October 27, 2007

Impact evaluation of Pratham's "Learning to Read" and "Reading to Learn" interventions

Abhijit Banerjee, Rukmini Banerji, Madhav Chavan, Esther Duflo, Michael Walton

J-PAL South Asia at IFMR, and Pratham Resource Center

This note describes a collaborative research program over more than three years (including two and half years of field research) between the Abdul Latif Jameel Poverty Action Lab South Asia at the Institute for Financial Management and Research (J-PAL South Asia, for short) and Pratham to evaluate the impact on reading and maths test achievements of two of Pratham's programs: "Learning to Read" and "Reading to Learn". These programs lie at the centre of a nation-wide effort called Read India, whose objective is to teach all Indian children to read and do basic maths in a three-year period.

At the end of this period, four randomized evaluations will have been conducted in four Indian States, we should have a good understanding of the causal impact of Pratham's Learning to Read and Reading to Learn interventions in different contexts and with different delivery mechanisms, results will have helped shape specific design choices by Pratham (and other actors in the education sector) on the roll-out of reading and maths programs, and at least four young Indian and two international researchers will be fully trained and able to conduct impact evaluations with minimal supervision.

This note describes the context, key questions, the design of the evaluation and plans for implementation. Some of the interventions will only be fully specified in the process of implementation, and the plans here will be periodically updated.

In October 2007, the Hewlett Foundation approved a preparation grant for the initial process of preparation of survey material and organization of the research process with Pratham and State Education services. This paper includes both the preparation and the implementation of the research in order to have an integrated account in one place.

Context

India's deficit in basic education has long been recognized. The PROBE report of the late 1990s documented the scale of the problem and explored the underlying influences. It emphasized the problems of a dysfunctional education system and the influence of unequal social and cultural structures, across gender, caste and class. Since then here has been substantial progress in enrollments—including of girls and poorer groups. But problems of low attendance and low quality remain pervasive. Widespread teacher

absenteeism is only the most dramatic indicator. Once appointed, teachers face little or no incentive to perform. They also typically feel disempowered and demotivated.

Since 2005, Pratham's Annual Status of Education Report (ASER) has provided a quantitative picture of both attendance and learning achievements, based on household visits and testing of children in a representative sample of households throughout India. In rural India less than 7 percent of children aged 6-14 were out of school in 2006, but there were large deficits in basic capabilities in reading and maths, with about a third of children in school in Standards 1-5 falling below the most basic of expectations on learning achievements (Table 1). There is also substantial variation across states. Since these basic learning capabilities are fundamental to all further levels of learning, these numbers suggest a large fraction of Indian children will enter the workforce with inadequate capabilities for all but the most unskilled of work. This is particularly troubling in a period in which India will be going through a major structural transformation in the pattern of production and employment.

Table 1. Indicators of learning deficits in rural India in 2006—national averages

	Average	Range across states
% children aged 6-14 out of school	**6.6**	**0.4 - 12.8**
% in Std 1-2 who cannot read beyond words	**27.0**	**1.4 – 45.6**
% in Std 1-2 who cannot recognize numbers	**40.7**	**6.1 - 61.1**
% in Std 3-5 who cannot read a Level 1 text	**34.1**	**8.3 – 66.6**
% in Std 3-5 who cannot do subtraction	**34.9**	**6.9 – 54.1**

Source: Annual Status of Education Report (Rural) 2006. Pratham.

There have been a variety of responses to this malaise. Some households have been leaving the public schooling system, with rapid growth in private schooling, even in poor rural areas. Private schools typically pay lower wages to teachers, but achieve somewhat better learning outcomes. According to ASER 2006 (rural) some 20 percent of the rural school population in the 7-14 age range was in private schools. But that still leaves almost 80 percent of kids in public schooling. The Federal Government has a major program (called Sarva Shiksha Abhiyan) to provide additional resources and support to state-level education systems. Many State governments have also increased efforts on basic schooling. The most striking example is the campaign by the Madhya Pradesh government in 2005/06. This was associated with a remarkable 30 percentage point increase in the measured proportion of children able to read alphabets or more between ASER 2005 and 2006. Other initiatives have focused on the use of "parateachers" in school, volunteers out of school, and attempts to get greater community involvement to put local pressure on school attendance by children and teaching performance.

Pratham's has been developing new reading materials that are designed to be more accessible and effective in current conditions in India's education system, whether these are used in school or with volunteers outside school. The focus of existing interventions has been on learning basic reading and maths skills—these are termed "Learning to Read" or L2R. Pratham is currently working on developing a range of additional materials designed to provide core skills for students in elementary education who can already read—these are termed "Reading to Learn" or R2L. The implementation of the first phase of this program in 100 districts is being supported by a separate grant from the Hewlett Foundation; work in about 200 other districts is being financed from other sources.

