

Jan Kallwejt

In 1993, after five years of grad school and low-wage postdoctoral research, Michael Kremer got a job as a professor of economics at MIT. With his new salary, he finally had enough money to fund a long-held desire: to return to Kenya's Western Province, where he had lived for a year after college, teaching in a rural farming community. He wanted to see the place again, reconnect with his host family and other friends he'd made there.

When he arrived the next summer, he found out that one of those friends had begun working for an education nonprofit called ICS Africa. At the time, there was a campaign, spearheaded by the World Bank, to provide free textbooks throughout sub-Saharan Africa, on the assumption that this would boost test scores and keep children in school longer. ICS had tasked Kremer's friend with identifying target schools for such a giveaway.

While chatting with his friend about this, Kremer began to wonder: How did ICS know the campaign would work? It made sense in theory—free textbooks should mean more kids read them, so more kids learn from them—but they had no evidence to back that up. On the spot, Kremer suggested a rigorous way to evaluate the program: Identify twice the number of qualifying schools as it had the money to support. Then randomly pick half of those schools to receive the textbooks, while the rest got none. By comparing outcomes between the two cohorts, they could gauge whether the textbooks were making a difference.

What Kremer was suggesting is a scientific technique that has long been considered the gold standard in medical research: the randomized controlled trial. At the time, though, such trials were used almost exclusively in medicine—and were conducted by

large, well-funded institutions with the necessary infrastructure and staff to manage such an operation. A randomized controlled trial was certainly not the domain of a recent PhD, partnering with a tiny NGO, out in the chaos of the developing world.

But soon after Kremer returned to the US, he was startled to get a call from his friend. ICS was interested in pursuing his idea. Sensing a rare research opportunity, Kremer flew back to Kenya and set to work. By any measure it was a quixotic project. The farmers of western Kenya lived in poverty, exposed to drought, flood, famine, and disease. Lack of paved roads hampered travel; lack of phones impeded communication; lack of government records stymied data collection; lack of literate workers slowed student testing. For that matter, a lack of funds limited the scope. It was hardly an ideal laboratory for a multiyear controlled trial, and not exactly a prudent undertaking for a young professor with a publishing track record to build.

The study wound up taking four years, but eventually Kremer had a result: The free textbooks didn't work. Standardized tests given to all students in the study showed no evidence of improvement on average. The disappointing conclusion launched ICS and Kremer on a quest to discover why the giveaway wasn't helping students learn, and what programs might be a better investment.

As Kremer was realizing, the campaign for free textbooks was just one of countless development initiatives that spend money in a near-total absence of real-world data. Over the past 50 years, developed countries have spent something like \$6.5 trillion on assistance to the developing world, most of those outlays guided by little more than macroeconomic theories, anecdotal evidence, and good intentions. But if it were possible to measure the effects of initiatives, governments and nonprofits could determine which programs actually made the biggest difference. Kremer began collaborating with other economists and NGOs in Kenya and India to test more strategies for bolstering health and education.

At home, meanwhile, his work was helping to inspire a small movement of economists and other social scientists—playfully dubbed the “randomistas,” in reference to the randomized nature of the studies. In 2003, a few years after Kremer had moved across town to Harvard, three like-minded economists at MIT launched a research institution, now called J-PAL (the full moniker is the Abdul Latif Jameel Poverty Action Lab, named for the late father of a donor), to promote the use of randomized controlled

trials on questions of poverty and development. They work closely with an independent sister NGO, Innovations for Poverty Action (IPA), which implements evaluations in the field. Kremer joined both groups as an affiliated researcher.

In the decade since their founding, J-PAL and IPA have helped 150 researchers conduct more than 425 randomized controlled trials in 55 countries, testing hypotheses on subjects ranging from education to agriculture, microfinance to malaria prevention, with new uses cropping up every year (see “Randomize Everything,” below). Economists trained on randomized controlled trials now work in the faculties of top programs, and some universities have set up their own centers to support their growing rosters of experiments in the social sciences.

