

The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high quality, educational resources for free. To make a donation or view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at [ocw.mit.edu](http://ocw.mit.edu).

**PROFESSOR:** So today, we're showing just a very short video I just wanted you to hear. And we're going to be talking about Pratham today. We've had a very long relationship with Pratham over the years, we being J-PAL. That is me and Professor Banerjee and several graduate students.

And one of the leaders of is this person, Rukmini Banerji And it's sort of useful to hear this. It's a very short segment where she explains a little bit what she sees the education problem being about. And then we'll take it from there.

[VIDEO WITHHELD]

**PROFESSOR:** So I kind of wanted to take one lecture to go in detail to all of our work with Pratham, not only because I think Pratham is an amazing organization, but also because it will expose you to how we learn about the problem, how we build from one evolution to the next, one project to the next, how one project is bringing some more questions then we started with, and how the next project tries to get at these questions, also how the partnership with the local partners develops and works and all that.

And to this, we are going to be talking about this quality problem, where we are trying to address the basic point that she makes. Kids are in school.

At the moment, she talked at the Clinton Global initiative, that was 90%. But the enrollment rates are even higher now, close to 100%. In some districts, it's more than 100%, because some of the five-year-olds also go to school, because that's kind of a convenient way to get babysitting.

But despite that, the achievement is low. She did point out the results from the ASER survey. So every year, since 2005, Pratham puts together teams of people,

local people, who go from local universities, et cetera. And they go in every single district in India.

How many districts did she say there was in India? She mentioned it at some point. You see that everything [INAUDIBLE] if you are listening to the video. She said it at some point, how many district they work in and how many districts they have.

If I'm not mistaken, she said something like 600. So they are sending people to every district in India, teams of college students, local volunteers, et cetera. They have a sampling. Of course they don't go to every village. But they do a sample in each district. And they administer a very simple test of reading skills and math skills.

So the reading skill test looks like a page. On the left side, you have a very simple story. Like she said, I go to school. My brother goes to school. We like going to school. On the right side, you will have single words and then single letters.

First, you start the child with a very simple story. If they can do that, you bring them to a slightly longer story and try to measure their comprehension. If they come to the very simple story, you ask them to read a word. If they can't read a word, you ask them to read a letter. So you, in this way, classify your child as a nothing reader, a letter reader, a word reader, a small paragraph reader, or a story reader.

And the simple, small paragraph reader is what you should be able to do at the end of the first grade. And what they find is, on average, in 2005, that about 35,000 children, aged 7 to 14, most of them are in school and could not read a grade 1 paragraph and 60% could not read a grade 2 story.

So these are numbers on one of the ASER survey. So blue is better in the maps I'm going to show you. And this is the fraction of kids who are out of school. And so that's quite a blue map, because there are not very many kids who are out of school.

And the colors here are the same. This is now the reading ability. This is the fraction of kids who can read a standard one text. So it is the fraction of kids who can read or understand one text among the kids who are enrolled in standard three to five.

So among children who are enrolled in standard three to five, how many of them can read this grade one paragraph? And you can see that now we have a lot of red in this card. The red is a bit distributed in different places, but mostly it's quite red.

If you know India, does this map look like what you would expect for any social indicator within India? For anyone who knew India or about India, because this is a question that you might not know the answer to.

So it doesn't really, because the folklore in India-- and if you look at immunization, that's really what you get-- is that sort of the bad states are these ones, Uttar Pradesh, Bihar, Rajasthan, Jharkhand.

These are what they call the BIMARU state, Bihar, Rajasthan, Uttar Pradesh. This is like where you have a lot of boys and not very many girls, suggesting infanticide or selective abortion. This is where you have very low immunization rate, bad outcomes, of course, as I mentioned.

And these are the good states, Andhra Pradesh, Karnataka, Tamil Nadu, Kerala. Kerala is, in particular, one of Amartya Sen and John Rawls' favorite states for its successful education. Kerala does well here, almost everyone can read.

But what is surprising is you have Tamil Nadu. They're very proud of themselves. A lot of the kids are in school, et cetera. It's very red. Less than half the kids enrolled in [INAUDIBLE] three to five cannot read a sentence.

So that upset the government of Tamil Nadu a great deal. So when the first ASER report came out, they didn't believe the result. That is not possible. They went ahead and did their own test, but they found the same thing again. So this suggests that there is something different.

It's not only like disorganized places that are unable to deliver to their kids. There is something maybe deeper than that. And Bihar did surprisingly well in not being at the absolute bottom of the barrel in this graph. So Bihar was happy.

Another thing that's is maybe even more troubling is this survey has been done from 2005 to 2010, every year. It's been released in January every year, on the Republic Day in India. And there is no progress.

So India is progressing across all these dimensions, becoming richer, there are fewer poor people. We're seeing they're eating less and less, which is not progress, per se. But education, nothing, no progress.

Also troubling is it's not an Indian exception. Surveys very similar to ASER have been done in Kenya and Uganda and Tanzania and in Pakistan. And they all find similar kind of things. A lot of children are in school, most of them can't read anything. Yep?

**AUDIENCE:** Is there still progress across the board on average, like in the nation? Or the states are static as well?

**PROFESSOR:** So you're exactly right. That's a very good question. It's no progress across the board in the nation, but the states are not static. Some states start doing well for a while, and means some states are moving down to compensate.

So, for example, in the last report, 2010, Punjab had a big progress. So that's one of the things they are trying to look at is why is Punjab suddenly progressing? What are they doing that is different? And I'm going to be able to talk a bit about that.

And in fact, in the Punjab example, what they seem to be doing that is different is really to try to insist on these core competencies in the schools.

They have a period of time, every school day, two hours a day, which are devoted to teaching the kids at whatever level they are forgetting the curriculum. And that is what they think is the reason why Punjab is progressing.