Pratham's approach involves specific interventions within the existing education system. It does not seek to effect systemic reforms of teaching incentives (that are likely to be infeasible in the short run.) The intervention could however still be effective through several channels: if the teaching material can work even with low-quality teaching; if the availability of better materials and learning results increases teachers' intrinsic motivation; or if the volunteer model is effective either as a supplement or complement to in-school learning.

Research questions

Overview

The research team has already conducted an initial evaluation of the volunteer version of the Learning to Read program in Jaunpur District in Uttar Pradesh. In this program, the community is first mobilized via a participatory exercise of testing children for their reading skills. The low level of achievement of the children is then discussed in the village, and Pratham offers to train some volunteers to hold reading camps. Reading camps are then held in the village, and children who cannot read attend.

Our findings from the existing experiment suggest that the program is effective for children who attend the camps. We estimate that the average child who could not read anything at baseline, and attended the camp, would be 60% more likely to read at the letter level after a year than a comparable child who did not attend the camp. The average child who could read at either the word to paragraph level is 61% more likely to read at the paragraph level, and 44% more likely to read at the story level. To put these results in perspective, in the control villages, 43% of the children who were reading nothing at baseline read at least at the letter level by the endline. If we add the treatment effect of the read program to this "natural improvement", it would imply that an average child unable to read at baseline who attends the read program is then able to recognize letter with 100% probability, and an average child who could read only letters would be able to recognize words with 100% probability.

The main downside of the volunteer-based program is that the number of children who attend the camp is limited, in part because of the limited number of volunteers that can be found and trained, in part because parents need to be mobilized to send them: only 8% of the village's children attend the camp (and 13% of those who could read either nothing at all or only letters). For this reason, Pratham is now interested in assessing alternative models with broader reach.

Public sector schools would be a natural place to reach a larger number of children, and teachers, who are already paid to teach, would seem to be natural people to train to implement the intervention. Pratham is thus planning to make the teachers in the state education systems key agents of their program. This is at the centre of Pratham's Read India programme, whose objective is to have all Indian children with basic reading and maths skills in two school years. The upside is clearly the number of children who can be reached. The downside is that teachers' level of motivation and incentives to run the program well is not clear, given the current incentive structure in the education system. It may also depend on the capacity of the States to monitor and motivate teachers.

The key questions that arise when implementing the Learning to Read (L2R) full scale are thus the following:

- What is more effective in getting children to read and do basic maths: running the programs through volunteers or through teachers?
- For the teaching option, what is the impact of just providing the L2R materials, compared with combining provision of materials with tailored teacher training and monitoring?
- Is a volunteer program a useful *complement* to a school-based programme, either as a form of remedial education for kids attending school, or to reach out-of-school kids?
- Do outcomes depend on the capacity of the State government and pre-existing schooling and socio-economic conditions?

A further important question posed by the program is, what next? Do children who have learnt to read retain the skill when this basic learning phase is completed, especially if there is no follow up? Assuming that that they do retain the skills, how do they use this skill to get more out of the school system? Can that process of further skill-building be reinforced?

As noted above, Pratham's "Reading to Learn" (R2L) program is being designed with interventions designed to ensure that children make use of their reading skills to learn other materials in school. Since Pratham has mainly concentrated on the L2R program until now, the specific interventions in this program are not completely defined yet—see below for initial ideas. It is anticipated that they will be ready in time to be incorporated into the evaluation described here.

The same set of questions over delivery mechanisms apply to R2L, since the effectiveness of alternatives could differ at the next learning level. In addition we are interested in longer-term impacts of L2R in the absence of appropriate follow-up interventions, i.e. without R2L?

Finally, Pratham is also interested in experimenting with the use of a school newspaper to provide a continuing source of reading material. This would involve partnering with commercial newspapers that have reach into villages, providing reading material that would be distributed at cost, using existing newspaper networks.

Evaluation strategy

Overview.

The evaluation will provide an assessment of causal impacts of alternative ways of delivering L2R and R2L interventions. It will be integrated into the rollout of Read India. The evaluation itself will be undertaken around the second and third school years of Read India, that is 2008/09 and 2009/10. It will be aligned with the rhythm of the school year, inclusive of time to prepare and to assess interim results. The remainder of the 2007/08 school year will be used for preparation and piloting. Thus the work will both provide a rigorous evaluation of Read India interventions, and inputs to the evolving strategy of Pratham, and potentially other actors in the education sector.

