INSTITUTE OF DEVELOPING ECONOMIES

IDE Discussion Papers are preliminary materials circulated
to stimulate discussions and critical comments

IDE DISCUSSION PAPER No. 551

The Impact of a Computer Based Adult Literacy Program on Literacy and Numeracy: Evidence from India

Ashwini Deshpande, Alain Desrochers,
Christopher Ksoll and Abu S. Shonchoy*

January 2016

Abstract

With over 700 million illiterate adults in the world, many governments have implemented adult literacy programs across the world, although typically with low rates of success partly because the quality of teaching is low. One solution may lie in the standardization of teaching provided by computer-aided instruction. We present the first rigorous evidence of the effectiveness of a computer-based adult literacy program. A randomized control trial study of TARA Akshar Plus, an Indian adult literacy program, was implemented in the state of Uttar Pradesh in India. We find large, significant impacts of this computer-aided program on literacy and numeracy outcomes. We compare the improvement in learning to that of other traditional adult literacy programs and conclude that TARA Akshar Plus is effective in increasing literacy and numeracy for illiterate adult women.

Keywords: *Adult Literacy Program; Adult Education; ICT; Women; India*
JEL classification: *O12*

Research Fellow, Micro-economic Studies Group, Development Studies Center, IDE (Abu_Shonchoy@ide.go.jp)

The Institute of Developing Economies (IDE) is a semigovernmental, nonpartisan, nonprofit research institute, founded in 1958. The Institute merged with the Japan External Trade Organization (JETRO) on July 1, 1998. The Institute conducts basic and comprehensive studies on economic and related affairs in all developing countries and regions, including Asia, the Middle East, Africa, Latin America, Oceania, and Eastern Europe.

INSTITUTE OF DEVELOPING ECONOMIES (IDE), JETRO
3-2-2, WAKABA, MIHAMA-KU, CHIBA-SHI
CHIBA 261-8545, JAPAN

The Impact of a Computer Based Adult Literacy Program on Literacy and Numeracy: Evidence from India[1]

Ashwini Deshpande
Delhi School of Economics, University of Delhi

Alain Desrochers
University of Ottawa

Christopher Ksoll
University of Ottawa

Abu S. Shonchoy
Institute of Developing Economies-JETRO and University of Tokyo

Abstract:

With over 700 million illiterate adults in the world, many governments have implemented adult literacy programs across the world, although typically with low rates of success partly because the quality of teaching is low. One solution may lie in the standardization of teaching provided by computer-aided instruction. We present the first rigorous evidence of the effectiveness of a computer-based adult literacy program. A randomized control trial study of TARA Akshar Plus, an Indian adult literacy program, was implemented in the state of Uttar Pradesh in India. We find large, significant impacts of this computer-aided program on literacy and numeracy outcomes. We compare the improvement in learning to that of other traditional adult literacy programs and conclude that TARA Akshar Plus is effective in increasing literacy and numeracy for illiterate adult women.

JEL: O12

Keywords: Adult Literacy Program; Adult Education; ICT; Women; India

[1] Corresponding Author: parves.shonchoy@gmail.com. We acknowledge the excellent research assistance provided by Sahib Tulsi, Meghna Bohidar, Uday Anand, Wei Yang, Sakshi Bhardwaj, and Vinitha Rachel Varghese. We would also like to thank Alka Srivastava, Ranjit Khosla, Col. MS Ahluwalia, Irfan Khan, and other Development Alternative members for their useful comments and feedback. Shaher Banu Vagh kindly provided us with the comparison of the EGRA/EGMA style literacy tests and ASER tests.

1. Introduction

The overwhelming majority of the world's 774 million illiterate adults live in developing countries (UNESCO, 2012). Adult literacy programs aim to improve the skills, and hence the earning potential and other socio-economic outcomes, of illiterate adults. However, traditional adult literacy programs, typically operated by governments, have been largely ineffective due to low enrollment, high dropout rates, and rapid skill depreciation (Abadzi 1994, 2003; Oxenham *et al.* 2002; and Ortega and Rodríguez 2008). Impact assessments of adult literacy programs measure direct effects, viz. the acquisition of literacy, and indirect or extended effects, such as intra-household sharing or child health outcomes. Recent advances in adult literacy programs have sought to integrate information and communication technology (ICT) into effective teaching methods. ICT could improve the quality and effectiveness of learning with the aid of interactive tools, the use of animation, and the implementation of effective teaching principles.

To the best of our knowledge, there is no rigorous evidence of the effectiveness of computer-based adult literacy programs in delivering high-quality literacy and numeracy teaching. In this paper, we seek to fill this void in the literature by investigating the impact of TARA Akshar Plus (TA+), a systematic computer-based adult literacy and numeracy program conducted in the state of Uttar Pradesh in north India,[2] on the literacy and numeracy of neo-literate adult women. TA+ has been implemented by the Indian NGO Development Alternatives (DA), which claims that the success rate of this program is over 90 percent.[3] Despite over 100,000 participant learners, the program has never been scientifically evaluated. Our study provides the first thorough assessment of TA+ and the first rigorous evidence of the effectiveness of an ICT-based adult literacy program on literacy and numeracy.

We employ the random assignment of 717 illiterate women from 18 villages into two groups: (i) a treatment group that had the possibility to undergo the TA+ program immediately and (ii) a control group that only had the possibility to undergo the program later. Respondents in the treatment and control groups were tested pre- and post-intervention. We present results on the impact of TA+ on learning outcomes by combining the random assignment with the individual-level test results.

[2] TA was developed by UK-based Social Enterprise Readingwise, which uses interactive computer-based learning modules.

[3] http://taraakshar.org/index.php/results/, accessed February 3, 2015

Our main results are as follows. We find that TA+ substantially affects literacy and numeracy in the short run, especially in basic literacy and numeracy skills such as reading letters and words and counting and number recognition. However, the effects on more advanced other dimensions of reading and numeracy, such as reading paragraphs and addition/ subtraction are smaller.