So what we are going to do today is-- do we have evidence to suggest that, in fact, they're right? And that is the key problem and, therefore, the key of the success of Punjab. Meanwhile, some other state, like Uttar Pradesh, would be going down.  
Ben?

**AUDIENCE:** I have a question. Is the student to teacher ratio more or less? I know these states are huge. But are they more or less the same across the different states?

**PROFESSOR:** So it's a very good question. I think, they are, probably, generally more or less the same. And they are not horrible. Because, actually, being a teacher is not a bad job. So there are a number of teachers who are available.

So you don't have the very, very, large class size that you have, for example, in a country like Kenya, where, after free primary education, the class size in grade one was easily 70 to 80 children per classroom. You have less of that in India. Yeah?

**AUDIENCE:** So it seems like the enrollment is high, but do they also track that the attendance is high?

**PROFESSOR:** So that's a very good question. And the answer is that, yes, they check attendance. And attendance is low. So attendance is low, both for teachers and for children. So in Bihar, for example, we have a survey of children's attendance.

The rate of absence is between 30 to 50% for kids. And we've done the same thing in several countries of the world. Generally, children's absence is very, very high. So kids are enrolled, but they don't show up. And it's not that they are ghost enrollees. This 30% attendance is not concentrated across 30% of kids who never come.

It's every child who is absent one day, coming one day, absent on day, coming one day. And that is one problem. So we have a lot of absence of children. And we have a lot of absence of teachers as well. So a survey was done in four to five developing countries, by the World Bank, including India. They found a very large teacher absence rate.

So in India, it was about 25% absence rate. And that's about what ASER found as well, of teachers. So you show up to a school during school hours, and you have three of four teachers who are actually in the school.

The second problem is that once they are in the school, it is not a given that they are actually teaching. So another one out of these four teachers is actually in school

but not teaching. There was a funny report, a few years ago, describing what they were doing, ranging from drinking tea to drawing political posters.

The reason being that a lot of these teaching jobs are patronage jobs, which are filled by the political parties, so they need to draw their posters sometimes. So that means that the teachers are present teaching half the time. The students are, themselves, present about half the time.

That means the students get about a quarter of the time that they should get in front of a teacher. This is a where, [? Hookman's ?] point, that you need a bit of time of student-teacher interaction to get teaching. It actually becomes quite pertinent. Yep, sorry?

**AUDIENCE:** I have a question about the standard curriculum. What kind of things does it teach? Is there any [INAUDIBLE] benefits from it? Are there any kids who are up to the curriculum?

**PROFESSOR:** So that's a very good question. What is the curriculum? So the curriculum in India-- and that actually is true in the countries I know-- tends to be very ambitious. And there are various reasons why those curriculum are very ambitious.

They are probably more ambitious than what is being taught in the US. Actually, significantly more ambitious than what is being taught in US primary schools. We once showed up to a school. We were with the Pratham team. The visit was announced. The teacher wanted to look good.

And his idea of looking good was to draw things like that on the board, in this grade three class, so all the kids were like looking around. And it doesn't mean that this is what he actually does on a day in day out basis, but it means it is what they suggest to do.

So it's not necessarily a bad curriculum. In fact, India has a lot of good, very good, very serious education experts who are thinking about educating the child to a real, true understanding of science. But for most of the children, it's completely above their head, very quickly.

And then some of the children benefit from it. We can see an example of that very quickly. We'll see an example of that very quickly. Yes, [INAUDIBLE]?

**AUDIENCE:** Is their proficiency in math as bad as in [INAUDIBLE]?

**PROFESSOR:** It tends to be worse. So math is a bit worse than reading. For example, if you look at division, which is the equivalent of your standard two kid, you have like 20% of kids can do division.

So that's where we are. So this quality of education, I was saying the other day, has not concerned the policy establishment, very much, for many years. The policy establishment has been more concerned about getting as many children as possible in schools, enrolled in schools. With the understanding that, oh, if we can manage to get them in schools, somehow learning will happen.

And the Turkey experiment is, in a sense, an archetypal example of that, which is let's make education compulsory, let's have boarding schools, let's have buses that are going to bring the kids to the boarding school. And then black box, learning happens. You have an educated child.

But it's been about 10 years since academics have been a bit puzzled at that. And looking at the quality of education saying, how do we make education actually work? What's important? And the first thing that jumps at you, when you go to a school in India, in Kenya, even in rural Morocco, which is a richer country, is that there is nothing in the school.

The school is usually very bare. You have fewer desks than students. People are bunching like three to a desk. They will usually have a blackboard, very few textbooks, and, of course, nothing like computers and things like that.

So the first generation of studies was about, well, they have none of that. How can they learn? Maybe what they need is more material.

And in fact, when the idea of randomized evaluation started to be applied to

development, Michael Kremer, who is a professor at Harvard, wanted to do a demonstration project, wanted to show that it is possible to do a randomized evaluation. And it's interesting. And he wanted, for demonstration purposes, an example where it would clear that it has an effect.

And so he asked around, et cetera. And he found out that what would be obvious is textbooks. In Kenya, no one has textbooks. If only we give textbooks to the kids, obviously, they are going to learn better. So that was his first example that he chose, which was chosen specifically to demonstrate success.

And running his experiment, he was learning how to do an experiment at the same time. So his first experiment was very small, 14 schools, 7 treatment, 7 control, and found no effect.

He thought about it, and he said, oh well, of course, it's because I have so few schools. And you remember, when you have schools, you need to take into account, when you calculate your standard error, that all these children are similar.

So he said, oh well, my experiment is not big enough. So let's take a bigger sample of schools. He took 100 schools. He took half of them and again found no effect. So there he thought, well, maybe the test is too hard. I'm using the test, the normal test. Maybe that is too hard. Let me use a different test that is going to be able to discriminate progress even at a lower level.