TO FIND OUT WHAT WORKS ON THE GROUND, YOU NEED TO CLIMB DOWN FROM THE IVORY TOWER AND DO SOME SERIOUS LEGWORK IN THE PLACES YOU’RE TRYING TO HELP.

Their results have challenged—or, in some cases, confirmed with hard data—widely held beliefs about aid strategies that command billions of dollars in annual outlay. It turns out that retrospective analysis of a program’s impact or even suggestive case studies from a few targeted households can be worse than useless in understanding how a program actually affects a community in the real world.

J-PAL researchers will be the first to caution that each study is specific to its context. What works in one community may not in another. But in the realm of human behavior, just as in the realm of medicine, there’s no better way to gain insight than to compare the effect of an intervention to the effect of doing nothing at all. That is: You need a randomized controlled trial. And that means you need to climb down from the ivory tower and do some serious legwork in the places you’re trying to help.

The first thing you need to know about randomized controlled trials, especially those pertaining to economics and human behavior, is that they’re hard—very hard. To evaluate the textbook campaign, ICS Africa had to collaborate with the Kenyan education ministry to choose 100 schools in the rural Western Province, ensure the textbooks got to the assigned schools, and develop and administer tests to thousands

of students whom ICS then tracked for the next four years. And that was simple compared with some of the other developing-world trials that IPA researchers have gone on to construct. For example, when two researchers were trying to study the effects of different pricing models for getting people to use bed nets to protect against malaria, the local nonprofit that sold subsidized nets declined to help them with the research or even sell them nets at the discounted rate. So the team had to spend a year recruiting a local research staff and drumming up funds to buy thousands of bed nets. Then they had to win the cooperation of 20 different prenatal clinics in Kenya, and then they had to oversee the experiment as it tracked the behavior of 10,000 pregnant women.

But there's a beautiful utility that can emerge from accumulating data, as interventions that researchers expect to be marginal—or, in some cases, weren't even thought of when the study began—reveal themselves to be highly effective. For example, after ICS's textbook campaign, it worked with Kremer to test a host of other strategies for increasing school participation rates in western Kenya, from subsidized meals to free uniforms to merit scholarships. One of the most cost-effective ways to boost attendance came as a big surprise: treatment for intestinal worms, which caused absenteeism to drop by one-quarter. And it wasn't only the schools receiving treatment that benefited. Attendance also rose at nearby schools as the overall transmission rate in the region dropped. The researchers calculated that, on average, deworming "buys" one extra year of school attendance for just \$3.50, less expensive than any other intervention tested. This unexpected finding has led researchers to found an initiative called Deworm the World, which has worked in partnership with governments and NGOs to treat 37 million children.

Similarly, a 2004 experiment to promote water treatment wound up suggesting a solution that the researchers hadn't imagined. In this case, the larger goal was to combat diarrheal diseases, which kill millions of people every year, especially children under 5. Chlorine treatments can render water safe; but despite years of education efforts in Kenya, few people purchased and used the chlorine solution, even though it was widely available.

To test remedies, the researchers identified 88 springs that supplied nearly 2,000 households in western Kenya. Surveys of local women, who usually collect the water for

the family and monitor children's health, found that 70 to 90 percent knew about the chlorine product but only 5 percent used it, and IPA's in-home tests detected chlorine in the water of just 2 percent of households. These women knew how to make their water safe, but they weren't doing it.

After that, the researchers spent four years testing different interventions. Giving away the chlorine solution helped in the near term, but when the free supply ran out, usage fell off. Half-off coupons for chlorine were a bust; out of 2,724 coupons handed out, just 10 percent were ever redeemed. The study also tested whether local "promoters," sent door to door with one free voucher per family, might succeed in evangelizing the use of chlorine among their neighbors. Promoters did make a difference in the short term; in this cohort, 40 percent of household water samples showed evidence of chlorine. But that number fell significantly when the vouchers ran out.