The rest of the paper is organized as follows: section 2 contains a brief review of related literature; section 3 gives the background to the TA+ program; section 4 describes the design of the experiment; section 5 describes the data and presents descriptive statistics; section 6 presents results of the impact analysis and presents a comparison of the learning effects of TA+ to results from other similar programs; and section 7 concludes the paper.

2. A Brief Review of Related Literature

The set of studies that measure the direct or immediate impacts of adult literacy programs, viz. acquisition of literacy and/or numeracy, primarily consists of studies that suffer from some or all of the following problems: very small sample sizes, flawed experimental design (e.g., lack of a comparison group), and poorly designed assessment tools (Carron 1990, Ortega and Rodríguez 2008). An exception is Banerji et al.'s (2013) rigorous evaluation of literacy classes on language and mathematics scores in the states of Bihar and Rajasthan in India.

The set of studies measuring the indirect or extended effects of adult literacy programs includes Blunch and Portner's (2009) assessment of a large-scale Ghanaian adult literacy program on household consumption; Blunch's (2013) assessment of maternal participation in adult literacy programs on child mortality in rural Ghana; and Banerji et al.'s (2013) assessment of the positive impact of maternal literacy on children's mathematics scores in India. These studies find that adult literacy programs can have a positive effect on participants' literacy, though the increases are small in magnitude. A third branch of the literature on adult literacy focuses on outcomes such as inter-household sharing (Maddox 2007), individual earnings (Basu *et al.* 2001), and children's height-for-age (Gibson 2001) mainly through a comparison between literate and illiterate adults; these correlations studies are not necessarily attributable to specific AL programs.

A larger branch of the literature highlights the strong, robust positive correlations between education and several desirable socio-economic outcomes. The positive correlation between

maternal education and child outcomes has been well documented (see, among others, Strauss and Thomas, 1995; Gakidou et al., 2010; Paxson and Schady, 2007; Masset and White, 2003; Garcia & Jacinto, 1988, 1991; Handa, 1999; Thomas 1990; Levin and Haddad, 1994). Female education is also positively correlated with lower fertility; improved health, hygiene, and education; better saving practices, and increased gender equity (Senauer *et al.* 1988; Thomas 1990; Hopkins *et al.* 1994; Strauss and Thomas 1995; Handa 1999; Masset and White 2003; Paxson and Schady 2007; Gakidou *et al.* 2010).

Turning to ICT-based programs, there is generally little evidence on the effectiveness of such technology-based literacy programs. In a pilot study conducted in Niger, Aker, Ksoll, and Lybbert (2012) provided some early evidence that teaching adults to use mobile phones in the context of adult literacy classes can increase math and reading test scores by 25 percent. Chugdar (2014) reported basic cross-tabulations from a survey of 409 illiterate adults in semi-urban locations in Gujarat, India, highlighting the potential role of mobile phones to address adult literacy. The somewhat-related literature on the use of computers in schools reveals mixed findings with respect to their effectiveness. Lei *et al.* (2013) and Mo *et al.* (2013), who examined the "One Laptop One Child" policy in rural Chinese schools, and Linden (2008), Osario and Linden (2009), Angrist and Lavy (2002), and Beuermann *et al.* (2015) found that the introduction of computers in schools have either no or mixed effects on children's learning outcomes (e.g., test scores) or cognitive skills mainly due to the failure to incorporate computers into the educational process. Banerjee *et al.* (2007) and Barrow, Markman, and Rouse (2009) found positive impacts of computers on learning, e.g., algebra scores, among schoolchildren.

3. Background

About one-third of the world's illiterate population lives in India, where illiteracy is still largely a female phenomenon. The female literacy rate—defined as all females aged seven and over who can read and write—is 65 percent, whereas the male literacy rate is over 80 percent. Women constitute 64 percent of the total illiterate population aged seven and above (Census 2011). While India has been successful in raising the primary enrollment rates of boys and girls through programs such as *Sarva Shiksha Abhiyan* and the mid-day meal scheme (raising the net primary enrollment rate to 93 percent in 2011[4]), progress on adult literacy has been comparatively limited (Kapur and Murthi 2011).

[4] World Development Indicators: http://data.worldbank.org/indicator/SE.PRM.NENR, accessed February 4, 2015.

The National Literacy Mission (NLM) was launched in 1988 to increase functional literacy among 15- to 35-year-olds in India. Because these individuals are in the "productive and reproductive period of life," the NLM offers them "a second chance" at functional literacy. The Indian government conducts literacy programs for both their intrinsic and instrumental values. Thus, the NLM defines functional literacy as a composite set of goals: "Self-reliance in the 3 R's [reading, writing, arithmetic]; becoming aware of the causes of deprivation and moving towards the amelioration of their condition by participating in the process of development; skill improvement to improve economic status and general well-being; and imbibing values of national integration, conservation of environment, women's equality, and observance of small family norms, etc." [5,6]

The NLM initially had the aim of increasing literacy rates to 75 percent by 2007. To date, however, only the state of Kerala has been declared "totally literate." This slower than anticipated progress has been attributed to large class sizes, inflexible schedules, poorly designed curricula, and, consequently, low participation. Partially in response to concerns regarding the poor quality of existing programs, the Indian government has extended its approach by supporting e-learning programs, including the TA+ literacy and numeracy program.

TA+ is one of a handful of e-learning programs accredited by the NLM. TA+ currently comprises two modules: TARA Akshar, a 36-day literacy program, and TARA Ganit, a 18-day numeracy program. TA+ began operating in 2005. As of 2015, 150,788 women have participated in its adult literacy classes, with a reported success rate of 95 percent. The TA+ program is funded both by the Indian government and by international organizations. The learners we studied were part of a UNDP-implemented program in the eastern part of Uttar Pradesh (UP). Uttar Pradesh is a north Indian state with high illiteracy rates, where the TA+ program has reached approximately 103,708 learners. [7]

4. Experimental Design

The sample for our experiment came from two phases of the program. During the first phase, 238 women from six villages in the Sant Ravidas Nagar district of UP participated in the baseline survey in September 2013. These women had indicated their interest in participating in the

[5] http://www.nlm.nic.in/nlmgoals_nlm.htm, accessed 2nd February 2015.
[6] As one can see, achievement of all these goals at the same time is not straightforward, and it is not clear whether the program prioritizes any particular component of this multi-faceted set of objectives.
[7] http://taraakshar.org/index.php/results/, accessed September 30, 2015

literacy program, and they were illiterate. Using a public lottery,[8] 139 of these women were chosen for the treatment group and 99 were chosen for the control group during Phase 1.