And he got his experiment started. And again, he found no effect. So there he started saying maybe there is no effect, even in a larger sample, even with a better test. So maybe there is no effect.

So eventually, what he found is that, if you focus on the children, who were already doing well at the baseline, then they benefit from the textbooks. So that's a very long winded answer to your question, which is there are children who benefit from the curriculum and, therefore, benefit from the textbooks. But it's a small minority.

And the explanation he gave is the textbooks are in English, which makes sense, because the curriculum in Kenyan schools is in English. But a lot of children don't

even speak English, because English is their third language.

They first speak their mother tongue, the local language. Then they learn Swahili in the first few grades. And then English is introduced in grade one as a language, and then as a language of instruction from grade three or grade four.

But the problem is that because the kids aren't actually learning effectively, by the time they reach grade five or six, they actually don't know English, so the textbooks are of no use.

So exposed, it could be understood what was going on here. And then they tried a bunch of stuff. Flip charts also have no effect. Cutting the class size in two, if you make no other differences, also has no effect.

And a little bit worrisome, things like the deworming program, which increases attendance, but it didn't increase test scores. So it seemed that these extra school days also had no effect.

So changing inputs just didn't work. So maybe it's a problem that is just incredibly difficult. So we were there a few years ago. So what this common with all this intervention is that they are just changing the inputs.

They are doing more of the same, adding more textbooks, adding more teachers, adding more resources. But they are no change to the pedagogy and no change to the incentives. So no one is given incentives. The teachers are not given incentives. The students are not given incentives. The parents are not given incentives. And the pedagogy and the curriculum stays the same.

So that's kind of where we were about five, six years ago, a little bit depressed. Picture us, this was our first set of randomized experiments and nothing works. So we were thinking, maybe, this is going to be the end of randomized experiments, because you can't be a doomsayer forever, otherwise people really hate you.

So that's why we were a little bit sad. Rukmini came to visit us, shortly after that happened, a few years after the establishment of Pratham, came here to MIT.

Because a former MIT student had gone to work for them as an intern. A former undergraduate student of ours had a relationship with them.

So Pratham started in 1994. At that time, it was established by UNICEF to help some kids in Bombay. It would be your typical, small NGO doing some work in Bombay, particularly what they call bridge classes. Which is take a kid who is out of school, give them a course for a few months so that they can try and go back to their regular school. That's what they were doing.

But they were ambitious. So there was the Rukmini that you heard and Madhav Chavan wanted to make a real difference. So since then, they're grown substantially. They've reached millions and millions of children. I think there are about 38 million children, who are reached by the Pratham program, one way or the other, so about half the population of France, overall population of France, just in terms of scale.

And so that's the largest non-governmental organization to do education in India, probably in the world. And their motto is "every child in school"-- that half of the world would agree with them-- "and learning well"-- is where, maybe, they have a little difference.

So they came to us towards the beginning. And that's the reading that read for today. They wanted to evaluate the Balsakhi program, which was the flagship program at the time. Balsakhi means the friend of the child.

So the Balsakhi is a young woman, from the community, so an adult. Like she said, you need some adult time. Barely an adult, some 18-year-old, usually having only high school education, so grade 10 to 12.

And Pratham would give them a very short training and then dump them into the school, with which they had an agreement with the teachers that the school would let them pull out the kids who were lagging behind in grade three or in grade four.

So why is this a good demographic, the 18-year-old, grade 10 educated woman? Why is that a good goal to work with? Yep?

**AUDIENCE:** She's more familiar with the community, so she can probably relate to [INAUDIBLE].

**PROFESSOR:** So that is one reason is that she is more known to the community, because she's local. She's not much older. And she's not intimidating or scary for the parents, who, maybe, can work better with the kids. That's a very good reason. What's another reason?

**AUDIENCE:** I would say for the teacher's as well. Maybe the teachers are not as intimidated that the balsakhi will go and teach these students some really innovative thing that would make [INAUDIBLE].

**PROFESSOR:** Yes. So for the teachers, she might be not very intimidating, which will cut both ways. On the one hand, it's more easily acceptable for the teachers. On the other hand, there is a tendency that teacher would use the balsakhi to make the tea or other activities like that.

And we are going to see some of that happening, not actually in Baroda. Yep?

**AUDIENCE:** So she's locally trained, which makes sense in kind of [INAUDIBLE] perspective.

**PROFESSOR:** Yes. She is locally trained, only for two weeks. This is very cheap. And there is another aspect of the cost. Yeah?

**AUDIENCE:** Not about the cost, but she might be more motivated to help the [INAUDIBLE] children, because it's not like this is a second job she can perform. It's not a fallback option. It's something that she wants to do that is a period in her life. It seems a bit time sensitive.

**PROFESSOR:** Yes. So she might be more motivated for this reason. Also for what you said earlier, that she's close to them, et cetera. But also this is something that she effectively chose to do at this point and not something that is just putting money in a bank account.

**AUDIENCE:** She's not highly qualified, so the [INAUDIBLE].

**PROFESSOR:** Right. She's not highly qualified, so she's cheap. What is another reason why she is cheap? Why don't these people go and do something else?

**AUDIENCE:** Also there's a high turnover in these [INAUDIBLE] make a certain income to teach, so you're not really relying on this thing. It's [INAUDIBLE] quality of the teaching. It doesn't have to be like a small, really enthusiastic [INAUDIBLE]. So it's easy to get [INAUDIBLE].

**PROFESSOR:** Yeah. They have high turnover, which could be a plus or a minus. On the minus side, of course, you don't get the experience. On the plus side, you don't get tired. And you don't rely on people who have a huge vocation. You rely on people getting enthusiastic for one year.

So very much along the model of Teach For America, where you get MIT, graduating from school, energetic about wanting to do this. A bit of what you're saying, which is it's what I want to do at this point in my life. And then by the time you don't want to do it any more, you stop.