To explore the potential heterogeneity in effects in different contexts, the evaluations will be undertaken in a total of four districts in four States, with different educational conditions, socio-political contexts and administrative capabilities. In the first evaluation period (the second year of Read India) it is planned that the districts will be in Bihar and Maharashtra (subject to final discussions with the various actors involved, especially in the State Governments). Bihar is a poor State, with high education deficits, historically weak institutions and severe distributional conflicts, across class and caste lines. However, the current government is undertaking a variety of measures designed to promote development, and has a substantial education focus. Maharashtra is relatively well-off and with relatively strong administrative capacities. Education deficits are lower, but there is significant variation across districts, and in almost all there is a question of the reaching the most deprived group of children, even for basic reading. There is also particular interest in the state in going beyond basic skills.

Table 2. Indicators of learning deficits in rural India in 2006—Bihar and Maharashtra

	Bihar		Maharashtra	
	Average	Range across districts	Average	Range across districts
% children aged 6-14 out of school	**12.8**	**3.1 – 30.1**	**3.8**	**0.3 – 13.8**
% in Std 1-2 who cannot read beyond words	**29.9**	**7.5 – 48.7**	**12.4**	**0.0 – 26.0**
% in Std 1-2 who cannot recognize numbers	**44.0**	**20.2 – 65.4**	**26.4**	**3.6 – 45.8**
% in Std 3-5 who cannot read a Level 1 text	**29.4**	**15.9 – 43.3**	**20.4**	**8.8 – 66.0**
% in Std 3-5 who cannot do subtraction	**27.6**	**14.5 – 43.0**	**29.6**	**9.1 – 63.2**

Source: Annual Status of Education Report (Rural) 2006. Pratham

The results of the first evaluation period will be used to select the most effective L2R and R2L interventions. These will be universalized in the first two districts, allowing various combinations of two-year effects to be assessed. They will also be introduced in two other districts, in other States, that will have by then have experienced the national roll-out of L2R. This will allow a tighter focus on the net effects of R2L, in addition to providing more information on any heterogeneity in effects across different contexts.

Period 1: Preparation and Evaluation of Alternative Delivery Mechanisms for L2R and R2L in two districts (October 2007-March 2009)

In this period we will conduct two randomized impact evaluations in the initial two districts of both L2R and R2L. Preparation will involve both definition of the interventions, especially for R2L, and setting up of the baseline for the evaluation.

L2R. The educational intervention for L2R has been largely defined: it involves reading and maths material, specially developed by Pratham, in combination with an educator, either in the form of a teacher or unpaid volunteer.

R2L. Pratham is designing and piloting various Reading to Learn (R2L) interventions and this will continue until early 2008. The focus will be on creativity and experimentation in designing interventions, and there will not be any formal impact evaluation of these piloting activities.

The current plan is to design R2L around a series of activities that will require writing, talking, and "doing". Current plans are to design learning tools that focus on four categories of learning:

- Reading to solve (solve problems, puzzles),
- Reading to know (gather information)
- Reading to think (involving a passage that asks you to think about something)
- Reading to do (involving practical experiments)

Pratham will develop training materials, such as sets of cards, in each of these categories.

Alternative delivery mechanisms. The L2R and R2L educational interventions will then be rolled out under alternative delivery mechanisms. This is the heart of the challenge in India, given the problems around the interface between educator and child, flowing from the dysfunctional education system and influences of an unequal social context. For both L2R and R2L, delivery can be by a teacher or volunteer, a combination of both, and with or without specialized training and monitoring. To answer the questions posed above, the following combinations are of practical interest.

(1) Teacher with materials and training/monitoring

The first school-based intervention will involve providing teachers with L2R and R2L teaching materials, a training session on how to use the materials, and monitoring and support by Pratham field workers. The training session will be incorporated into the standard State-financed annual teacher training. The monitoring will be conducted by Pratham field workers at a level that is potentially scaleable at the State and national level.

(2) Teacher with materials and no training/monitoring

It may be that the provision of materials alone is sufficient to increase learning outcomes. It is important to know this, since providing materials has substantially lower administrative requirements than specialized training and monitoring. So a second treatment will involve handing out the materials to teachers, again in the context of the annual State-financed teacher training, but without any additional training and ongoing support from Pratham.