The literacy program was implemented in October and November of 2013 before the endline study in December 2013. The program during the time of implementation spanned 26 days, and it was divided into three parts. Part 1 involved computer-based instruction for 100 minutes per day. The first 10 days were devoted to learning the Hindi alphabet, and subsequent days of instruction were focused on decoding and writing syllables and words. Each instructional session involved the following activities: a) four minutes of watching a video to facilitate the encoding and recall of characters, syllables, and words. The program heavily uses an idea from psychology that so-called "memory hooks" are an effective tool to learn. In the current context characters are linked with objects that start with that particular character. To exemplify this with an English example: The letter "s" might be shown as a snake curled in the form of an "s"; b) 12 minutes of work with large flash cards with letters, characters, syllables, or words; c) 20 minutes of writing practice; d) 20 minutes of work with small flash cards to facilitate recall of letters, characters, syllables, or words introduced; d) 10 minutes of the learner identifying letters, characters, syllables, or words using the computer software; e) 20 minutes of writing practice; f) 10 minutes of quizzes and practice through peer learning; and g) four minutes for a follow-up video. Part 2 spanned 10 days with 100 minutes of instruction each day. This part of the program was primarily devoted to practice with reading (70 minutes) and writing (30 minutes). Part 3 lasted six days and focused on learning recovery. This part was designed to assist learners who were late in joining the program, missed classes, or struggled with any part of the teaching material. There were also review days interspersed during the course of the program: the sixth, tenth, twentieth, and twenty-fifth days of the program were reserved for this purpose. On the twenty-sixth day, all learners were assessed on their ability to recognize Hindi characters, write words, write phrases and sentences, and apply reading and writing skills beyond the program coverage. We do not use make use of those assessments as they were not conducted by us.

TARA Ganit, the numeracy program, lasted 13 days during implementation, with each instructional session lasting 105 minutes. The learners' numeracy skills were assessed on the last day. Every session involved the following activities: a) 13 minutes on a story; b) 10 minutes of work with big flash cards; c) 12 minutes of writing practice; d) 10 minutes of work with small flash cards recalling numbers introduced; e) 10 minutes on the learners identifying numbers using the computer software; f) 10 minutes of instruction on how to write numbers alphabetically; g) 10 minutes repeating the digits taught; and h) 30 minutes of writing practice. The seventh, ninth, eleventh, and thirteenth days of the program were reserved for revising the

[8] Where a public lottery was not possible, the research team conducted a draw of names through a computer-based algorithm.

numeracy learning. During this numeracy training, the learners were taught multiplication tables up to 10.[9]

While the first phase of the program documented (highly) statistically significant learning effects (see Appendix Tables A1 and A2), the research team felt that a larger study based on more than six villages would have larger external validity. As in the first set of villages, interactions with villagers indicated that some illiterate women were not allowed to participate by their families or did not want to[10]; therefore, the research team decided to quantitatively investigate participation in the program.[11] The research design was extended to a further 12 villages. The second phase of the program was implemented in June, July, and August of 2014 after baseline surveys had been administered. As in the first phase, surveys were also conducted at the end of the program.

Before the administration of the surveys, together with DA, a list of illiterate learners was drafted in each of the villages. Some women immediately indicated that they were not interested in participating, and they were not part of the lottery that selected learners. We therefore exclude these women from the analysis in the current paper. Thus, the current paper includes only women who were interested in participating in the program or did not immediately state that they were not interested. These women were grouped by the hamlet in which the classes would take place, based on DA's suggestion, in recognition of the stated unwillingness of the women to join classes held in other hamlets. In some cases, the lists of women were also separated by *jati* (caste) when DA thought respondents would refuse to join a class with learners from other castes. Each of these lists was then part of a (separate) public lottery in the villages/hamlets to determine assignment to the treatment group. After the treatment group was chosen, the remaining women were assigned a priority number on the waiting list. A number of women who were selected for the treatment group, however, backed out and did participate in the treatment, often because the household did not permit their participation. These women were replaced by contacting women on the waiting list in the order

[9] In Phase 2, Tara Aksar Plus spanned 10 additional days, which were dedicated to periodic tests (not additional days of instruction). However, in all our regressions, we adequately control for this difference between the two phases.

[10] Major reasons for non-participation that respondents stated during the survey were a) not interested in learning at this age (32%), b) having inconvenience or family responsibility at home like having a baby or older person to look after or being pregnant or could not take time out from household chores (16%) and c) TA center is located at an inconvenient location (14%).

[11] The results of this investigation are in Ksoll and Wang (2015).

in which they had been drawn after the treatment group.[12, 13] Any woman on the list created for the lottery who was invited to participate in TA+ was considered a treatment respondent. The remaining women on the waiting list - who were not contacted - constitute the control group.[14] In total, a further 479 women participated in the baseline survey for the second phase.[15] Of these, 264 were part of the treatment group and 215 part of the control group.

4.1 Attrition, Participation, and Non-compliance

Appendix Table A1 shows the number of observations in our sample in baseline and endline separated by phase and round of survey. Of the 238 respondents from the first phase, we interviewed 232 at the endline (with two control and four treatment observations not participating in the endline interview). Of the 479 respondents in the second phase baseline, we interviewed 430 at the endline (with 25 control and 26 treatment observations not participating in the endline interview).[16]

Not all the respondents selected for treatment participated, and a very small number of the respondents in the control group participated in the literacy program. In the first phase, 127 of 135 women in the treatment group participated in the treatment (of which five attended only a few classes). In the control group, 10 women participated in some TA+ lessons, with six attending a few classes. In the second phase, 176 women of the 238 assigned to the treatment group participated in TA+, with 23 attending only a few lessons. Eight women from the control group participated, with two of these attending only a few lessons. In the estimation, we treat women who attended *any* classes as participants.