And there is one more thing about them, specifically, which makes them cheap.

**AUDIENCE:** Everything about them. I mean I think there's like a good chance that they have families or that they're part of a family, and maybe that's why they're stuck in the area. And they don't necessarily have the ability to go somewhere else.

**PROFESSOR:** Exactly. That's a very important point. Most of these women are unmarried. They are living at home. They are sort of waiting to be married. And there is actually not much they can do. Because the parents won't let them leave.

And mostly, their parents won't let them work, because working is not becoming for a certain, sort of lower, middle class category woman, waiting to be married. You shouldn't be like working. You should be able to be-- it's not a good signal for the family that they're not able to provide for you and all that.

But this is hardly seen as working. This is more like volunteering and helping your community and things like that. So they can do that. So they are sort of available.

So one part of the genius of Pratham is that they identify a group of people who are actually sort of available for free. Here, they pay them, but in future programs, that I'm going to talk to you a minute after, they're not even paid.

They are just unpaid. They do that for a while. And that get some, presumably, other rewards of doing that, which are more in the form of intrinsic reward of seeing the kid progress, and, therefore, you also select people who are motivated in that dimension as one of you pointed out.

So that's all of the pluses. The minuses you already mentioned. The turnover is one of them. You can't get them to get experienced.

The second minus is that they're, of course, much less educated. Maybe there is some good reasons why a teacher needs a college degree. Maybe it's actually terrible to take the kid out of the classroom to put them in front of someone who knows nothing. So that's kind of the risk that would potentially be there.

So that's the program. And when they came to see us, we were wondering how to evaluate this program, with them, in an experiment, in a way that is not going to cause a problem for them in the city, that's not going to hinder their work too much.

And that's experimental design that we decided to adopt in a new place. So Vadodara and Gujarat, those were places where they had not worked at all. Had about 100 schools. A little more than 100 schools divided in two groups, randomly, with the computer, group A schools and group B schools.

The group A schools got the balsakhi for grade three children. The group B schools got the balsakhi for grade four children. So what's an advantage of this design from the point of view of political acceptability? Why did we go like this, instead of doing the more standard thing, which would be to say, well, in group A treatment, and grade three and grade four get the balsakhi and group B, no one gets a balsakhi?

**AUDIENCE:**

Well, in a way, if you don't give one group any treatment at all, you're almost sabotaging your education [INAUDIBLE].

**PROFESSOR:** Sabotaging is a strong word. But exactly the idea is that this allows you to be present in every school. So Pratham is not infinitely rich. They can easily argue-- and I think that was very true at that time-- that look, we can give you one balsakhi per school. We can't afford more.

So you either put it in three or you put it in four. We can't give you both. But at least every school is being engaged with. So they are engaged with the system as a whole. And no school is disadvantaged.

That makes it easier for the school system to say, yes, they are just working with us. They are part of us. It also makes data collection much easier, because you're involved with everyone.

What is a potential danger of this design? The plus is what I just said. What's the potential minus? Since you have a balsakhi in every school, what's the potential minus?

**AUDIENCE:** They might not just be teaching grade three. They might be overlapping grade three and four or something.

**PROFESSOR:** Exactly. The balsakhi might take both groups, so that your control group becomes partly treated. Or even if it's not the balsakhi, the head teacher might say, oh great, we have a balsakhi for grade three. So we're going to take the teacher out of grade three and use it to divide grade four in two. And so grade four would be associated.

Or they could say, well, since the grade three got the balsakhi, when some other organization comes with computers, well, we'll put them in grade four. So there is a danger of resources being reallocated across grades in a way that contaminates your treatment.

So fortunately, that was not really an issue there, because in Vadodara, the way the school system is organized, a school is made of one teacher per grade, regardless. Sometimes in the same school building, you have more than one school. And that happens sometimes in the US as well. You have this charter school and the regular school sharing the same school building.

In Vadodara what happened is, if it's a big school, it's actually a big school building, but under totally separate administration. So a school is made of five teachers and a head teacher. That's it. So there is less scope for reallocating, because you have your one teacher anyway.

And they're also not very imaginative in terms of using the resources in the most optimal way. So that turned out not to be a problem.

So the way we evaluated the program is then by comparing grade three, group A's treatment compared with control. And in grade four, that's the opposite. We also worked in Bombay. So that's the second year we have asked a group.

So if you were a kid who was in grade three, in group A, you are treated in grade three and then, again, you are treated in grade four, because you would have moved one grade.

If you were a kid who entered in grade four, you would never be treated. Or if you were a kid who entered grade three in group B, you also would never be treated. So we have kids treated one year, kids treated two years, and kids treated one year.

And then in Bombay, we did something similar. This is Shobhini She used to work at Pratham. And now she's one of the executive director of J-PAL South Asia. You can say, hi, to her.

And in Bombay, we did something similar, except we started in grade two, three, and then moved it to three, four. The reason why we switched is that they realized that they didn't have right pedagogy for engaging with the grade two children, that the grade two children were too small, and they didn't know how to deal with them.

The initial plan was to do two, three and then reverse again, two, three. And then they turned out that they weren't happy with they're grade two, so they moved three, four.

So that's kind of the idea. So this is the design. So in principle, now, we could say,

well, they're good. We have the designs. Now, all we need to do is to compare treatment and control after one year and after two years, and we're going to be done.

But unfortunately, when you run real experiments, you get real problems. I've yet to meet an experiment where there was no problems. And this was, at least, my first big one. So we had a lot of problems that I want to talk about.

The first one is the way to evaluate the program was to administer a test in the school. And we had some issue with the test. The first one is that, as I was saying, a lot of kids are absent. And if we only tested the kids who were present in school, we have some attrition.