Finally, the researchers tried a new strategy, one that they had come to late in their design. Collaborating with engineers at MIT and Tufts and local artisans in Kenya, they developed chlorine dispensers that could be installed at the spring. Each tank was large enough for a whole community; two turns of a knob would release the right amount of chlorine to treat the standard 20-liter jerrican that women typically carry water in. In communities with these dispensers, the percentage of households with detectable chlorine in their water jumped to 60 percent, versus 58 percent in a group where seven free bottles were actually delivered to their homes. And the dispensers cost less than a third as much.

Finally, the researchers tried a new strategy, one that they had come to late in their design. Collaborating with engineers at MIT and Tufts and local artisans in Kenya, they developed chlorine dispensers that could be installed at the spring. Each tank was large enough for a whole community; two turns of a knob would release the right amount of chlorine to treat the standard 20-liter jerrican that women typically carry water in. In communities with these dispensers, the percentage of households with detectable chlorine in their water jumped to 60 percent, versus 58 percent in a group where seven free bottles were actually delivered to their homes. And the dispensers cost less than a third as much.

The researchers developed the dispenser with one big insight about human behavior in mind: We're far more likely to develop a good habit when we have cues to remind us

what we should be doing, and when no inconveniences stand in the way. People saw the dispenser right by the spring, they saw their neighbors using it, and it was easy to use it themselves. But before the campaign, there would have been no way to justify why community dispensers should have been pursued over strategies suggested by other theories (that people would purchase chlorine once they had become used to using it, that a nonzero price would confer a sense of value, that social pressure could overcome cost barriers). In the absence of a controlled trial, our theories are just more noise.

The randomista movement is here to stay, bringing data and evidence into conversations that for too long have been dominated by poignant anecdotes and academic speculation. On a stormy September afternoon, I visit an empty New Hampshire summer camp to watch as J-PAL and IPA train the next generation of researchers. Many of the 68 attendees hold recent undergraduate or graduate degrees in economics from top-tier universities. Heidi McAnnally-Linz, the global outreach coordinator for IPA, passes her seven-week-old daughter to her husband for safekeeping and takes her place at the front of the room. Projected on a large screen beside her is a slide. On the left is a single orange and a price: \$1. On the right is a bag of oranges and the same price: \$1. At the bottom of the slide is a question: Which would you buy?

The trainees chuckle. There's no need for an economics degree to answer that.

"We use this to explain our work to funders," Linz says, and advances to the next slide. On the left, a photo of smiling schoolchildren in Africa, labeled "Free School Uniforms," shows a price of \$5,000. On the right, a picture of a teacher handing a pill to a student, labeled "Deworming Children," also shows a price of \$5,000. Below is the same question: Which would you buy? Answering that question took nearly a decade of research involving tens of thousands of Kenyan students—but the value of the answer is incalculably high.

The young staffers come from different backgrounds, but they're all hoping to use their economics training for good. Madeline Duhon has just defected from the private sector, where the explicit goal of her department was to raise the company's stock price. "That wasn't a huge motivator for me," she says. Andreas Tiemann recently accepted a position at IPA after working as a research analyst for J-PAL in Chile. "It's definitely not a 9-to-5 job," he says, pointing out that one research team he knows is trying to study nomads in Mongolia.

Tiemann is here to teach trainees how to design surveys so that the questions elicit good data. He's also trying to drive home just how seriously J-PAL and IPA take quality control, with reentry of survey results in duplicate—and sometimes triplicate—to ensure the accuracy of the data. The trainees are also learning how to identify communities for an evaluation, how to randomize, how to protect participants' privacy. They brush up on their database and statistical skills and think about how to work with other researchers and local staff. When the training is over, some will depart for two-year assignments in other countries; others will stay to work at the J-PAL or IPA offices in the US. Many are headed for higher degrees and research projects of their own.

