

Rachel G.: Thanks. Welcome to everyone. Innovate, test, scale. It sounds like an obvious idea, and yet, it's a pretty radical departure for development where too often we go by hunch, anecdote, and claim a success.

[00:00:30] I hope you've had the chance today to hear from a lot of people about some the investments that are being made in this approach of innovate, test, scale. We've certainly learned a lot about how to combat crime, how to improve education, and how to empower women.

[00:01:00] But, I want to talk about what we've learned about the process. What we've learned about the process from going from innovation to scale. And what innovations lead to scale, what types of partnerships are needed, and can donors do to catalyze this process.

[00:01:30] So, my first example, which you have heard about a bit today, is the graduation program, which started with the Bangladeshi NGO, BRAC, who noticed that a lot of people in Bangladesh could not benefit from the classic micro-credit because they were too poor. And they decided to innovate and they produced a program where they gave people an asset, a stream of income, and a lot of support to help them graduate from extreme poverty. And this was evaluated carefully within RCT, and it was proved to be very successful.

[00:02:00] And then, very deliberately, they then decided, a group of researchers decided to test this in multiple different countries with different organizations. And again, with RCT's and it was spectacularly effective. And then, the organizations who had been tested used the evidence to get additional funding and to scale it, including was support from USAID DIV and the Office of Microenterprise Development.

[00:02:30] So, does this give us our recipe for using evidence to change lives? The model that we have in the graduation program, as you evaluate a program, you work with the organization that is going to scale the program, you replicate it in multiple contexts, and that leads to improving millions of lives. Well, I think this can be a very effective approach. I think it was very appropriate to use this approach when scaling the graduation program. Why? Because it was a very expensive program, and therefore, the burden of evidence was very high. It was also a very complicated program, so you can see the benefits in a complicated program of scaling with the organization that was tested, but I think it's wrong to conclude that this is the only way to go from innovation to scale.

[00:03:00] So, I'm going to talk about some examples, where the evidence was not replicated in many contexts and I think that was completely appropriate. Where the scaling was done, not by the organization that was tested, but by a different organization. And even, I'm going to talk about some examples, actually ones that had the biggest effect on lives, which didn't even evaluate a program.

[00:03:30] So, let's start with the Generasi example, which is one that was not replicated in different contexts. This was program that was designed with the government of

[00:04:00]

[00:04:30] Indonesia and researches in the World Bank, and the idea was to give a grant to communities to improve health and education. And then, additionally, the grant, how much money they got in future years depended on how well they did in improving education and health. And it was tested using a randomized control trial in 264 sub-districts across Indonesia in five different provinces with a total of 1.8 million people in the original RCT.

[00:05:00] So, when they got the results it showed that this was effective. They had a pretty good idea that it was going to be effective across different parts of Indonesia. Here the testing was subsidized by a consortium in funders including the government of Indonesia, and many others, including the Australian government, who is helping to fund this conference, too.

So, yes, it might be interesting to see if this program worked elsewhere in other contexts, but I think it was completely appropriate to scale it even without testing it in multiple contexts.

[00:05:30] Now, Generasi is an example of the benefits of testing a program with a government can then take that program to massive scale. And working with the Indonesian government, you really work at massive scale.

[00:06:00] But, deworming is an example of a very effective scaling where it was scaled by a different organization than the one that was tested. So, in this case, school based deworming was tested with a small NGO, but it was scaled by governments. And just in 2016, just in India, 179 million children were dewormed by the government of Indonesia with technical assistance from evidence action, and that technical assistance was supported by USAID.

[00:06:30] Now, people may object and say, "How can you know that the impact will be the same, if it's scaled by government, when you tested it with an NGO?" Well, that may be an argument when you're testing a program where there's very complicated and subtle interaction between the people implementing the program and the beneficiaries, but guys, we're talking about a pill here. Right? Do we really think the pill is going to have a different effect on people if the distribution of the pill is organized by a government or an NGO? I don't think we do. And what's important is to test to make sure that we know that the pill is getting to children, and evidence action puts in place very strict monitoring to make sure that that happens when the government takes over the program.

[00:07:30] So, there's another reason why this program scaled very effectively. And that was this chart, or the data behind this chart, which shows that school based deworming is by far the most cost effective way to get children into school, of all the different programs of being rigorously evaluated. And I think, just to explain the chart for a minute, \$100 spent on deworming lead to 12 additional years of schooling. I mean, that's just phenomenal. And that means that if in a different context it's slightly less effective, or much more expensive, you can do pretty much any sensitivity analysis you want with these numbers, and I've done that. It's still going to be massively cost effective and still blow away all the other programs.

[00:08:00]

[00:08:30] So, I think one of the lessons here is that, if we test low cost programs that are easy to implement, it's easier to, then, transfer them to other contexts and other organizations. There's also a lesson for donors here, in that funding technical assistance to governments to help them introduce programs can be a very high leveraged investment.

Now, I'm going to talk about another example, which is teaching at the right level, it's something that you've heard a lot about today, but I want to make the point here that some of the most effective use of evidence goes well beyond a program.

[00:09:00] So, the story starts with an evaluation of a program that was not effective. This was the impact to providing textbooks to schools in Kenya, and the researchers found that only the very top of the class benefited. And in digging into the data, they realized that most of the kids were so far behind the curricula, they couldn't even read the textbook. So, the next piece of the puzzle is, based on that, people looked for, what can we do to help children who are falling behind? And Pratham was doing a program to help children who were falling behind in India, and that was tested rigorously and found to be very effective.