Maybe in urban India, absenteeism is less bad, but maybe 20% of kids are absent on a given day, so they don't take the test. So why is that potentially a problem to have missing kids in the test? Yep?

**AUDIENCE:**

Maybe the kids that are absent are those that are more likely to not [INAUDIBLE] class, and so they either chose or their parents chose that, maybe, they could better use that time working from home. And so all the results are biased. Then you're only really testing the kids that understand more and you're ignoring the kids that understood the least.

**PROFESSOR:**

Exactly. So that could be the first problem, which is those kids, whom we are testing, they are not representative, because maybe they're the one who understand the most.

And that could lead to an even bigger problem, because suppose the intervention is effective and most kids now understand better. Then what could happen with the attrition in the treatment group versus the control group?

Now you all know very well, it's the kids who don't understand who don't come. And now the kids who understand better are in the treatment group. What might happen to attrition?

**AUDIENCE:** The attrition would be less for the students who are in the treatment group.

**PROFESSOR:** The attrition would be less for students who are in the treatment group. And the kids who are coming tends to be the weaker kids. So now we are comparing the strong kids in the control group to the strong kids plus some of the weak kids in the treatment group.

So that will tend to make our estimate look smaller relative to what it should be. Because now we have a different population of students in the treatment and control groups. So we have a bias here. It goes in the direction of not finding an effect. So if we still find an effect, we're happy. But it's not the right estimate.

So to solve this problem, what did we do? Did we rely, uniquely, on the test in the school? Yeah?

**AUDIENCE:** You found the kids, even if they happened to be at home or elsewhere and administrated the test [INAUDIBLE].

**PROFESSOR:** We went and looked for them, wherever they were, except if they'd left for the village. So the attrition ended up being low, less than 10%, and similar, now, in treatment and control school. Because it's not because people don't understand that they leave for the village. It's because they have something to do there.

So that was the first thing. The second problem we had is that the testing instrument-- actually, not in year 1 but in year 0-- we designed a test which was at grade level. We designed a test to test the curriculum.

So it's like, if this is what should be tested, what's on the board, here, we tested at that level. So if they were supposed to learn Euclidean geometry, that's what was tested. And we had a big problem, which is that the teachers cheated like crazy.

So the reason we saw that is that all of the tests were the same. So clearly, the teacher had written the answer on the board, and all the students were copying down. In one class, all the students had the same name. So the teacher clearly wrote the name, like a sample name, on the board. And all the students dutifully

wrote down the name that they saw on the board.

So we said that's not working. There is a paper by Steve Levitt, who shows very subtle things, how you can detect cheating. That was easy. It didn't require a lot of imagination. It was pretty obvious that cheated like crazy to make their kids look better.

So we thought about how we can solve this problem. We are going to administer a new test without the teacher, that some special Pratham staff is going to come to administer the test. And we did that. And then we realized that most of the students now had 0 on the test.

So this test was way too hard, which is why the teacher had been cheating, because they were a bit embarrassed. So the solution here was to develop a much easier test that covered the competencies starting from, can you write your own name, to-- I don't know--  $2 + 2$ , finishing with a few questions such as as, Nancy goes to the market. And she buys three onions at 15 rupees. And how much change does she have, that kind of thing.

So that's the problems that we had. Further problems we had, we had a problem in Bombay. Because the second year of the evaluation turned out to be 2002 to 2003.

Before that, September 2001 happened. We were walking in a quite Muslim neighborhood of Bombay. A lot of kids were named like Osama and things like that. People were pretty upset about the US at the time.

And so Pratham, maybe because of us, maybe just it would have been anywhere, was somehow seen, maybe, as an American presence. So some schools said that they don't want to work with the program. So about 30% of schools said that they did not want to work with the program. Or, in another set of schools, someone was identified, but they couldn't read, themselves, so they couldn't really be entrusted to teach reading to the kids. So they also didn't do it.

So can we just drop all the schools that refuse to participate from the program? Can we just drop them from the analysis?

**AUDIENCE:** Well, that's a selective bias. There is a some reasons for which they dropped, which make them inherently different from the other schools.

**PROFESSOR:** Exactly. So we can't do that, because that's a selection bias. Maybe these are the weaker schools who are refusing the program. Maybe they are the stronger schools, because they are independently minded. Who knows?

But what is clear is that's it's certainly not random. So we lose random control comparison. So what we do? Well, the first thing you can do is to say, well, I'm going to not measure the effect of the balsakhi program.

I'm going to measure the effect of my intention to treat the schools with the balsakhi program. So we called that intention to treat. What an intention to treat is a Pratham official coming to the headmaster and saying, hey, we have a balsakhi for you.

So this, we can do in an unbiased way, right? Because if we compare all the test scores in the school where someone has come and offered the program to the test scores in the school where no one has offered the program, there is no selection here.

But it's not the effect of the program. It's the effect of my attempt to provide the program. So that's why we call it an intention to treat. It's not the treatment effect. It's the wishful thinking. It's like my attempt to treat you.

Now, I wrote, here, on the board, if we wanted to know the effect of the program itself not the intention to treat, we can divide the intention to treat effect by the fraction of schools who accept the program.

So, for example, if 70% of schools accept the program, and we find an effect of 10 on the test scores, a gain of 10 points in the intention to treat, the intention to treat is 10 points. That's due to only  $\frac{3}{4}$  of the schools.

So we can blowup the estimate by dividing by  $\frac{3}{4}$ , which is about multiplying by 1.25. So instead of 10, the effect is 12.5. So this is what's called a Wald estimate.

When you divide your intention to treat by the fraction of the take-up of the program, which, here, we call that the first stage.

So in the first stage, you offer a program, and then some fraction of people take it. In the second stage, that has an effect. And the combination of the first stage and the second stage produces the intention to treat.