J-PAL and IPA aren't just carrying out trials; they're also coordinating—and championing—the use of randomized controlled trials throughout the social sciences. As part of that effort, they're building a master registry, overseen by the American Economic Association, to provide a transparent record of trials across disciplines. By showing all the ways that rigorous studies can be carried out with modest budgets and small staffs, these organizations have allowed a tool once confined to medical trials to help investigate just about anything.

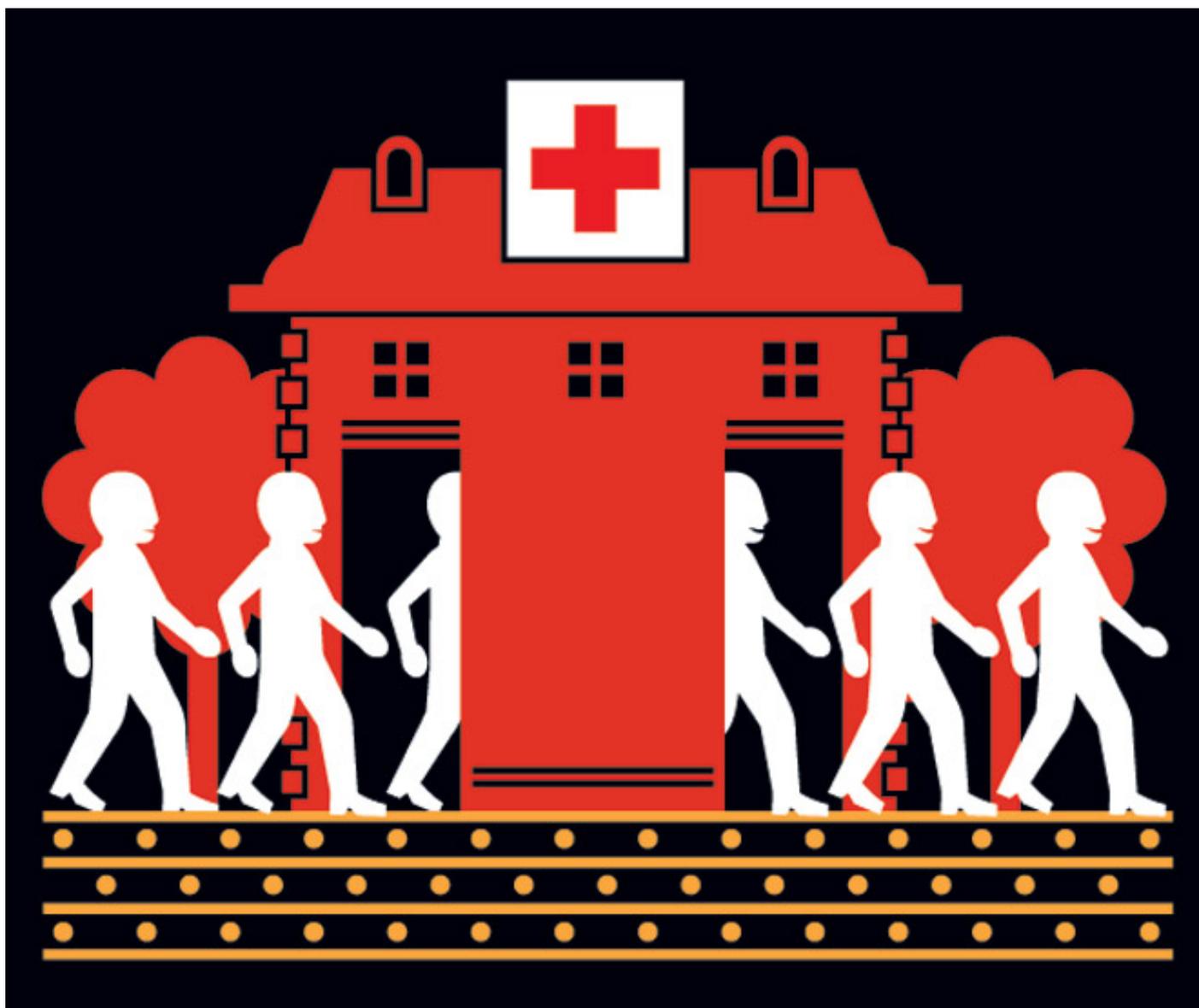
One recent trial in China gauged the effect on productivity, work-life balance, and employee turnover of allowing call-center staff to work from home. Another evaluated a policy of publicly releasing audits and revealing corruption of elected officials in Brazil, to see how it affected the reelection of incumbent mayors. Several are currently testing rainfall and crop price insurance programs for farmers in India, Malawi, and Ghana. Another is evaluating whether paying local residents to refrain from cutting trees can curb deforestation in Uganda. And the breadth of questions being investigated is expanding quickly.