(3) Volunteer with materials and training/monitoring

School-based systems may be ineffective, owing to the incentive and social issues noted above. So we are still interested in comparing the teaching intervention with a volunteer model. The third intervention thus involves unpaid volunteers, and is a modification of the one assessed in Jaunpur, Uttar Pradesh. In UP the volunteer model was embedded within a series of interactions with communities involving information and participation. The current evaluation will involve a streamlined process, with Pratham seeking unpaid volunteers and selecting about one per village to work with children who cannot read. It is again judged to be scaleable. The form of the intervention is still being specified: it is expected to involve both: (i) providing additional, remedial, reading support to in-school students; and (ii) identifying out-of-school students and teaching them to read. These are expected to occur both in the summer period of April-June 2008 and over the school

years. There may in addition be activity designed to (iii) motivate parents to send out-of-school kids to school. The first two of these would involve Pratham's standard L2R materials. Volunteers will be trained, monitored and supported by Pratham field staff.

(4) Teacher and volunteer and training/monitoring

The fourth intervention will involve a combination of (1) and (3), with both in-school and volunteer-based delivery, use of the teaching materials, and with Pratham training, monitoring and support. This is the most administratively intensive—though there are economies of scope for Pratham field staff working with teachers and volunteers in the same village. We are interested in exploring the complementarities between the teaching and volunteer model, and the hypothesis that this mix may be the most effective both at reaching large numbers of children and catching those who fail to learn from school, whether they attend or not.

(5) Comparison group with no intervention

Finally there will be a control group that has no intervention related to Read India. This will, however, be subject to other influences taking place in the district, including any State-wide efforts to promote basic education. The particular nature of any State-based initiatives will be documented as part of the process evaluation, so that we have a good understanding of the nature of the control, as well as any "contamination" from treatments, if teachers in the control learn about and obtain the teahing materials, for example. The effect of the above treatments will thus be *in addition to* such State-wide initiatives, as well as to other influences affecting learning of children in the districts.

Volunteer-based interventions are expected to start around April 2008 and school-based interventions around July 2008.

Testing. Testing will be both household- and school-based. The household-based testing is to ensure that we have an unbiased assessment of the effects on *all* of the target population, that is all children in a village. While some interventions are expected to have impacts within the government schools, the volunteer-based interventions are designed to reach out-of-school kids. In addition, there can be spillover, or external, effects from the school-based interventions, for example if other children switch into the government schools (from out-of-school or private schools), or if there are reading or maths benefits on siblings.

The school-based testing will be undertaken both in the school itself, and through finding kids not at school on the testing day from the enrollment information. This will use a sample drawn from the enrollment register. School-testing has the disadvantage of being only for kids enrolled (and reachable) at the government schools. However, school-based testing is relatively low cost and has the advantage of replicating what governments typically track—that is performance in schools—and so is likely to be of administrative interest to the State government. It will also allow construction of tables of progress by

grade that has been a politically effective feature of the ASERs. Finally, in combination with the household-based testing, it will allow us to form a direct assessment of biases in school-based testing, both in relation to the portion of children in the household-based test attending government schools, and to the overall child population. The last will provide an estimate of any externalities of the program—through comparison of the treatment effects of the in-school population with the village-wide population.

The household-based sample of children will be tested in a baseline conducted as close to the end of the 2007/08 school year as possible, around March 2008, in order to capture the status of children (in and out of school) at the end of the preceding educational period. There will then be an "end-line" to assess the impact of the interventions (actually a "mid-line" for the two-year intervention—see below) that will be conducted around December-January 2008/09. The reason for this early timing, before the end of the school year, is to provide sufficient time to undertake an initial statistical analysis, so that results can feed into planning for the second period, that will be conducted in the March-June period.

The school-based sample of children will be tested at the beginning of the school year, in July 2008, in December-January 2008/09 (coinciding with the household-based sample) and again in July 2009. The purpose of the additional test is to provide a direct estimate of depreciation in learning attainment, that is of increasing concern amongst educators. It would, however, but too costly to add yet another test in households..

Survey instruments. The evaluation will use a range of survey instruments to support comparison of the effect of each treatment and interpretation of results. This will include the following:

(i) For all children, we will use Pratham's own tests. This will involve the existing test for reading and maths, used on a large scale in the ASERs, that is a rapid, practical test on an ordinal scale, and takes 10-15 minutes per child. This is particularly appropriate for testing of L2R and will also us to link our results to the ASER results at the local and national level. It will be supplemented by tests related to learning related to R2L, that are still under development. These tests will be undertaken for all children in both the household and school-based samples.

