Stanford SOCIAL INNOVATION Review

Informing and inspiring leaders of social change

Scaling

Getting to Scale: Size Matters. Shape Matters More.

Scale is a verb, not a noun: The trajectory and curve of impact are more important than the numbers.

By Kevin Starr | Jul. 20, 2021

(Photo by iStock/bacalao64)

worth of social impact.

Scale is one of those words that means something different to just about everyone. But to measure it, to figure out concretely whether an intervention has scaled, you have to decide what you mean. So when Michael Kremer et al set out to measure the social return from innovation investments made by USAID's Development Innovation Ventures (DIV), they selected four interventions that had "scaled" because they had reached at least a million people and used that benchmark to estimate the social return on DIV's investment (see this excellent blog post by Jocilyn Estes, David Evans and Sarah Rose summarizing that work). As they note, the choice of a million was "an arbitrary cut-off motivated by the costliness of detailed data collection." And on that basis, a dollar spent by DIV is judged to generate $17

Whether $17 is the right figure or not, the research makes it clear that the return on investment is a lot. And it's great to have a persuasive piece of evidence around the value of smart innovation investment, a powerful resource to make the case to others. (My own dream is that

USAID will reform its current approach of hopscotch projects of dubious impact, hugely increase DIV's budget, and restructure around scaling solutions like ones DIV has gestated along with many others that Michael Kremer and colleagues have proven out.)

While there is zero chance my little fantasy is going to budge USAID, the study does provide a nice vehicle to examine this critical notion of "scale." It's great that the authors clearly define what they mean: Something has scaled when it's reached a million people. But in a world of seven billion people—and a lot of big problems—a million might be just the start. We need to think about scale in terms of both the size of the problem and the potential of the solution. With this in mind, what matters is the shape of the impact curve and the prospect that the slope of additional impact over time will get steeper until the solution approaches its full potential.

In other words, scale isn't a number, it's a curve: To "scale" is to achieve sustained non-linear growth of impact and an increasing momentum that persuades us that something really big—the achievement of its full potential—may well happen.
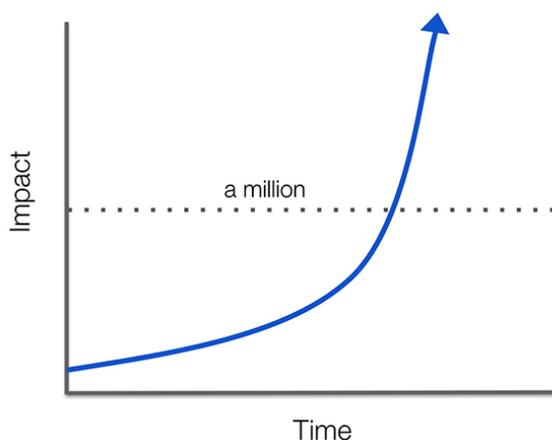

Calling scale "a million" doesn't even *begin* to do all that. It fails to anticipate the impact S-curve—impact that takes off, then flattens out—that afflicts so many interventions. For example, take chlorine dispensers that have been cited as a big win, simple devices placed at community water sources that make it easy to squirt a dose of purifying chlorine into your filled water container. Well-designed RCTs in rural communities in Kenya, Malawi, and Uganda have demonstrated impressive impact, for an intervention that is effective, cheap, and easy for users. This is all great, as far as it goes.

Mulago supported the work. We thought it had huge potential. It seemed like it would work wherever there is an unprotected water source, implying tens, maybe hundreds of millions of people. The problem was that there was no real strategy to take it to its potential. In the wake of these high-profile studies, and with the impressive efforts of the NGO Evidence Action, the reach of chlorine dispensers grew rapidly, then—as far as I can tell—peaked at about four million users worldwide. Nobody emerged to do it or pay for it at a scale that approached its potential: There was no plausible way to get to a profitable business model and governments showed little interest in doing it *or* paying for it. Some NGOs took it up, but they rarely

scale anything up in a concerted way, and even if they wanted to, there was no identified payer willing to fund the effort at a size that mattered. As a result, dispenser implementation became a patchwork of NGO projects paying for it with whatever dough they could get their hands on. It was a sign of desperation when implementers raised money through a shaky carbon credit deal that paid out on the basis of "avoided boiling." The idea was that the dispensers would obviate the need to purify water by boiling: the problem was that these populations had never boiled water before and there was no conceivable reason to believe they'd ever start.