Indeed, for some of the newest experiments in the database, researchers will be able to leave their passports at home. J-PAL has recently launched a North American wing, and it's working with groups such as the Lab for Economic Opportunities at Notre Dame and the Crime Lab and the Urban Education Lab at the University of Chicago. With 15 percent of Americans living below the poverty line, a growing income gap, and the highest incarceration rate in the world, the US has its own dark corners that could use some statistical illumination.

Yes, randomized testing can be expensive and time-consuming. But by surfacing the small details that can determine a campaign's success or failure, it is potentially saving billions of dollars and years of effort. And it's providing something invaluable, something beyond intuition and anecdote: hard statistical proof on the most effective ways to change the world.

RANDOMIZE EVERYTHING

Controlled trials are shedding light on human behavior—and how to nudge it in a positive direction. Here are some choice examples. —J.B.

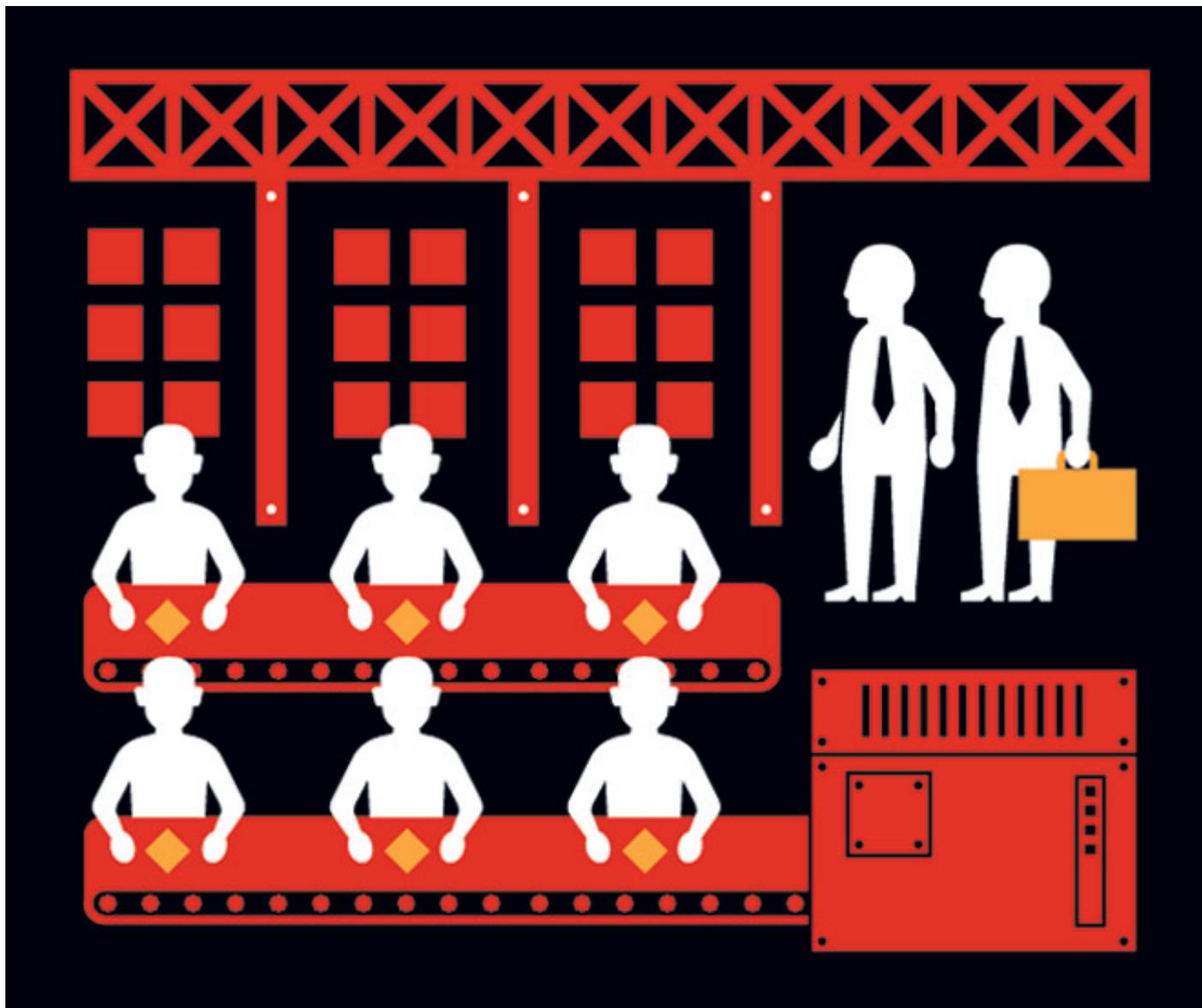


Incentives for Vaccination

India, 2004–2007

In rural Rajasthan, 134 villages were randomly assigned to one of three groups: a control, a group that received monthly visits from a vaccine team, and a group where the visits came with a free bag of lentils, plus a \$2 set of plates as a prize for families that completed the full course of shots. Incentivized villages vaccinated at twice the

rate of those that got only the monthly visits.