(ii) For the household-based sample of children, the Pratham test will be complemented by international tests of cognitive ability and achievement. We are obtaining advice from international educators on this, for example from Harvard's Graduate School of Education. The purpose of this is to provide some validation of Pratham's core survey instrument and explore any influence of cognitive abilities (though these may be hard to separate from achievement.) For some parts of the test we may be able to put achievements in international context, but educators warn that this is problematic for reading, owing to differences in languages; it is less problematic for maths. This will use instruments appropriate to the full range of capabilities expected amongst primary school

children. We will review, adapt and pilot alternative survey tools in the October-December 2007 period. Educators who have reviewed Pratham's test consider it be a sound instruments. However, if we find substantial problems of validity we will feed this back to Pratham (and we will also have the second achievement test for our analysis.) These achievement and cognitive tests will be complemented by tests of child self-confidence to investigate the potential two-way relationship between confidence and learning.

(iii) Additional household information will be collected on parental characteristics (including educational attainment and current literacy) and other socio-economic conditions, including measures of assets, social position (including caste). It is also planned to include a cognitive test on mothers, and questions on home-based activities or interactions that could affect learning. We are interested in this information because socio-economic status, mother's literacy and social distance (with teachers) have all been documented to be influences on learning—and often of greater influence than school conditions. In the analysis this will both allow us to interpret interaction effects between the treatments and household characteristics (are the treatments more or less effective for groups of lower socio-economic status, for instance?), and to provide greater precision in the estimates of average treatment effects.

(iv) Surveys of school conditions, of headmasters, teachers and observational studies of teaching will be conducted at base and end-line. School conditions may directly influence the effectiveness of the treatment, or proxy the underlying quality of teaching; this will include potential influences on school governance, such as the presence and activity of village education committees. We are also interested in surveying teachers. This is for two reasons: first, to explore whether there are interactions with measures of teacher characteristics (that can be linked to outcomes for children through averages for schools); second, to explore the hypothesis that teacher motivation can be influenced by more effective teaching. While observational studies may influence teaching behavior, educators argue that they do provide some information, and they will help us in the interpretation of what is going in the classroom affecting any effects that we find. In similar vein volunteers will be surveyed and observed.

(v) Focus group discussions with groups of parents, children and teachers. This will provide a complementary source of information on how the different groups see the learning process and the quality of schooling. It may provide hypotheses to be explored in the quantitative analysis, to the extent that this matches variables measure in the survey work. And it could lead to modifications of the survey for the second year.

Of these survey instruments, the basic Pratham test has been used in the ASERs, and the socio-economic household survey has been developed from the survey used in the Jaunpur survey. Others are being developed, adapted from international models and Indian practices, and piloted in the preparation period, from October to around January.

Draft survey instruments for households, teachers, headmasters, school conditions are available on request.

There will also be a full documentation of process for the interventions in the evaluation, involving descriptions of the interventions and control conditions throughout the period. This will include discreet low-key monitoring of school attendance by teachers and students, undertaken by village residents ("monitors") hired under the project, and managed by the field-based research associates.

Costing. All four delivery mechanisms will be fully costed *ex post* with respect to both financial costs and administrative burdens, for both the State and Pratham so that comparative cost-effectiveness analysis can be performed.

Randomization. Randomization will be undertaken within one district in each State at the level of villages, where a village is a defined administrative unit, as well as a social and geographic grouping. One village may have more than one school (and these may be of varying quality, especially where geographic segmentation across castes is reflected in school location.) As discussed above, it is appropriate to use the village as the basic unit of analysis, since we are interested in the total effects on children in an area. And this will also be inclusive of spillover effects, for example on attendance at public schools, on parental awareness, on siblings. All schools will be included in the survey.

The actual process of selection will involve three stages.

(1) *Selection, with Pratham, of a district within each State and 3-5 blocks within the district, to provide a reasonable spread of geographic areas.*

Both district and blocks will be chosen to be not atypical of conditions in the State: we do not need them to be precisely representative, since in the end we are examining conditions within the district and blocks covered by the evaluation—of which State-wide effects are one source of influence. There will also be attention to administrative feasibility, for example in terms of reasonable geographic proximity of blocks, since transportation is a major financial and time cost.

(2) *Randomized allocation of villages within the selected blocks to the four treatments and control.*

The randomization is in principle straightforward, and will be undertaken to allow pairwise comparision between any treatment, and full ranking of all treatments and control. Randomization for the volunteer model will be based around Pratham's current model for identifying volunteers: this involves sending out calls for volunteers through a variety of networks and then selecting one individual per village. For villages involving a volunteer treatment, the volunteer would get the training, materials and support for Read India. For others, volunteers will be allocated to other activities, such as working with pre-school children, that does not affect this evaluation.