Now, don't get me wrong, four million families protected from waterborne disease is a lot, and the dispenser effort was totally worth the effort. If we're talking about scale, though, we have to note that dispenser use flattened out into an impact S-curve that fell orders of magnitude short of what we believed to be its potential. A million was great when there was momentum, but when that momentum stalls far short of the potential—to say nothing of the need—it's not so much. And now, chlorine dispensers have become a kind of zombie intervention: not inanimate yet, but not particularly going anywhere. The innovation will remain animate as long as there is someone willing to scramble for grants to keep it going, but if it was supposed to make a big and lasting dent in a big problem, it failed.
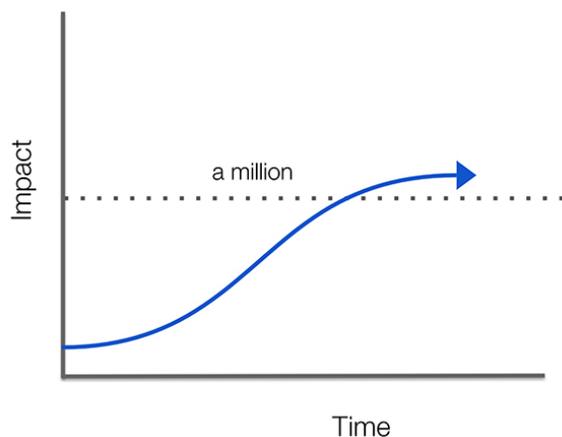
Another example is the discussion of "digital attendance monitoring" for health workers in Indian clinics, an effort to get health facility staff to show up for work by using a thumbprint reader to document their attendance. The intervention was said to have "reached" 1.8 million people. But really, it didn't "scale" at all: 1.8 million only describes the size of the population included in the pilot study! And while it was *initially rolled out* at a scale over one million, it never got replicated – I'm not sure this even qualified as an S-curve. Worse, it was discontinued in the pilot. As portrayed in this paper, it represents an abandoned pilot. An abandoned pilot, however successful, is not an example of successful scaling, and "a million reached" is irrelevant here. Abandoning promising pilots has long been one of the worst habits of the development sector all over the world, and it's a real problem that we have RCTs all over the place that are not backed by an organization committed to scaling up positive results.

Using "a million" to indicate scale also neglects lasting impact. For example, take another of the key investments cited, "Road Safety Stickers," which are placed in public minibuses in Kenya to encourage passengers to speak up against reckless driving. In Kenyan

passenger vans—called matatus—drivers are incentivized to drive like maniacs, so the prominently-placed stickers remind passengers that they can and should tell the driver to slow the hell down. It's a super cheap, relatively easy intervention and the ROI is impressive. So far, so good. But does the intervention create *lasting* impact? There is the problem of habituation: we tend to ignore ubiquitous warning labels over time, perhaps especially when they get dirty and otherwise messed up. And you have to have a system that makes sure they're reliably in place. And, of course, nothing has changed about the incentives that lead to maniacal driving in the first place. The intervention failed to replicate in Tanzanian and Ugandan RCTs, and the program has been suspended in Kenya, so I might be right. Looks like another impact S-curve.

One exception cited in the paper is affordable eyeglasses, as taken on by VisionSpring: Here, a scale-obsessed, capable organization is providing an example of a sustained, steepening impact curve. Eyeglasses have a jaw-dropping ROI, because we're talking about recovered livelihoods, helping someone resume something they *already* know how to do. But however useful glasses are to people, you still have to actively and effectively sell them at the "last meter." That means that if you want to scale affordable eyeglasses, you have to recruit existing distribution channels where you can be confident that glasses will be marketed and sold to the people who will benefit most.

That's exactly what VisionSpring is doing, and that's what could keep impact on a steepening curve. Their biggest hit so far has been distribution through BRAC in Bangladesh, where community health workers sell their glasses as part of a basket of health goods. We love VisionSpring and have funded them for years. It can be hard to figure out the best way to scale an innovation, but we've happily supported their efforts to figure it out—dead-ends and all—because they've done the necessary work in a smart way, the kind of the dogged effort it takes to create and maintain a steepening impact trajectory. We'll have the prospect of getting eyeglasses to problem-solving scale when more organizations start plugging them into the right distribution channels, and even more so when the right distributors do it on their own.

If innovation matters, then high-impact innovations have to scale up enough to make a big dent in the problem. As weird as it is to say it, a million may not matter so much if the impact curve flattens long before the potential is achieved. With eight billion people on the planet there is a need to take scalable solutions all the way to their full potential, and no solution gets anywhere near its potential unless there is an organization obsessed with getting it there.

The curve matters more than the number, and the failure to keep those curves steep is a broad indictment of our whole funder/doer ecosystem. If an organization with an idea, or a researcher with a result, can't identify promising doers and payers at scale, the result will likely be an S-curve that flattens out long before anything gets solved. Innovation doesn't matter unless there's a promising place for it to go.

Read more stories by Kevin Starr.

Kevin Starr (@mulagostarr) directs the Mulago Foundation and the Rainer Arnhold Fellows Program.

If you like this article enough to print it, be sure to subscribe to SSIR!

Copyright © 2021 Stanford University.
Designed by Arsenal, developed by Hop Studios