask: “How do we help our clients understand the value of our program?” In this formulation, the ultimate outcome is not explicitly defined, and there is an assumption that the best way to secure the outcome is the program (or product) in question. Say the relevant program is a financial education workshop. In this case, we do not know what behaviors the workshop is trying to encourage and whether classroom education is the best solution. We must define the problem *only* in terms of what behaviors we are trying to encourage (or discourage), such as getting people to save more.

2. Diagnose. This intensive phase generates hypotheses for behavioral reasons why the problem may be occurring. To identify potential behavioral hurdles, this approach draws insights from the behavioral science literature and what we know about the particular situation. For example, in the ASU work-study project, we hypothesized that many students intended to apply but failed to follow through because they procrastinated past the deadline or simply forgot it. Both are common behavioral underpinnings for such an intention-action gap.

After generating some initial hypotheses, the next step is to conduct qualitative research and data analysis to probe which behavioral barriers may be most prevalent and what features of the context may be triggering them. Here, “context” refers to any element of the physical environment, and any and all experiences that the consumer or program’s beneficiary is undergoing, even her physical or mental state in the moment.

Qualitative research usually includes observation, mystery shopping (purchasing a product or experiencing a program incognito to study it firsthand), and in-depth interviews. Unlike typical qualitative research that asks many “why” questions, the behavioral approach focuses on “how” questions, since people’s post-hoc perceptions of why they did something are likely to be inaccurate.

3. Design. Having filtered down and prioritized the list of possible behavioral barriers via the diagnosis phase, we can generate ideas for solutions. Here many of the structured creativity techniques of HCD prove useful. When possible, it is best to test a few ideas rather than to guess which solution seems best. Solutions also change during their journey from the whiteboard to the field, as numerous operational, financial, legal, and other constraints invariably crop up. Such adaptations are critical to making them scalable.

4. Test. We can then test our ideas using RCTs, in which we compare outcomes for a randomly selected treatment group vis-à-vis those for a control group that receives no treatment or the usual treatment. Although RCTs in academic research are often ambitious, multiyear undertakings, we can run much shorter trials to secure results. An RCT run for academic purposes may need to measure several long-term and indirect outcomes from a treatment. Such measurement typically requires extensive surveys that add time and cost. For iterating on a design, by contrast, we may only measure proximate indicators for the outcomes we are seeking. These are usually available from administrative data (such as response to an e-mail campaign), so we can measure them within days or weeks rather than years. We measure long-term outcomes as a final check only after we have settled on a final solution.

When RCTs are impossible to run even for early indicators, solutions can be tested that approximate experimental designs. A more detailed description of these other methods is outside the scope of this article but is available through the academic literature on program evaluation and experimental design.

If the solution is complex, we first test a crude prototype with a small sample of users to refine the design.⁹ We can also test components of the design in a lab first, in the way that engineers test wing designs in a wind tunnel. For example, if we are designing a new product and want to refine how we communicate features to potential users, we can test different versions in a lab to measure which one is easiest to understand.

5. Scale. Strictly speaking, innovation could end at testing. However, scaling is often not straightforward, so it is included in the methodology. This step also has parallels with engineering physical products, in that designing how affordably to manufacture a working prototype is, in itself, an invention challenge. Sometimes engineers must design entirely new machines just for large-scale manufacturing.

Scaling could first involve lowering the cost of delivering the solution without compromising its quality. On the surface, this step would be a matter of process optimization and technology, but as behavioral solutions are highly dependent on the details of delivery, we must design such optimization with a knowledge of behavioral principles. For example, some solutions rely on building a trusted relationship between frontline staff and customers, so we would not be able to achieve a cost reduction by digitizing that interface. The second part of scaling is encouraging adoption of an idea among providers and individuals, which itself could benefit from a scientific, experimental process of innovation.