For the *implementation* of treatments this completes the selection process, since treatments are at village level. Where there is more than one school in a village, all schools will be included for the school-based interventions, as noted above. For the *evaluation* we will be surveying a sample of households within treated and control villages. This involves a third stage.

(3) *Selection of households within villages.*

This will be undertaken on the basis of a household listing prepared as part of an initial village census, that will be conducted as part of the survey approach. (If a household listing already exists we will use that, but our understanding is that this is not the case for Bihar or Maharashtra.) Since the listing itself will not be random, given geographic location patterns within villages, the sample will be constructed through selection of every *n*th household in the list (starting at a random place), where *n* is based on the required number of households from the power calculation. This effectively simulates a stratified design to the randomization. A smaller number of replacement households will be chosen in similar fashion (say every *k*th household) to replace those households who are unavailable for survey. All children in the 6-14 age bracket in selected households will be tested, given the low marginal costs of testing once an enumerator is working with a household.

Power calculations. The design needs to have sufficient villages, households and children to be able to detect a minimum level of effects of the treatments at standard levels of statistical significance. For this we need information on the likely variance of effects (where effects are given in terms of *changes* in learning achievements), and the extent of correlation of these effects within villages (since households from within a village are potentially more similar to each other than those in the total population, and a larger sample is required to achieve the same power.) To provide the parameters for the power calculation we used the results of the analysis of the volunteer treatment in Jaunpur, UP, from which we took the variance, intra-cluster correlation and actual observed effect (for the minimum detectable effect). We used as a basis for calculation the probability of children moving up one or more learning levels on the five-point ordinal scale used in Pratham's testing instrument (from letter, to word, to para, to story etc.) The results were:

- percentage of improvers (by 1+ read level) in treatment: 54.6%
- percentage of improvers (by 1+ read level) in control: 47.5%

So the power calculations were based on detecting at least this impact of the treatment.

In the case of Jaunpur, the units were panchayats. Since panchayats in UP are roughly two villages, the calculated intra-cluster correlation of 0.0197 may underestimate intra-village correlation in Bihar and Maharashtra. On the basis of a power of 80% and significance level of 5%, this leads to the alternative combinations of number of villages and number of households within a village given in Table 3; this is for each of the two

district-level experiments.[1] We have assumed equal size for the four treatments and control group, since we are equally interested in comparisons between treatments as between treatments and control. The budget calculations are based on 200 villages per district/State and a total household sample size of 6400 for each district. However, a final decision on the balance between numbers of villages and number of households per village will be made in discussion with Pratham on the basis of implementation costs.

Table 3. Required number of villages and households per village from power calculations (selected combinations)

Number of villages	Number of households per village	Total sample size
100	173	17300
125	82	10250
150	54	8100
175	40	7000
200	32	6400
225	27	6075
250	23	5750

For test results there will be approximately twice as many children as households (based on an average of two children aged 6-14 per household, from the Jaunpur analysis). However, we do not plan to reduce the household sample size, leaving this as a comfort margin for the power of the analysis (that should more than offset the smaller size of units than in the Jaunpur data noted above). Thus while we expect significant intra-household correlation across children, this is fully covered by these power calculations.

Villages versus administrative clusters. The alternative to using villages as the unit is to use the administrative clusters of schools used by the State education services, each of which group about 10 schools. Use of administrative clusters as the unit would have would have facilitated allocation of treatments and controls across teacher-based interventions—since there could be some spillover across schools within the same administrative cluster. However, randomizing across clusters would have led to a need to effectively cover the whole district, or have a very large sample size in order to have sufficient power. This would have brought large additional implementation costs, both for the monitoring of the cluster-based treatments and for the survey management (especially because the sample frame for households from each administrative cluster would be spread across several villages.)

[1] The calculation from the Jaunpur data involved one treatment and one control. Since we will be working with four treatments and a control, we multipled the Jaunpur results by 2.5.

Period 2: Extension of successful interventions in initial and new districts (April 2009- March 2010)

The second period of analysis will involve selection and extension of the most successful interventions in the first period, both to explore longer-term effects and test the robustness of results in additional settings. This will draw on an initial analysis both of impacts of the alternative treatments and of their relative cost-effectiveness. This analysis will be conducted in early 2009.