4.2 Tests

We administered a battery of literacy and numeracy tests to all the women (in the treatment and control groups) before and after the program. The literacy tests were developed, tried, and tested by Pratham[17] based on the model used in the DIBELS (Dynamic Indicators of Basic Early

[12] We follow Card, Ibarrarán, and Villa (2011), who suggest this approach to ensure costly slots in labor market programs are filled while preserving the treatment and control group research design.

[13] In the first set of villages from the second phase, the staff from the NGO did not make as many attempts to recruit the women on the waiting list as in the second set of villages. Empirically, we account for this through the strata/hamlet fixed effects.

[14] To be exact, in some cases TA instructors could not make contact with an individual on the waiting list and so skipped to the next participant. Our information on that is incomplete, we do, however, know who was the last individual contacted on the waiting list. We take the conservative approach of assigning to the treatment group all women ranked earlier in priority than that last individual.

[15] For a few of the hamlets, the number of women was not large enough to support a control group. We drop hamlets when there were fewer than three learners left for the control group.

[16] The second phase was added for purposes of larger external validity even though, as shown in the appendix, a separate analysis of the first phase showed large, significant effects from the program.

[17] http://www.asercentre.org, accessed January 25, 2015.

Literacy Skills) tests (Good, Kaminski, Simmons, & Kane'enui, 2001). The same model was used in the development of the Early Grade Reading Assessments (EGRA) and Early Grade Math Assessments (EGMA), which are extensively used in international literacy and numeracy research. Appendix A contains a more detailed description of the literacy and numeracy measures, though we outline the tasks here.

The literacy tests were timed tasks. The learners were given one minute to read 52 letters (Task 1), 63 syllables (Task 2), 52 words (Task 3), 48 non-words (Task 4), and a 64-word Grade-1 level paragraph (Task 5). On Task 5, examiners marked as correct/wrong only those words read within a minute but allowed the learner to complete the paragraph even after the first minute so the learners could answer the comprehension questions. The test also included a number of "stopping rules" so the test did not progress to more difficult tasks if the learners could not achieve a minimum level on an earlier task.[18]

The mathematics tests were not timed. Respondents were asked to complete eight tasks in Section I: (i) count three objects orally; (ii) recognize single-digit numbers ranging from 0 to 9; (iii) recognize ten randomly selected two-digit numbers; (iv) count objects and circle the correct written number; (v) count objects and *write* the correct number; (vi) fill in the missing digit in two series; (vii) add two to three single and two-digit numbers to each other; and (viii) subtract single and two-digit numbers. Based on the results of these eight tasks, we created an overall math score.[19]

The baseline survey consisted of questions on household composition, asset, and other socio-economic characteristics of the women, which are described in more detail in the following section.

5. Data and Sample Analysis

We present summary statistics for both the treatment and control groups, illustrating that the TA program targets a very disadvantaged population.

[18] The precise stopping rule was as follows: If a learner could not read more than three letters in a minute, the reading test was stopped at Task 1. If the learner could not read a single word in a minute, the reading test was stopped at Task 3. If the learner could not read more than three sentences without any time restrictions, the reading test was stopped at Task 5.

[19] Because piloting suggested extremely low math skills, we also asked students in Phase I to count up to 30 and in steps of 10 to 100 as very first task In the fieldwork, very few people could not count to 30, so we dropped this part for Phase II. We also implemented the counting in steps of 10 to 100 for Phase II, but we do not present these results here because they are not part of the EGRA-type test. In Phase II, we also administered a verbal mathematics test with 15 questions, but do not report these results for the same reason.

The process of selecting participants in the treatment group through a lottery was meant to create two groups with very similar observable and unobservable characteristics. A standard check of whether the randomization procedure was properly conducted involves investigating whether these two groups have similar observable characteristics, e.g., demographic variables and baseline test scores. If one lottery had been held for all potential participants across all villages, then a simple t-test would be appropriate. However, as Duflo *et al.* (2008) note, one should include subgroup indicators (called *strata*) into a regression.[20] This is also true for baseline balance tests. In our case, the strata fixed effects indicate the hamlet where the respondent lived or went to attend class.

[Table 1a and 1b about here]

Table 1a reports the summary statistics of individual- and household-level characteristics for the treatment and control groups as well as a whether the difference between the two groups is significant. Of the women, 93 percent had never attended school. Four percent belonged to the upper/general caste, 51 percent to the Other Backward Classes category (OBCs), and the remaining belonged to the Scheduled Caste and Scheduled Tribe (SC/ST) category.[21] The average age was 34 years. Approximately 90 percent of the women were married, 5.5 percent were unmarried, and 4 percent were widows. The average number of children in the household was 3.3. For an overall idea of household well-being, we computed scores for each household on the Grameen Banks's India Progress-out-of-Poverty Index (PPI).[22] A PPI of 20 corresponds to a roughly 90 percent chance of being under the poverty line using international US$2/day/person poverty lines. The average PPI score of households in our sample was 25. Of the households, 22 percent had a "Below Poverty Line" (BPL) card.[23] The overall housing condition of our sample reflects the PPI index: 26 percent had access to electricity; approximately 57 percent resided in a brick house; only 4 percent had toilet facilities at home;

[20] Technically speaking, whether it is necessary or simply advisable to include strata fixed effects depends on whether the same proportion of participants is chosen in each lottery. Imbens, King, and Ridder (2006) showed that including strata fixed effects for different lotteries generally lowers the estimated standard errors for the estimated coefficients and is thus advisable. When the lottery is conducted for separate subgroups separately and the number of participants chosen for the program as a proportion of all possible participants differs by subgroup, it is necessary to include strata fixed effects to avoid biased results.

[21] Scheduled Castes and Scheduled Tribes are "lower-ranked", stigmatized and marginalized castes (jatis) and tribes, eligible for affirmative action. Other Backward Classes (OBCs) are castes and communities that are not stigmatized like the SCs, but low enough on various socio-economic indicators to warrant additional affirmative action. "General" castes in the popular parlance is used to refer loosely to upper castes (For more detail on the caste system and its economic consequences, see Deshpande 2011).