So now to go from the intention to treat the effect of the program, you do a scale-up by dividing by the take-up. So this is, in five sentences, you're instrumental viable estimate,

We could spend about a whole semester on it. There are conditions under which this is valid to do this. There are conditions under which this is not valid. And this has interpretations which vary. So Josh Angrist, in the econ department, is the one who did most of the work on this.

So the point is that-- but think of it in this simple way, for this kind of noncompliance problem-- is let's just normalize the estimate by dividing by the take-up. It would also work if you have some school in the control group that managed to get the balsakhi anyway.

Then you could divide by the difference between the take-up in the intention to treat group and the intention not to treat group. So if 10% of schools manage to get the program in the control group, you would divide your intention to treat by 27% minus 10%, that's 65%. And you would blow it up that way.

What it amounts to doing is to say, this is my estimate of the effect of the program, assuming that the only reason why they're in the intention to treat effect is because those who might try to treat are more likely to be treated.

So here are the results. The results are expressed in standard deviation of the unique test scores. What this means is that we take the mean of the control group and divide by the standard deviation. So the advantage of doing that-- maybe you can see that more in recitation-- is that every education program does that. So we can compare our results to what other people did.

For year one, this is the standard deviation. This is the test score in treatment and the test score in control. That's the difference, 0.17. That's the standard error of this difference below. So you divide one by the other. You get a t-stat of about 1.7. So that's significant but at a 10% level.

In year two, kids progressed. And the difference is much larger. So you get an effect of 0.4 standard deviation in math and a 0.25 standard deviation in language and something similar for the standard four classes.

So those are a large effect. To scale that up, one of the most famous education experiments is the Tennessee STAR class-size reduction experiment, where that reduced class-size in Tennessee from 20 to 8. It was very, very expensive program that had an effect of 0.2 standard deviation.

So 0.2 is considered to be a reasonably large effect. Here, in the second year of the program, we are way above that for math. Yep?

**AUDIENCE:** How big is the standard deviation relative to the average score?

**PROFESSOR:** So this is all normalized. So this is, in a sense, already standardized. This is already standardized to the average and to the standard deviation.

**AUDIENCE:** But there is like a wide distribution in the classroom?

**PROFESSOR:** You mean, is it wide? Yeah, so you can see, it's not very wide, in this case. Because the students are so weak that, I think, the distribution of test scores, in terms of points, is not that large compared to what you could find elsewhere. It's kind of a relatively tight distribution.

So in terms of points, it's useful. There are other ways to look at this than standard deviation. For example, the fraction of kids who can do these kinds of things, so in percentage points. How many more kids can do division? How many more kids can do addition? Which are also useful to look at, which I don't have with me. That's a good point.

So these estimates are large, but they're not very precise, because there's a lot of differences between kids. And one thing that really helps to control this noise is to control for how good where they before. Because test scores of children are extremely stable over time.

The biggest predictor, sadly, of how a child does is how they did last year. And so, when we control for that, we get much more precise results, which are here. So these are a bunch of results.

So this is the improvement in year 1, for both cities together, where we get an improvement of 0.2 in math, 0.3 in language-- sorry, opposite, 0.2 in math in year 1, 0.3 in math in year 2, for verbal, 0.7, not significant in math, for verbal, 0.7 for year 1, 0.15 for year 2.

And then we can look at various aspects of it, where we get a slightly bigger effect for Vadodara and Bombay but, generally, quite similar results for math. And language has bigger effect in Vadodara than in Bombay. Bombay starts from a higher level. But across the board, except with Bombay year two, you find pretty large effects, quite large effects.

So this is very different from what we had with the test scores. Now, you might still wonder that these are for all the kids involved in the program. And you might still wonder whether the effect comes from helping the lower achieving kids, which is what you're trying to do, or it's coming from helping the higher achieving kids by removing these disturbing, lower achieving kids from their classroom for two hours a day.

That would be fine. But the distributional effect would be very different. If the idea is that you just send the kids to recess, and they kind of hang-out, and learn strictly nothing, meanwhile, your 20, high achieving kids progress, that could give us those mean effects. But that is not necessarily something that we would be very excited about.

So for that, what is important is then to look at what is the effect by sub-group? And

what was the finding when we looked at sub-groups, who benefited?

**AUDIENCE:** The lowest one third benefited the most.

**PROFESSOR:** The lowest one benefited the most. And who was more likely to go to the balsakhi? Also the lowest one, fortunately. So when you look at the lowest one, you start finding a pretty large effect for the bottom third and a much weaker for the top third.

For example, if you look at year 2, the effect for the bottom 1/3 is 0.5 standard deviation. That's starting to be a huge effect, 0.5 standard deviation progress. 0.3 for the middle one, and 0.03 for the top one, no effect.

And this is your chance to go to the balsakhi. So your chance to go to the balsakhi declines with the group. So the point estimate of the effect declines at about the same rate as the point estimate for the balsakhi.

So the next thing we did is a little bit the same exercise as adjusting the Bombay estimate, to go from this estimate of the effect of putting a balsakhi in the classroom, to are you going to the balsakhi?

So basically, the effect, this 0.5 standard deviation, is due to about 20% of kids who are actually going to the balsakhi. Because there doesn't seem to be an effect for the other children, which we see from looking at the pattern declining as the pattern of take-up declines.

So now you can divide this 0.5 by 0.2. And you get the effect for people who go to the balsakhi of one standard deviation, which are very, very, very large effects, in the education literature. And if you divide 0.32 by 0.16, you'll also find about 1.

And this is, of course, insignificant. But if you divided 0.4 by 0.06, taking the estimate seriously, again, you would find the same type of effect.

So this is a program that was highly effective, for kids who were sent to the program, but had no effect on the kids who, in principle, should have also benefited from the reduction in class size.