A CLOSER LOOK AT THE METHODOLOGY

To be fair, it is sometimes impossible to go through the full, in-depth behavioral design process. But even in these cases, an abridged version drawing on scientific insights rather than creativity alone is

always feasible. Notice that the define, diagnose, and design stages of the behavioral design process apply the scientific method in two ways: They draw on insights from the scientific literature to develop hypotheses, and they collect data to refine those hypotheses as much as possible. The first of these steps can be accomplished even in a few hours by a behavioral designer with sufficient expertise. The second component of data collection and analysis takes more time but can be shortened while still preserving a scientific foundation for the diagnosis and design. Field testing with a large sample can be the most time-consuming, but lab tests can be completed within days if time is constrained.

Two sorts of hurdles typically confront the full behavioral design process: lack of time and difficulty measuring outcomes. In our experience, time constraints are rarely generated by the problem being addressed. More often, they have to do with the challenges of complex organizations, such as budget cycles, limited windows to make changes to programs or policies, or impatience among the leadership. If organizations begin to allocate budgets for innovation, these artificial time constraints will disappear.

To better understand working under a time constraint, consider ideas42's work with South Africa's Western Cape to reduce road deaths during the region's alcohol-fueled annual holiday period. The provincial government had a small budget left in the current year for a marketing campaign and only a few weeks until the holiday season began. The ideas42 team had to design a simple solution fast; there was no time to set up an RCT with a region-wide marketing campaign. The team instead used an abridged version of the first three stages to design a solution grounded in behavioral science. Quick diagnosis revealed that people were not thinking about safe driving any more than usual during the holidays, despite the higher risk from drunk driving. To make safe driving more salient, ideas42 designed a lottery in which car owners were automatically registered to win but would lose their chance if they were caught for any traffic violations. That design used two behavioral principles coming out of Prospect Theory,¹⁰ which tells us that people tend to overestimate small probabilities when they have something to gain, and that losses feel about twice as bad as the equivalent gain feels good.

Applying the first principle, we used a lottery, a small chance of winning big, rather than a small incentive given to everyone. Using the second, we gave people a lottery ticket and then threatened to take it away. Since an RCT was not feasible, we measured results by comparing road fatalities in the treatment period with road fatalities in the same month of the previous year; this showed a 40 percent reduction in road fatalities. There were no known changes in enforcement or any other policies. While ideas42 was not able to continue to collect data in subsequent years, because its contract ended, the program saw success in subsequent years as well, according to our contacts in government.

ADOPTING BEHAVIORAL DESIGN

If you were convinced of behavioral design's value and wanted to take the leap, how would you do it? There are resources available, and many more are still in the works. Behavioral insights are not yet readily available in one place for practitioners to access, but are instead spread out over a vast literature spanning many

academic disciplines, including psychology, economics, neuroscience, marketing, political science, and law. Results from applications of behavioral science are even more distributed because many are self-published by institutions such as think tanks, impact evaluation firms, and innovation consultancies.

To mitigate this problem, ideas42, in partnership with major universities and institutions that practice behavioral design in some form, is building an easily searchable Web-based resource as well as a blog that will make it possible to find ready-to-use behavioral insights in one place. In the meantime, some of these organizations, including ideas42, also offer classes that teach elements of behavioral design as well as some key insights from behavioral science that practitioners would need in order to do behavioral design. As the practice of behavioral design is adopted more widely, and its use generates more insights, it will become more powerful. Like technology, it will be able to continue to build on previous discoveries.

Organizations and funders would also do well to adopt the behavioral design approach in their thinking more generally. Whenever someone proposes a new approach for innovation, people scour the methodology for the secret sauce that will transform them into creative geniuses. In this case, the methodology applications of behavioral science, in themselves, do have a lot to offer. But even more potential lies in changing organizational cultures and funding models to support a scientific, evidence-based approach to designing interventions. Here are three suggestions about how organizations can adopt behavior design:

Fund a process (and people good at it), not ideas. | Today's model for funding innovation typically begins with a solution, not a problem. Funders look to finance the testing or scaling up of a new big idea, which by definition means there is no room for scientifically analyzing the problem and then, after testing, developing a solution. Funders should reject this approach and instead begin with the problem and finance a process, and people they deem competent, to crack that problem scientifically. To follow this path, funders must also become comfortable with larger investments in innovation. The behavioral design approach costs a lot more than whiteboards, sticky notes, and flip charts—the typical HCD tools—but the investment is worth it.