In the districts covered in Period 1, from a research viewpoint it would have been preferable to have kept the control without the intervention for a further school year, in order to get an estimate of two-year effects relative to this comparison group. Pratham's judgement, however, is that this would be ethically unacceptable given the overall roll-out of Read India, and consequently politically impossible—and especially so if the interventions analyzed in the preceding school year prove successful. So for the villages covered in the evaluation districts in the first year, the most effective intervention will be universalized. This may involve modifications to the mix if impacts are heterogeneous across children or the two districts. For example, the volunteer model might be found to be most effective for the L2R intervention for kids who can't read, while the in-school, materials plus monitoring model is most effective for R2L with kids who can read. The most effective intervention may also differ between Bihar and Maharashtra. The selected Period 2 intervention may thus have modifications from the Period 1, and could differ across districts.

The survey structure from Period 1 will be essentially replicated, though now the end-line from Period 1 becomes the baseline for Period 2 (or the mid-line for the overall evaluation). A further end-line survey will be conducted at the beginning of 2010. (As before, the precise timing will be determined with Pratham, depending on whether they are still looking for specific design advice for future years.)

While the Period 2 intervention will have been universalized across villages, the underlying structure of the experiment will still exist, and we will be effectively finding out the consequences of five *combinations* of interventions across two plus years, of which one is a control-plus-intervention. So we will still have rich information on longer-term effects.

As an important additional exercise, the intervention mix determined to be most successful in Period 1 will be introduced into two districts in new States. This is expected to be in districts that will have already received the general roll-out of L2R as part of the Read India program. In these new districts, there will be one treatment and one control group. If we find that the most effective intervention differs between Bihar and Maharashtra, then there would be a case for having different treatments in the two new districts, linking these to the district characteristics.

This part of the evaluation is first and foremost an exploration of the external validity of the results of the experiment in Period 1, applied to two more socio-economic and administrative contexts, and interacting with the generalized experience of L2R.

The structure of the survey work and analysis would essentially replicate the approach detailed in Period 1. This will involve a new sample, with the associated effort on an initial village cense. However, the required sample will be smaller since only one treatment is currently planned. The power calculations will also be updated with the information gathered in Period 1.

While the above plan has been designed as a logical sequence for finding the most effective intervention, there may have to be mid-course adjustments as information comes in.

Potential evaluation of the impact of a children's newspaper

As noted, Pratham is exploring developing a newspaper targeted specifically at young children that could be distributed regularly in schools. The production and delivery would be undertaken by a commercial newspaper (at cost), while the content would be provided by Pratham. The concept is that a regular stream of new reading material will provide additional incentives for reading—and material for teachers. Once the intervention is designed, including funding for its provision, we will also consider integrated this into the evaluation, for example offering it in some of the treatment and control schools within the above plans, probably in Period 2.

Period 3: analysis, documentation and synthesis (April 2010 – December 2010)

The final period will involve analysis, interpretation and preparation of research reports. This is expected to include a range of activities and outputs, targeted at distinct audiences. This will include:

- The Indian education community. Pratham is already significantly networked into the community of practitioners within India. Results of the evaluation will be taken to policy and design debates within this communities through the direct participation of Pratham personnel, backed by preparation of tailored summaries for this community, that present the results, documentation of how the interventions were linked to state education processes, and description and interpretation of the educational processes involved.
- The broader development community in India. A different kind of product will be developed for the non-education specialists within India, targeting outputs such as the Economic and Political Weekly, and other general magazine and press. JPAL South Asia will take the lead on this.
- The academic community. Academic papers will be written on the experiment and results targeting publication in leading economic journals.

- The broader development community outside India. Accessible summaries of the experience will be prepared for the international development community, including the types of brief descriptions now produced by JPAL in Cambridge, MA.

To support the dissemination, resources have been included for consultants the writing of the various reports in readable and accessible form. They would work with the research team.

The Research Team

The evaluation will be led jointly by J-PAL South Asia and the Pratham resource center. The principal investigators will be Abhijit Banerjee (J-PAL, MIT), Rukmini Banerji (Pratham), Madhav Chavan (Pratham), Esther Duflo (J-PAL, MIT) and Michael Walton (KSG, Harvard and the Centre for Policy Research, Delhi). No time cost has been budgeted for Banerjee, Chavan and Duflo, who will work on this project pro-bono.

We expect that Rukmini Banerji and Michael Walton will the primary actors in the implementation of the project. Michael Walton will be located in India for the entire duration of the project, and devote about a quarter of his time to it. In addition to Rukmini Banerji's overall role, Pratham will make additional adjustments in its staffing allocations, at district and block level, to assure the quality and quantity of management of the variety of interventions, that are significantly more complex than its regular roll-out of the Read India program. Lindy Miller, head of JPAL South Asia at IFMR, will help coordinate the administrative side of the project.