So what did we learn? We learned that it is possible to make a lot of progress, with a grade 10 educated woman trained for two weeks. And we also learned that teachers are not very good at exploiting freed-up resources. Yep?

**AUDIENCE:** How much variability was there between the different thirds? If the top third isn't that much better off than then the bottom third, then that would say something different than, if the top third is very well educated whereas the bottom third is really not.

**PROFESSOR:** So the bottom third, basically, knows nothing at all, for starters. Like they can't do anything at the beginning. They can't recognize letters, can't recognize numbers. Whereas the top third is not at grade level but can at least do something.

So what I'm saying is not that we shouldn't care about the top third. We should care about them. But that program was really targeted toward the bottom.

And so the big difference, in the context of this program, is that the top third didn't get the benefit from the program, because they weren't sent to the program. Except that there could have been indirect benefits from the fact that the teacher's now have a much smaller class size. Instead of 40 kids of very heterogeneous levels, they now have 20 of better level. So they could have adjusted their teaching to do better by these kids. And it seems that they didn't.

So in terms of the whether this is an optimistic or pessimistic conclusion, it's kind of the glass is half-full or half-empty. The half-full part is you can do this. It's reasonably easy to make a lot of progress. Perhaps because kids start from such a low level, to go back to your question, initially.

If the standard deviation is low, because everyone is at 0, then it's easy to make some progress. And you're going to see it very quickly. On the other hand, the teachers seem to not really be using the resources in a way that allows them to take advantage of this.

So that's where we were at the end of this. And we had a bunch more questions that we wanted to ask. What do we still need to know? So suppose we would want to go from this program to, say, well, let's do a policy for all of India. What else do

we need to know before we move further? Yeah?

**AUDIENCE:** The thing, for me, is to understand what is the relation between attendance and how much those kids, those programs are actually effective? Because, I mean, you showed that enrollment, as a whole, doesn't have too much of an influence. But maybe attendance would help.

**PROFESSOR:** Right. So we could say, well, maybe there are other things to look at, children's attendance, teacher's attendance. Maybe the big difference with the balsakhi is that they were there. Maybe that's just it. Can we change their learning just with incentives?

**AUDIENCE:** Also, maybe look at the curriculum. Maybe the reason why the children are learning is because the material is more basic, whereas if you just stay in the classroom, the material is already too hard, and you're not able to do it.

**PROFESSOR:** That's the other big contender. So you have both contenders. One is, OK, the balsakhi are actually there. The kids are going. That's why they are learning. The second is that's not-- the incentive to this is really the pedagogy, which is about the learning material. And if we could train the teachers to train at that level, we would have the same effects.

**AUDIENCE:** The cost effectiveness of the program, [INAUDIBLE].

**PROFESSOR:** Right, so we could say, what's the cost effectiveness in the city? We could look at the cost effectiveness elsewhere.

**AUDIENCE:** Also, it could be interesting to see the effects of having more strict, I guess, passing policy. Because a lot of kids are advancing to the next grade without actually being prepared. And maybe making that a little bit more strict, so that they don't go to next level until they're actually prepared. That could improve.

**PROFESSOR:** Yeah. It's a very good point. In fact, the policy is not strict at all at the moment. Rukmini says in the short video, everybody passes, regardless. So maybe having something where, actually, repeating is allowed or remedial education of some kind

is provided would help.

The other question one might ask is, is it only an urban phenomenon or would it also work in rural schools? Do you need to pay the people or could you have volunteers? Would it be sufficient to distribute materials, if it's a material question?

Would you be able to motivate the teachers to do it, to focus on the students? Can you concentrate even more on the basics and make even more progress on the basics, by not trying to do remedial of everything, but just focusing on learning?

So those were the questions that Pratham had. Those were the questions we had. Another question you could ask is, on the basics, you could ask the opposite question.

You were saying that the students at the top, it's not that they are exactly ready to come to MIT. They are still way below grade level. Is there something that can be done for them? Or does this very simple pedagogy that Pratham has work only for to very low achieving kids?

So we had all these questions. But it was many years ago. We had a lot of time. So we started looking at them. So the first thing we did was a new evaluation in Jaunpur, in Uttar Pradesh.

In the first program, Pratham was running this program. And they called us to evaluate it. And we went and evaluated what it is they were doing. But with the Jaunpur program, we were working together.

So this study is, actually, authored by Rukmini Banerji, along with the rest of us. We were now all working together to try to figure out, can we learn more what's going on?

So Pratham renamed the balsakhi program as the Read India Program. And as they renamed it, they shifted to focus it on something even more basic and simple, on reading, and tried to spread out much more everywhere, not focusing on cities but also working in rural areas and working on a much larger scale.

So we worked in rural Uttar Pradesh. We had three groups of villages, here, in Jaunpur. One group, where we just went to see all the parents, had village meetings, the women and Pratham.

We went to see the parents and say, you know what? There are things you can do. You can advocate. We can lobby for more resources for your schools. You can get an extra teacher that the government will have to pay for. You can get scholarships, et cetera. That's the first treatment, to see whether parents would be able to engage with the system.

The second thing we had is we realized that parents were overestimating how much their kids knew. So a first step, that Pratham has discovered, to get people excited about reading, is to train parents to administer the small Pratham test. So you go and say, Ben, read this. And Ben is like staring blankly. And then they realize, oh, my god, the kids can't read.

And they have been giving their kids to the schools, faithfully, for years, and assuming that something would come out of it. And then they realize in this exercise that actually, not really. So the parents prepare a report card for the village. And then there was a lot of discussion.

And the third was the Read India volunteer. At the end of this process, Pratham asked anybody, are their volunteers to learn Read India? In this case, they were not paid. People just came up, boys and girls, usually youngish boys and girls. They got trained by Pratham and started running these camps for the students, Read India camps.