[00:10:00] Next step was an alternative way of teaching at the right level, which is to take a class and divide the class by how much, in this case, English they knew because that was the medium of teaching in Kenya. And again, simply dividing the class allowed the teacher to teach better to the level of the child, and again, that was very effective.

[00:10:30] Now, the Pratham program was tested in multiple variations in different parts of India. Another key aspect of this was that people just got descriptive data. Through the PAL network started testing and realizing that kids were falling behind the curricula all across the developing world. So, we realize this fundamental problem that the original textbook RCT showed is a problem in many, many contexts. Now, part of the story here is a classic scaling by the original organization that was tested, Pratham. And they have scaled up and reached 50+ million kids. And that's amazing, but that's not the end of the story.

[00:11:00] We also, then, have the Zambian government saying, "Well, we've got this problem, too, how can we learn both from Kenya and India to adapt programs to our needs in Zambia?" And some of you have heard about that in other panels. Then, RTI says, "Well, let's look at the original problem of kids not being able to follow the textbook in Kenya." [00:11:30] And, they go and redesign textbooks for early grade reading in Kenya. If you read about how they developed this, they spent a lot of time with kids in the Kenyan classroom. So, the textbooks reflect what the kids know and how they're learning.

[00:12:00] And finally, we see computer programs coming in and saying, "How can we use computing to do personalized learning that is really directed at the level of the child?" So, the lessons go far beyond, and the impact of the evidence goes far beyond, simply scaling a program.

In my next example of the benefits that go beyond a program, didn't even start with testing a program. It started with an idea, testing an idea, and also a massive argument.

[00:12:30] So, for those of you around and in development world, that time around 2000, there was this massive argument about whether to provide free malarial bed nets. And on one side, people argued, well, if you charge for bed nets, you're going to exclude the poor and you'll have very low take up. On the other side, people argued that, people only value something if they pay for it, and therefore, unless you charge, people, even if you

[00:13:00] give them the malarial bed net, they won't use it. And this was a pretty evidence free argument. It was massive, there was lots of heat generated, but very little evidence.

[00:13:30] So, then there were series of studies looking at a wide range of products in health asking the simple question, how much does take up for with price? And you can see, across all these different studies, you see this dramatic decline in the take up of the product, of the preventative health product, with very small increases in price. The scale there is, the furthest is \$2, so even a small increase in price you see a sharp drop off in take up. And they also found that people are just as likely to use the product if they got it for free.

[00:14:00] Now, this evidence helped clear the way for the massive roll out of free bed net distribution in Sub-Saharan Africa. So, these maps illustrate that. The one on your left shows the increase in bed net coverage across Sub-Saharan Africa between 2000 and 2015. And on the right, you see the decline in the incidence of malaria.

[00:14:30] So, recently there was a Nature article by Bhatt and colleagues modeling, estimating how much of the decline in malaria was due to the increase in bed net usage. And they estimate that 450 million fewer cases of malaria in that time span were due to the increase in the bed net coverage. And four million lives were saved from the increase in

[00:15:00] bed net coverage. So, that's a very powerful effect on lives from the evidence and shows the power of testing an idea, and not just a program.

[00:15:30] So, what are the lessons? Well, in particular, what are the lessons for donors? Given that we're sitting here at USAID. Well, I think one of the lessons is, it's important to be open, and this is very different from the classic way that donors often operate, but if you want to fund innovation it's important to be open to the ideas that are coming up about ... Because it's very hard to prescribe what innovation you want to see.

[00:16:00] The other lesson is the evidence as a public good. Yes, the organization that you evaluate benefits from it, from that evaluation, but as we see and the lessons go much wider. You know, Mindspark benefited from that testing of the textbooks, that they had to do with at the beginning, but they learned and they innovated from that lesson.

[00:16:30] The other point I want to draw, which is a little subtle, is some of these more theoretical tests, like the pricing ones, actually are more policy relevant than testing a kind of a program. Because they test a general idea, which then is more likely to be relevant to a larger number of contexts. I think we also see again, and again, in these cases that funding technical assistance and funding the governments to be able to take up the program, can be a very highly leveraged investment for donors.

[00:17:00] Earlier on, Michael Kremer talked about the importance of looking at this as a portfolio.

[00:17:30] And I think that's, again, a very important way to think about it, even just a few of these cases, scaling can justify the investment in the entire portfolio. So, it's really the long way to look at it to say, how many of these programs have scaled? The right way to look at it is, when these programs have scaled, does that justify the overall investment that is being made?

[00:18:00] And finally, this is a plug, for all of those organizations who have been a key to this, this really radical change in how we do development, to not only invest in the evidence, but really take it onboard themselves, and change the policies that they are making, the decisions they're making, internally. I think one of the really interesting things to see is, that having evidence shops within USAID, like USAID DIV, enables that group to have a much wider influence across USAID, advocating for the programs, and the lessons, that have come out of the research.

[00:18:30] And finally, I just want to say, thank you. Thank you to, on behalf of the world's poor, for this incredible investment and willingness to be open to testing that we've seen across the development community and many of the groups here, both to those who have invested in doing the evidence, but also, in particular, to the groups who have allowed themselves to be opened up to being tested because they have generated enormous public goods for the rest of the world. Thank you.