The PIs will be supported by two non-Indian research associates and two Indian research associates (that is one foreign and one Indian in each State), in Period 1, plus two additional Indian research associates in Period 2, since the action will then be in four States. In the first period, the non-Indian research associates are masters economics students with sound technical training in the techniques of field experiments and quantitative analysis—but not necessarily experience in field work. The Indian research associates will have an undergraduate or masters in a social science disciplines and have management, personality and language skills that will allow them to take the lead in the field work itself. We expect there to be a process of mutual learning in project implementation.

There may also be specific use of Indian researchers with experience in the complementary "qualitative" work, such as focus groups. There will also be a team of young Indian workers managing the intervention, most of them coming from Pratham. Ten of them have been trained at the J-PAL SOUTH ASIA impact evaluation training to be held in Chennai in July of this year. These evaluations will also be an occasion to train a group of researchers.

Throughout the duration of the project, regular meetings (probably at least every quarter) of all the researchers working in the project will be held. In these meetings, the researchers will exchange news of the evaluation as well as receive further training. With the combination of working on the project, and the training received in these meetings, we expect that after 3 years, all the young researchers associated with the projects would be capable of conducting similar projects on their own.

Allied research activities

In this entire period, Pratham will be implementing the L2R program in a much wider area, as part of the nationwide Read India campaign to have all children with basic reading skills by 2009. There will be a need to do continuous process evaluation there, both in order to monitor what is going on and to understand the nature of the programs as they are run in the rest of India. While this general process evaluation is not directly a part of this plan (or this budget), and will be carried out by a separate team, we expect to work with that team, to make sure that the right metrics get chosen and that the process data can be integrated into our evaluation. (This is in addition to the continuous process evaluation that will be undertaken directly under the project.)

There is interest in two categories of additional analysis on Read India: (a) validation of Pratham's testing results; and (b) an econometric analysis on the relationship between Read India and results.

Validation of Pratham's survey results. Note that there are two distinct issues:

a) Independent checking of Pratham's results *within* their existing approach to testing. This should involves replication of Pratham's survey methodology, in terms of sampling, survey instrument and details of survey approach (since even details such as timing of household visits can make a difference to results).
b) Cross-checking of the survey instrument. This should involve combining Pratham's existing survey instrument with others—including potentially the proposed general test of cognitive abilities that it is proposed to be used in this evaluation project.

The overall verification will not (and should not) be part of the core evaluation described here, but will be undertaken by an independent researcher. This can potentially be included in the budget and channeled through J-PAL SOUTH ASIA, but implementation would be separate.

However, this evaluation will provide two specific contributions to the validation, for the districts in which we will be working. First , since the testing will use (in part) the same instruments as used in ASER at about the same time, at least in December 2007 (for the initial baseline). This will require the ASERs being designed to oversample in the villages/districts that the project is operating in, so a direct comparison can be made.

Second, as discussed above, we will be using international testing instruments that can be used to benchmark Pratham's survey instrument.

Econometric analysis of Read India. The overall Read India and ASER program of work will create extensive data on both the evolution of Read India interventions and district level test results. This can support an econometric analysis of the relationship between Read India-related interventions and test results—controlling also for the range of district and other influences.

Research consortium. Pratham is also planning on building a research consortium, which would be largely delinked from Pratham and serve as an independent monitoring evaluation cell for Pratham's work, but potentially also work with other organizations on education and related issues. The ASER surveys, for example, would be carried out under its aegis. This is a separate activity, and our work will not be directly involved with this bigger project, though there may be links in the future. The investigators trained under this project, for example, could play an important role.

Budget

A summary budget is given within the proposal, and the detailed budget calculations are attached as an Excel worksheet. Note that the budget does not cover the costs of any large scale evaluation of Pratham's overall work under this program and does not cover either ASER or process evaluation outside the budget areas. It also does not cover the cost of running the programs in the intervention areas (or piloting) as this is covered elsewhere, including from Hewlett's other grant to Pratham.

The budget is based on required sample sizes detailed above, that draw on our previous experience and on power calculations from the Jaunpur, UP results. Actual survey costs will be based on tenders submitted by reputable survey companies. Given these, and other, uncertainties, a contingency of 10% has been included.

The budget also includes supervision and coordination costs for the overall project as well as the expected 6 research associates and junior village-level workers to supervise the field work. It includes the cost of capacity building and training. Domestic travel is budgeted for regular meetings of the evaluation team and site visits of the actors directly involved in the evaluation. International travel is budgeted for the US-based researchers and international research associates.