[22] For more information on the construction of the index, see http://www.progressoutofpoverty.org/ppi-construction.

[23] The BPL card is given to households deemed to be under the official Indian poverty line to identify disadvantaged households requiring government assistance.

38 percent had access to a safe source of drinking water, that is, access to water from private tap water or community well; and 1.5 percent used a clean fueling source for cooking, that is, they cooked with a gas cylinder or electricity. In terms of household assets, only 3 percent possessed any thermoware food heating equipment (e.g., thermos, casserole), 22.5 percent had either a TV or DVD player/VCD player/VCR, and 39 percent reported owning a personal mobile phone.

The third column of Table 1a, which reports an estimate of the difference between the treatment and control groups, shows that the treatment and control groups are balanced on virtually all characteristics.[24] The only difference between the two groups is that the treatment group is 5 percentage points more likely to own a sewing machine, a difference that is statistically significant at the 5-percent level. Similarly, Table 1b provides information on how the samples differed between Phases 1 and 2.

[Table 2 about here]

Table 2 gives an overview of the primary occupations of the women.[25] Approximately 76 percent of the women reported household activities (not working outside the home) as their primary occupation. Classifying women using broad occupational divisions, 18.71 percent of the women were involved in farm activities, 5.1 percent were involved in non-farm activities, and the rest were involved in non-income generating activities.

6. Impact Estimation

6.1 Empirical Framework

We first show that at the baseline the treatment and control groups were also similar, in the outcomes of interest, viz. literacy and numeracy test scores. Columns 1 and 2 of Table 3 report the baseline means of literacy and numeracy outcomes for the treatment and control groups, respectively. Column 3 reports an estimate of the difference between the two. The table shows that the baseline (pre-treatment) difference in test scores between the two groups was not significant, with the exception of the ability to complete subtraction problems, as there is some evidence that the treatment group solved more problems ($p < 0.1$).

[Table 3 about here]

[24] This estimate accounts for stratification and clustering (cf. Empirical Framework section for the specification).

[25] For this multiple response question, we designated the "primary occupation" as the occupation/activity in which a respondent spent most of her time.

We then estimate a simple treatment effect model before implementing instrumental variables, difference-in-difference, and fixed effects estimations. Formally, we first estimate the following regression, where *testscore*$_{i,j}$ is the test score of individual *i* located in village *j*:

$$testscore_{ij} = \alpha + \beta_{ITT} ITT_{ij} + \mu_j + \varepsilon_{ij} \qquad (1)$$

ITT_{ij} is an indicator variable that takes a value of 1 if a respondent was assigned to the treatment group and a value of 0 otherwise. Because all the women did not necessarily comply with this assignment, the variable β_{ITT} captures the "intent-to-treat" effect, i.e., the effect of being assigned to treatment. μ_j is the specific sub-village (usually the hamlet) part of the error term (as mentioned above, our sample is stratified at the sub-village level), and ε_{ij} is the individual-specific part of the error term. All our standard errors account for clustering at the village level.[26]

Next, we include a set of individual- and household-specific control variables that convert equation (1) into the following regression equation:

$$testscore_{ij} = \alpha + \beta_{ITTC} ITT_{ij} + \gamma X_i + \mu_j + \varepsilon_{ij} \qquad (2)$$

The variable β_{ITT} captures the intent-to-treat effect for control variables. These control variables include age, marriage status, household assets, type of water the household has access to, type of cooking fuel used, material of the household dwelling, dummy variables for broad caste groups (SC, ST, OBC), PPI index, and whether the household has a BPL card.[27]
As noted above, not all households complied with the assignment to treatment and control groups. In particular, close to 20 percent of the women assigned to the treatment group did not participate in TA+. Therefore, intent-to-treat will be an underestimate of the program for those who participated in the treatment. We therefore implement an instrumental variables strategy in which participation in TA+ is instrumented by assignment to the treatment group. In the first stage, we regress participation in TA+ on assignment to the treatment group and the other covariates used in the second stage.

$$TA_{ij} = \alpha + \beta_{Stage1} ITT_{ij} + \gamma_{Stage1} X_i + \mu_j + \varepsilon_{ij} \qquad (3a)$$

[26] More recently, Cameron, Gelbach, and Miller (2008) have argued that with a small number of clusters, it is preferable to base the standard errors on a *Wild cluster-bootstrap percentile-t procedure* and impose the null hypothesis as opposed to clustering. When we implement their suggestion with sub-village fixed effects that we need to avoid biased estimates while clustering at the village level (since all groups in a village share a teacher), the standard errors seem to be too *small.* We thus present the more conservative standard errors clustered at the village level.
[27] The list of control variables used in the regressions is shown in Table 1.

The instrumental variables approach then uses the predicted values instead of treatment assignment as an independent variable in the second stage.

$$testscore_{ij} = \alpha_{IV} + \beta_{IV}\widehat{TA}_{ij} + \gamma_{IV}X_i + \mu_j + \varepsilon_{ij} \qquad (3b)$$

The parameter β_{IV} captures the effect of TA+ for those induced into treatment by treatment assignment, and it is our preferred specification.[28] We correct the estimated standard errors for the two-stage procedure.

We implement a number of robustness specifications. The panel structure of the data allows the implementation of a difference-in-differences (DID) IV specification. Pooling the data from the first and second rounds, the equation to be estimated is as follows:

$$testscore_{ijt} = \alpha + \beta_{DDIV}\widehat{TA}_{it} + \delta_{DDIV}post_t + \theta ITT_i + \mu_j + \varepsilon_{ij} \quad (4)$$

where $post_t$ indicates the information from the endline survey implemented post-intervention and ITT_i indicates whether a women was selected for the program in the lottery. β_{DDIV} is the parameter of interest on the participation variable instrumented by $post_t * ITT_i$.