Embrace failure. | In a world where ideas are judged on expert opinion and outcomes are not carefully measured, solutions have no way of failing once they leave the sticky-note phase and get implemented. In a new world where ideas must demonstrably work to be successful, failure is built into the process, and the lessons learned from these failures are critical to that process. In fact, the failure rate can serve as a measure of the innovation team's competence and their bona fide progress. To be really innovative, a certain amount of risk and courting failure is necessary. Adopting a process that includes failures can be hard to accept for many organizations, and for the managers within those organizations who do not want their careers to stall; but as in engineering and science, this is the only way to advance.

Rethink competitions. | The first XPRIZE for building a reusable spacecraft rekindled the excitement for competitions, which

have now become common even outside the technology industry. However, competitions to invent new technology are fundamentally different: With a spacecraft, it is relatively easy to pick the winner by test-flying each entry. In the social sector, by contrast, competitions have judging panels that decide which idea wins. This represents a big-idea approach that fails to motivate people to generate and test ideas until they find one that demonstrably works well, rather than one that impresses judges. Staged competitions could work much better by following a behavioral-design approach. The first round could focus on identifying, or even putting together, the teams with the best mix of experience and knowledge in behavioral design and in the domain of the competition. Subsequent rounds could fund a few teams to develop their ideas iteratively. The teams whose solutions achieved some threshold of impact in a field test would win. Innovation charity Nesta's Challenge Prize Centre has been using a similar approach successfully, as has the Robin Hood Foundation, with the help of ideas42.

Revolutionizing how we innovate presents a huge opportunity for improving existing programs, products, and policies. There is already sufficient scientific research and techniques to begin making the change, and we are learning more about how to better devise things for human interactions every day. The more we use a scientific approach to innovate, and construct platforms to capture findings, the more science we will have to build on. This immense promise of progress depends on changing organizational cultures and funding models. Funders can and must start to bet not on the right "big ideas" but on the right process for solving challenges and on the people who are experts in that process. They must also not just expect failures, but embrace them as the tried and true means for achieving innovation. ■

NOTES

- 1 Amos Tversky and Eldar Shafir, "Choice Under Conflict: The Dynamics of Deferred Decision," *Psychological Science*, vol. 3, no. 6, 1992.
- 2 Sheena S. Iyengar and Mark R. Lepper, "When Choice Is Demotivating: Can One Desire Too Much of a Good Thing?" *Journal of Personality and Social Psychology*, vol. 79, no. 6, 2000.
- 3 For several examples, see ideas42's June 2016 report *Nudging for Success: Using behavioral science to improve the postsecondary student journey*, <http://www.ideas42.org>.
- 4 D.D. Baals and W.R. Corliss, *Wind Tunnels of NASA*, Washington D.C.: Scientific and Technical Information Branch, National Aeronautics and Space Administration, 1981.
- 5 Kevin Ashton, *How to Fly a Horse: The Secret History of Creation, Invention, and Discovery*, New York: Anchor, 2015.
- 6 Uri Gneezy and Aldo Rustichini, "A Fine Is a Price," *Journal of Legal Studies*, vol. 29, no. 1, 2000.
- 7 E. Pronin, T. Gilovich, and L. Ross, "Objectivity in the Eye of the Beholder: Divergent Perceptions of Bias in Self Versus Others," *Psychological Review*, vol. 111, no. 3, 2004.
- 8 This process was previously described briefly by ideas42 Managing Director Saugato Datta and cofounder Sendhil Mullainathan in their paper "Behavioral Design: A New Approach to Development Policy," *Review of Income and Wealth*, vol. 60, no. 1, 2014.
- 9 For good examples of quick tests, see David A. Asch and Roy Rosin, "Innovation as Discipline, Not Fad," *New England Journal of Medicine*, vol. 373, no. 7, 2015.
- 10 Daniel Kahneman and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, vol. 47, no. 2, 1979.