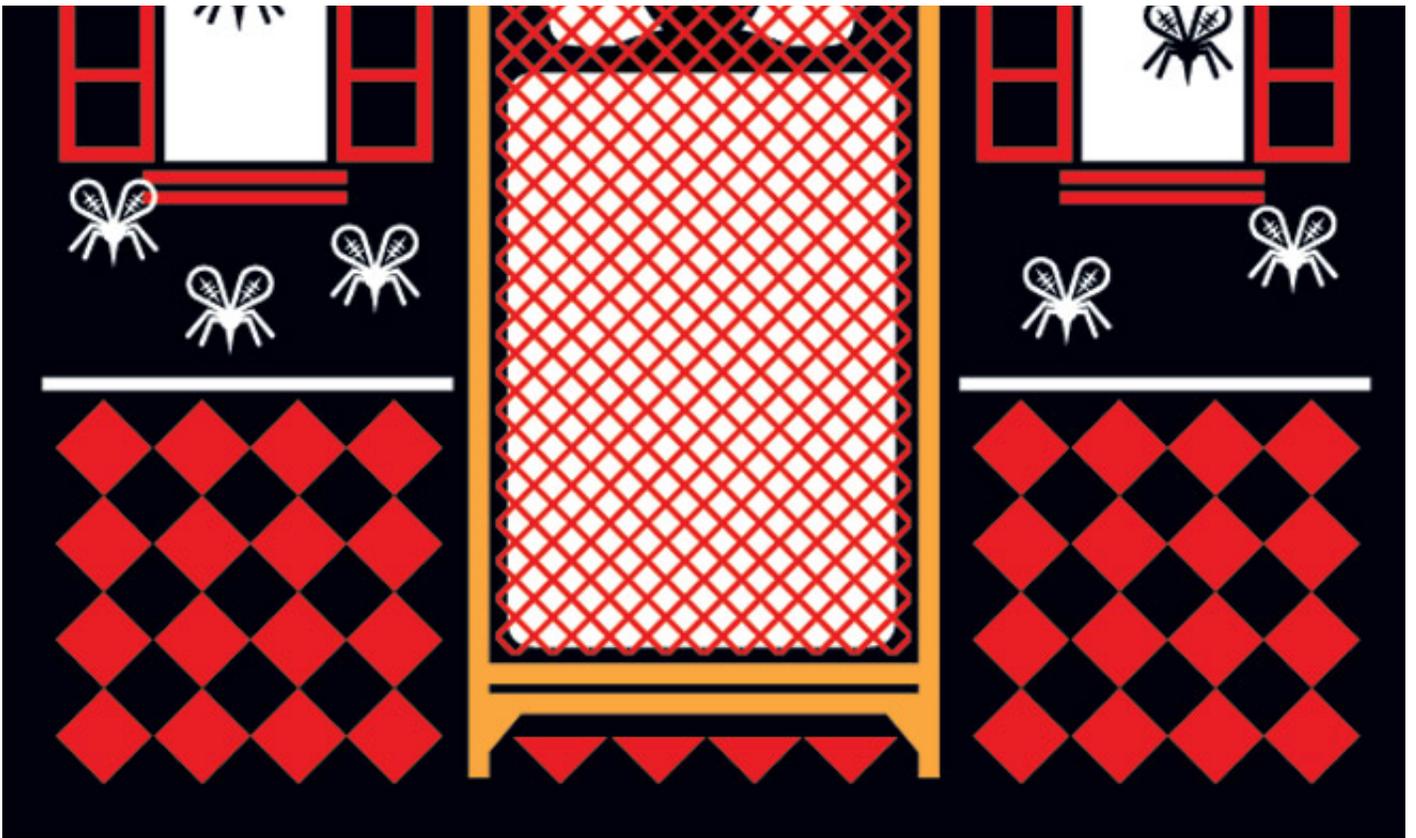


Management Consulting

India, 2008–2011

Manufacturing businesses in India are 40 percent less efficient than Western firms. So researchers identified 17 large manufacturers to get evaluated by consultants. They then assigned the firms to a control group or a treatment group, which received targeted advice to address their weaknesses. Productivity in the treatment group increased by 17 percent in the first year.