Our last specification deals with fixed individual heterogeneity. We estimate the following fixed-effects instrumental variables specification where we allow for an individual specific fixed effect μ_i.

$$testscore_{ijt} = \alpha + \beta_{FEIV}\widehat{TA}_{it} + \delta_{FEIV}post_t + \mu_i + \varepsilon_{ij} \qquad (5)$$

6.2 Results

6.2.1 Reading Scores

Table 4 shows the main results of the impact of TA+ on literacy measures. The different columns contain results on the different measures of literacy in the order in which they were administered (which, with the exception of non-words, corresponds roughly to the level of difficulty). The first two rows contain the mean and standard deviation of the control group using the endline data. The first panel contains the estimation results for the simple difference

[28] Given that there is very little non-compliance on the part of the control group, this is very close to the treatment effect on the treated (ToT); see Angrist and Pishke (2010).

estimate corresponding to equation (1). We find that the TA+ program improved literacy significantly across all dimensions of literacy. Compared with the control group, those assigned to the treatment group could read almost 9.4 more letters on average, 4.4 more syllables, and between two and three more words depending on the difficulty level. These estimates are all significant at the five-percent level. The results are robust to the inclusion of household and individual control variables, as seen in the second panel of Table 4.

The third panel presents the results of the instrumental variables specification. As expected, the coefficients are larger, reflecting the fact that the simple difference estimate does not consider that some observations assigned to the treatment group did not participate. In particular, participation in TA+ increases the number of letters read in one minute by an average of 12.4 letters (compared with a control group mean of 3.4), the number of syllables by 5.7 (over a control group mean of 1.7), and the number of words per minute read by between 2 and 3.7 depending on the difficulty of the text selected.

[Table 4 about here]

6.2.2 Numeracy Score

Table 5 shows the results of TA+ on numeracy outcomes. Similar to Table 4, the first two rows contain the mean and standard deviation of the control group. With the exception of counting objects (where the mean at the baseline was already very high), TA+ has significant impacts on all mathematical tasks. Focusing on the absolute levels of these effects, learners in TA+ recognize, on average, about seven out of 10 written single-digit numbers and two out of 10 double-digit numbers, compared with three and one, respectively, for the control group. These results are robust to the inclusion of household and individual control variables and instrumental variables specifications as reported in the second and third panels of Table 5, respectively.

[Table 5 about here]

6.2.3 Robustness Checks

Tables 6 and 7 report the results from the DID and instrumental variables fixed effects specifications with covariates. The results show that the estimated impacts are robust across a broad range of specifications and very similar to our previous estimate findings reported in Tables 4 and 5.

[Tables 6 and 7 are here]

We further tested the robustness of our findings by running phase-specific regressions using the specification used in the third panels of Tables 4 and 5. Appendix Tables A2 and A3 report the findings of these regressions on literacy and numeracy, respectively. The point estimates of the treatment effect on literacy outcomes are always positive and significantly different from zero with two exceptions: there is a positive but insignificant effect of TA+ on the number of words per minute read at the grade 1 and grade 2 levels during Phase 2. However, for the numeracy test, both phase-specific regressions provide evidence of significant program effects except for the two-digit number identification in Phase 2.

We notice that the results during the first phase are much larger than during the second phase, and the differences are particularly pronounced for the more advanced literacy skills. However, due to the limited number of villages, differences between the two are not statistically significant. We do not have a conclusive argument for why this is the case, though some differences between the two phases are noteworthy. First, the villages of the first phase were wealthier. Second, because the program had been implemented multiple times in these villages, the teachers had more experience. Third, teachers reported that during the second phase, they did not need to spend much effort to motivate learners. Fourth, the post-program survey was implemented within two days after the end of the program, whereas the tests in Phase 2 were implemented later on average.

Appendix Tables A4 and A5 show that the results are robust to the inclusion of baseline test scores.

For the main regressions up to now, we excluded learners who recognized more than 10 letters (of 52 in the Hindi alphabet), treating them as outliers. The official cutoff for the program was five to six letters. In Table A6 in the Appendix, we reported the effect of TA+ by including those initial high baseline score samples and testing for whether the effects on these learners were different. While high baseline learners seem to make more progress than low baseline learners (i.e., the coefficients are positive and substantial relative to the main effects of TA+), the difference in outcomes between high baseline learners and low baseline learners are only marginally statistically significant for the words per minute outcome. We conclude that there is weak evidence that learners with just a bit of knowledge at the baseline benefit more from the program. Given that we would consider most of these learners to be illiterate in that they cannot put the knowledge of a few letters to much use, a case can be made to extend the eligibility criteria to include the barely literate into the literacy program. An argument can be made that the difficulty of the program is tailored to learners with some background knowledge,

and that for the most illiterate, DA could consider a preliminary phase to bring completely illiterate adults to a minimum level of literacy before starting the program.

During our program implementation, we noticed that some learners only attended some classes and did not complete the entire program; however, in the previous instrumental variables regressions (for example the third set of regressions of Tables 4 and 5), we treated them as if they had been full participants. For Table A7 of the appendix, we estimate a regression considering only those participants who completed the program. Our point estimates show that the estimated impact is noticeably higher for those who completed the program in both literacy and numeracy tests compared with any type of participants.

6.3 Long-term Results

About nine to ten months after the endline (and 12 months after the baseline), we administered the same battery of tests to the women in the sample from the second phase. Tables 8 and 9 contain the results using the 12-month literacy and numeracy tests scores as dependent variables, where we implement a subset of specifications from Tables 4.[29] Surprisingly, for every dimension of literacy, the point estimates of the 12-month effects are *larger* than the estimates immediately after the program, which are presented in the second panels of Appendix Tables A2 and A3 (to compare IV estimates of Phase 2's immediate effect with the long-term effect). This is also true for the effects on math: the estimated impact on the 12-month total math score is larger than the estimated effect immediately after the end of the program. DA organizes reading groups that occur after the end of the program.[30] These relatively intensive reading groups might be the reason for larger effects on reading after one year. The larger effects on math scores may be because the numeracy skills are generally useful on a day-to-day basis and may be used by respondents.

[Tables 8 and 9 are here]

6.4 Reading Scores in Perspective

One of the shortcomings of research on adult literacy in developing countries is that there is no generally established benchmark for what literacy programs should specifically aim to achieve,

[29] The full set is available upon request, but the qualitative results are the same across specifications.