And what did we find? That's kind of the result in one graph, kids who, at baseline, could not read anything. And this their result at end-line. These three lines look the same, control, information only, information plus test provided no significant difference.

But there's a somewhat bigger jump for the Read India program. That jump, you might think, is really not that big. So first, I have to tell you, it's statistically significant.

This is different than that.

But second is, this jump is only due to 13% of kids who actually showed up. So only 13% of kids who couldn't read went to the reading camp. So we can do the same exercise, the same Josh Angrist exercise that we did for Bombay or we did for looking at the effect of the balsakhi, to look at the effect of Read India.

So we divide this little bar by 13%. What is it going to do to my little bar. It's going to make it look much bigger. So this is what we have here. And that's what we find.

So these are kids who couldn't read letters at baseline. And now I'm adding to the control group the little bar that have now become bigger. And we get to exactly 100%. So this suggests that 100% of the kids who actually attended the camp are able to read letters at end-line.

So this is a program which is trying to get the kids to read. And it doesn't get them to read full paragraphs, but it put them one level up. And you can do the same exercise. Kids who were at letter levels are able to read paragraphs and kids who were at paragraph level are able to read stories.

So this is a program that, as a program, is tremendously effective. But there were still some issues, which is that very few kids attended the camp. Why did only 13% of kids attend the camp? And that is something that was a puzzle for us.

Next time, we'll spend much more time on why did only 13% of kids come to the camp. Therefore, the overall effect was low. Next time, we'll try to understand that.

But today, I'll tell you what happened when we saw this result, which is we saw an effective pedagogy, where, if I managed to grab Ben by the collar and try to inculcate him how to read with Pratham technique, I can do it. He can learn letters.

But on the other hand, if it's left to volunteers and left to the effort of the parents, we only managed to get 13% of the kids. So the next step was is it possible to integrate this within the school system? Because enrollment is already high. Students are already a captive audience in the school. If we could use the teachers to do this,

then that would work better.

So off we went. Pratham got a bunch of money, by the Gates and Hewlett Foundations, to expand the program in about 100 districts. That's how they reach 38 million kids today. But they felt that they should try and work through the states.

And so here the trade-off was, well, if it does work, if we can make it work with regular teachers, then we will have many more kids reached. So that's the advantage. The 13% will become 80%.

On the other hand, the treatment effect might go from this very large treatment effect we have to a much lower treatment effect if the teachers are not willing or able to carry out the program. So that's the question.

So the next experiment, we worked in two states, Bihar and Uttarakhand. In the meantime, UP got cut in two. And Uttarakhand is one of the big, mountainous states, a former part of UP, very beautiful place with mountains and stuff.

So in Bihar, we contrasted four models in different state, a summer camp taught by government teachers, trained teachers to implement Pratham as part of their regular teaching, trained volunteers roughly like the Jaunpur model, and distribute only the material. Can you just distribute the material, the textbooks, and get an effect from that? That would be the best, because that's very cheap.

In Uttarakhand, we tried this same teacher training as in Bihar and the volunteers. But the volunteers were put in the school, with the idea that, maybe, in this way, the kids would come. And what did we find?

What we found is that, again, when you have the volunteer intervention, that again works extremely well, with effects similar to what we had in Jaunpur. So what we found in Jaunpur, we find the same thing in Bihar. It worked in the same way, with not very many kids going, but the kids who went really benefited.

And what is interesting here is that it's not only basic reading level any more. It also covers more advanced skills. And the gains were felt at all levels. So the same

approach of trying to teach what the kids don't know, it works at reading, but it also works at more advanced levels.

So great. We had a big effect on learning. We were happy. However, the teacher intervention had very, very little effect. So training the teachers had very, very little effect. There is one effect on Hindi written tests. So if you squint, you find a little bit of an effect of the teachers but, basically, not much.

And moreover, when you put the volunteers in the schools, the volunteers had no effect. In fact, what we found is when you put the volunteers in schools, the teacher is absent more. So they just stay home more. And the volunteers start teaching the kids. And then it has no impact.

So what do we conclude? Is it that the teachers are just horrible, and they can absolutely, never teach anything to the kids? Well, we don't think that's the case, because of the summer camp. In the summer camp program, it's the teachers who were trained to teach. A teacher had to volunteer, and they were paid extra to teach summer camps.

But they were still teachers. And they were trained. And they did it. And the summer camp effect was as large as the volunteer treatment. It means that teachers are able to teach kids to read if they want to. So the problem is that it's not that they can't. It's that usually they choose not to.

And why do they choose not to? That's what we are going to see next time. Where are we? We know that we can improve the quality of education. Because you can take a high school graduate, train them for two weeks, and they can do it.

And the puzzle is that, why is it not taken up more? Why is it not taken up more by the school system? When the teacher's are trained, it doesn't work. Why is not taken more by teachers themselves? Why is it not taken more by parents, who are not sending their kids to the reading camp when they have a chance? I mean some do but not that many.

And finally, they are private schools. A lot of these kids are going to private schools.

The private schools are somewhat more effective than the public schools but not tremendously more. So why aren't the private schools using those techniques?

After all, nothing stops them. They can fire the teacher if they don't do what they do. They are more flexible. So why aren't the private school adopting that?

So that's kind of where we are, which is there is an effective methodology, of very effective, you know. You can get everybody to read at least one more level, from where they were, in three months of a volunteer. We can increase their skills by one standard deviation.

And yet this is not taken up. So this is not the technology. The technology exists. But it's not adopted, neither at the individual level nor at the system level. So we'll see where we are next time. And we'll conclude with you.

**AUDIENCE:** How would you [INAUDIBLE]?

**PROFESSOR:** Teachers are very well educated. Teachers have a BA in teaching plus a discipline. And it's a pretty competitive job to get. And it's very well paid. But it's very well paid at entry, and then you don't increase much. And you can never, ever be fired. So it's a very desirable job.