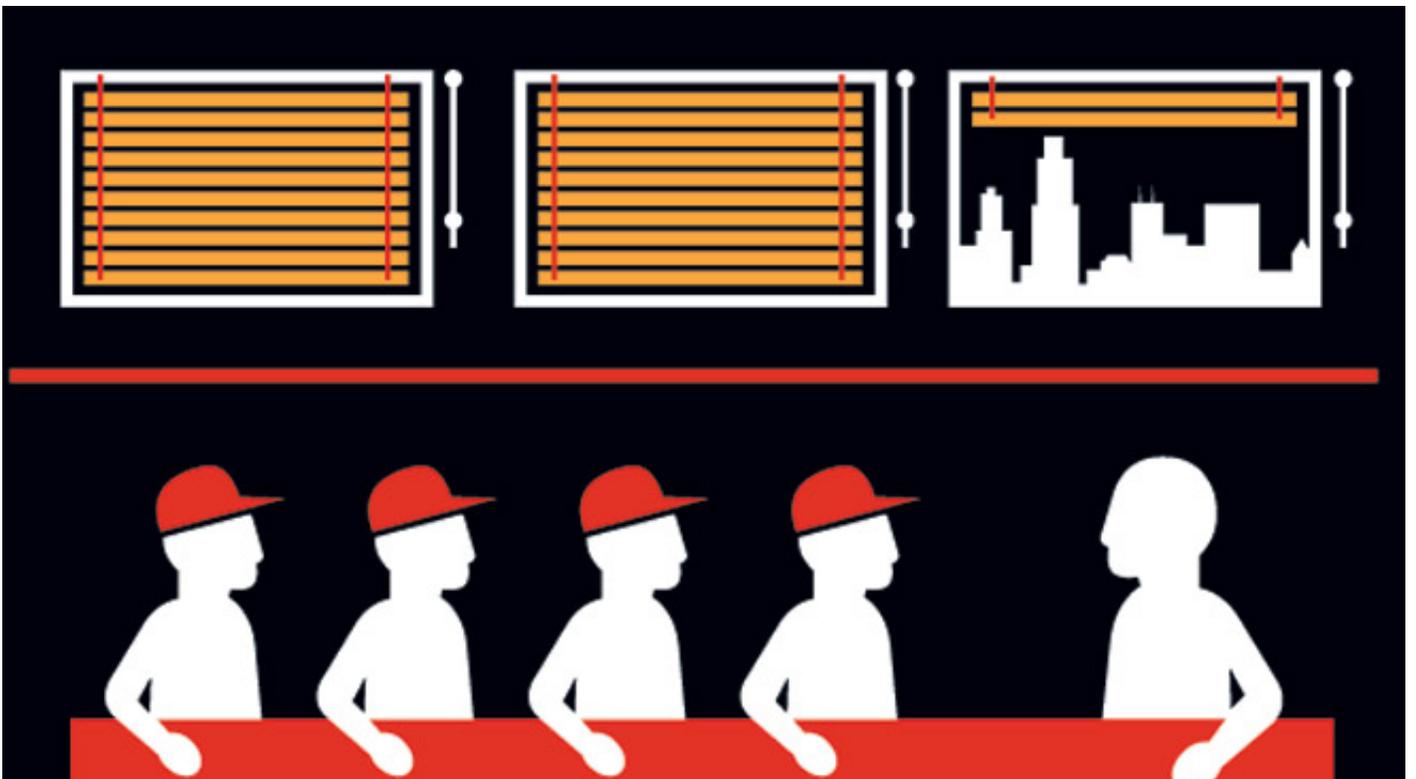




Bed Net Pricing

Kenya, 2006

Some economists have argued that giving stuff away cheapens it in recipients' minds. For bed nets, which help prevent malaria, researchers wanted to know what price would maximize use. So they randomly assigned 20 clinics to one of five groups; one was a control, while the others offered nets at various prices — free, \$0.25, \$0.30, and \$0.60. Their finding: Free won hands-down.





Mentoring High-Risk Boys

Chicago, 2009–2010

Researchers identified 2,740 boys in high-crime neighborhoods and assigned them to control or treatment groups. The treatment program, called “Becoming a Man,” aimed to help the youths make better decisions and develop conflict-resolution skills. During the study year, program participants had better school attendance and 44 percent fewer violent-crime arrests.



Rural Bus Service

Malawi, 2009

The lack of transportation in rural Africa hinders economic growth. In hopes of finding a business model for private service, researchers hired buses and assigned 406 households to either a control group or one of seven other groups offering passes at different prices. The results were eye-opening: No price created a sustainable business, meaning heavy subsidies would be required.

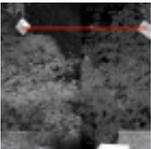
Related
You Might Like
Related Links by Contextly



[Solving Parrots Understand Cause and Effect](#)



[Trees Multitask — When the Stakes Are High](#)



[Tet-Block Problem With a Twist](#)



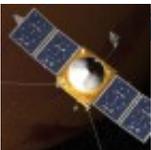
[Quartz Crystals' Could Upend Physicists' Theory of Time](#)



[Space Agencies' Curiosity: Who Will Suffer the Space Budget Axe?](#)



[A New Kind of Food Science: How IBM Is Using Big Data to Invent Creative Recipes](#)



[Mars Reconnaissance Orbiter: NASA Launches New Mission to Mars](#)



[Spectacular Video of Etna's Latest Eruption](#)