[30] The reading club in Phase 1 was called *apnipathshala*, and it continued for six months (two hours daily) after the TA+ program was completed. It was led by an unpaid volunteer in each TA center from the same village. In Phase 2, the reading club was called *gyanchoupalik*, which also continued for six months (three-hour-long sessions daily) after the program. It was led by a paid TA Saheli who was given two days of training prior to the beginning of the reading clubs and further periodic training. In Phase 2, each session included 45 minutes of reading, 45 minutes of writing, 45 minutes of games, and 45 minutes on mathematics.

though there are attempts to change this (Abadzi, 2012)[31,32]. We are aware of no comparable data that would help directly put the learning outcomes of the TA+ adult literacy program into perspective. To provide some context for the meaning of our coefficient estimates, we compare the reading levels of the adult learners in our literacy program to those of schoolchildren in Uttar Pradesh and India using the 2013 Annual Status of Education Report (ASER, 2013).

Comparison with child learning outcomes in India

Table 10 compares the results of learners in TA+ to the ASER results from 2013, which include rural students in both public and private schools.

The first three columns refer to grade-1 and grade-2 students for Uttar Pradesh (columns 1 and 2) and all of India (column 3). The rows indicate the levels of the ASER test: children are classified at the "nothing," "letter," "word," "paragraph" (grade 1-level text), and "story" (grade-2 text) levels based on defined performance criteria. For example, the inability to identify four out of five letters classifies the child at the "nothing" level.[33] From Table 10, we can see that 47% of children in grade 1 are classified under "nothing" in the all-India group (column 3).

Vagh (2012) compared the reading tests we use with the levels defined by the ASER tests for samples from the states of Bihar and Uttarakhand, two neighboring states of Uttar Pradesh. Using this comparison, we can benchmark our results against ASER levels by "converting" the results from the fluency test to ASER levels, which we do in the last four columns of Table 8.[34]

In particular, columns 4 through 7 present the computed proportions of control (columns 4 and 6) and treatment (columns 5 and 7) participants that fall into different ASER levels, based on the Bihar (columns 4 and 5) and Uttarakhand (columns 6 and 7) samples.

Comparing across columns, we note that TA+ does very well in moving learners away from the "nothing" category. TA+ learners perform approximately at the first-grade level of

[31] https://www.academia.edu/6922710/Standards_and_criteria_for_courses_aimed_at_teaching_basic_literacy

[32] Indeed Abadzi's(2004) survey of literacy programs notes that each of these programs have different measures that they measure up against.

[33] See www.asercentre.org for testing tools and the annual reports for test administration details.

[34] Vagh (2012) provides information on the number of letters read by students at the letter, syllable, word, and paragraph levels. We categorize our students into these categories, starting with the lowest so the average letters identified per minute for TA learners in that category corresponds to the average letters identified by children in the study by Vagh. For example, if children at the "nothing" level read 2.3 letters per minute on average, then we create our "nothing" category so the average score for the learners we assign to that category is 2.3. We do the same for the upper levels. The last category is the category that scores above the word level but does not reach the paragraph level on average.

schoolchildren in Uttar Pradesh and India. We note that the students in the first phase perform almost at the level of UP second graders.

[Table 11 about here]

The main shortcoming of this comparison is that we compare learning increases for TA+ with achievement levels for the schoolchildren. Thus, the extent to which children in Uttar Pradesh have a greater (or lesser) knowledge of letters before starting school relative to our control group will bias this comparison downwards (respectively upwards) and change the assessment of the effectiveness of TA+. In all, 24.5 percent of the schoolchildren in Uttar Pradesh go to kindergarten, often as a requirement before attending private school (ASER, 2013).
This absence of a "standard" profile of learning gains to which we could more convincingly compare TA+ learning outcomes is why benchmarks are necessary. Our study provides a first benchmark for future assessments of adult literacy programs.

6.5 Numeracy Scores in Perspective

We are in a somewhat-better position to compare the effectiveness of the TA+ program with respect to numeracy outcomes. To place the numeracy skills into perspective, we compare the effects of TA+ to those of a different program in India evaluated in a recent study (Banerji *et al.* 2015). This literacy program was conducted in 240 hamlets in two blocks (sub-districts) of the Purnia district in Bihar and two blocks of the Ajmer district in Rajasthan, and it was designed to determine whether a combined mother-and-child literacy program would improve child outcomes more than a mother literacy program alone or a child activity package alone. The "Mother Literacy" (ML) intervention provided daily literacy and numeracy classes. Volunteers were recruited from the community to teach classes for two hours per day. A version of Pratham's Read India methodology was modified to suit the interests of adults (see Banerji *et al.* 2015 for more details). In addition, a random sample of women also received a "Child Home Activities and Materials Packet" (CHAMP) that the women were supposed to use with their children to improve their children's outcomes. In general, Banerji *et al.* (2015) found that the combined ML and CHAMP intervention yielded the largest effects.

Banerji et al.'s actual numeracy test is different from ours in terms of the number and (perhaps) the difficulty of the test items within each task category (single-digit and double-digit number identification; single-digit addition; etc.). Therefore, we report a standardized measure of the effect of the program, called the "effect size" of both programs, as follows:

$$Effect\ size = \frac{mean\ of\ treatment - mean\ of\ control}{standard\ deviation\ of\ control}$$

We compare the effect sizes of the TA+ effects with the ones implied by Banerji *et al.* (2013)[35, 36] in Table 11. We compare the outcome of TA+ relative to the mother literacy program and the ML and CHAMP combination. Across all dimensions of numeracy, we find that the implied effect sizes of TA+ are larger than the program studied in Banerji *et al.* (2013). This suggests that the numeracy component of TA+ is more effective than traditional adult literacy programs.

[Table 12 about here]

7. Conclusion

The world has seen substantial progress in access to primary education, with primary school enrollment rates for children in developing countries now reaching over 90 percent[37]. However, a large number of youth and young adults who are illiterate have been neglected in this expansion of access. Recognizing this problem in India, the NLM has now extended its remit to include youth and adolescents within its target group. Nonetheless, adult illiteracy will only be reduced through *effective* adult literacy programs. This implies providing comparison benchmarks against which programs can measure themselves, perhaps by using the same assessment tools.

In this paper, we investigate the impact of Tara Akshar plus (TA+), a computer-based adult literacy and numeracy program. Through this evaluation, we also provide the first rigorous evaluation of a computer-based adult literacy program that we are aware of. TA+ operates in the very disadvantaged, rural setting of villages in eastern Uttar Pradesh, where literacy rates among females are extremely low. We find that TA+ is effective in imparting literacy and numeracy skills, especially given the short duration of its implementation.

While we document substantial effects of TA+, an open question remains that affects the literature on computer-based instruction more generally: to what extent can the success of TA+ be attributed to computer-based teaching? The success could occur because of the

[35] Banerji et al. (2003) also present results for reading, but their test expands on the ASER tests and their test items are not close to being comparable to the results we are presenting here. However, just in terms of effect sizes for literacy outcomes in general, we find a similar ranking as for the numeracy results.

[36] The results Banerji et al. report are in terms of the intention to treat effect. We make the numbers comparable by accounting for the fact that only 34 percent of their sample participated in the treatment. Assuming no spillovers and only one-sided compliance, the instrumental variables effect is about three times the intent-to-treat effect.

[37] http://www.un.org/millenniumgoals/2014%20MDG%20report/MDG%202014%20English%20web.pdf

effectiveness of DA or the efficiency of their teaching model. The only way to answer this question is for other agencies (either other NGOs or the government) to implement a similar computer-based, similar to Bold and Sandefur (2014) who investigate differences in implementation and effectiveness of the same program when implemented by a government or an NGO. We leave these questions for future research.

References

Abadzi, Helen. 1994. "What We Know about Acquisition of Adult Literacy: Is There Hope?" Working Paper. World Bank. https://uta-ir.tdl.org/uta-ir/handle/10106/24626.

Abadzi, Helen. 2003. "Teaching Adults to Read Better and Faster: Results from an Experiment in Burkina Faso." Working Paper. World Bank. https://uta-ir.tdl.org/uta-ir/handle/10106/24335.

Aker, Jenny C., Christopher Ksoll, and Travis J. Lybbert. 2012. "Can Mobile Phones Improve Learning? Evidence from a Field Experiment in Niger." American Economic Journal: Applied Economics 4 (4): 94–120. doi:10.1257/app.4.4.94.

Angrist, Joshua, and Victor Lavy. 1999. "New Evidence on Classroom Computers and Pupil Learning." Working Paper 7424. National Bureau of Economic Research.

Banerjee, Abhijit, Shawn Cole, Esther Duflo, and Leigh Linden. 2007. "Remedying Education: Evidence From Randomized Experiments in India." The Quarterly Journal of Economics 122(3): 1235-1264.

Banjeri, R. and James Berry, Marc Shotland (2015) The Impact of Mother Literacy and Participation Programs on Child Learning: Evidence from a Randomized Evaluation in India. Working Paper.

Barrera-Osorio, Felipe, and Leigh L. Linden. 2009. "The Use and Misuse of Computers in Education: Evidence from a Randomized Experiment in Colombia." World Bank Policy Research Working Paper Series, Vol. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1344721.

Barrow, Lisa, Lisa Markman, and Cecilia E. Rouse. 2008. "Technology's Edge: The Educational Benefits of Computer-Aided Instruction." NBER Working Paper 14240. National Bureau of Economic Research, Inc. https://ideas.repec.org/p/nbr/nberwo/14240.html.

Basu, Kaushik, Ambar Narayan, and Martin Ravallion. 2001. "Is Knowledge Shared within Households? Theory and Evidence for Bangladesh." Working Paper 82. Institute for Social and Economic Change, Bangalore. https://ideas.repec.org/p/sch/wpaper/82.html.

Beuermann, Diether W., Julian P. Cristia, Yyannu Cruz-Aguayo, Santiago Cueto, and Ofer Malamud. 2013. "Home Computers and Child Outcomes: Short-Term Impacts from a Randomized Experiment in Peru." Working Paper 18818. National Bureau of Economic Research.

Blunch, Niels-Hugo, and Claus C. Pörtner. 2009. “Literacy, Skills and Welfare: Effects of Participation in Adult Literacy Programs.” SSRN Scholarly Paper ID 905583. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=905583.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. “Bootstrap-Based Improvements for Inference with Clustered Errors.” The Review of Economics and Statistics 90 (3): 414–27.

Card, David E., Pablo Ibarraran, and Juan Miguel Villa. 2011. “Building in an Evaluation Component for Active Labor Market Programs: A Practitioner’s Guide.” SPD Working Paper 1101. Inter-American Development Bank, Office of Strategic Planning and Development Effectiveness (SPD). https://ideas.repec.org/p/idb/spdwps/1101.html.

Carron, Albert V., and Kevin S. Spink. 1995. “The Group Size-Cohesion Relationship in Minimal Groups.” Small Group Research 26 (1): 86–105.

Chudgar Amita. 2014. “Mobile phone use and ownership among illiterate adults and its implications for adult literacy training” International Journal of Educational Development, 34 (1), 20-29.

Daniel Ortega, Francisco Rodríguez. 2008. “Freed from Illiteracy? A Closer Look at Venezuela's Misión Robinson Literacy Campaign.” Economic Development and Cultural Change 57 (1): 1–30.

Gakidou, Emmanuela, Krycia Cowling, Rafael Lozano, and Christopher J. L. Murray. 2010. “Increased Educational Attainment and Its Effect on Child Mortality in 175 Countries between 1970 and 2009: A Systematic Analysis.” Lancet (London, England) 376 (9745): 959–74.

Gibson, John. 2001. “Literacy And Intrahousehold Externalities.” SSRN Scholarly Paper ID 252806. Rochester, NY: Social Science Research Network.

Good, R. H. III, Kaminski, R. A., Simmons, D., &Kame’enui, E. J. (2001). Using Dynamic Indicators of Basic Literacy Skills in an outcome driven model: Steps to reading outomes. OSSC Bulletin, 44(1), 1-24

Haddad, Lawrence. 1999. “The Income Earned by Women: Impacts on Welfare Outcomes.” Agricultural Economics 20 (2): 135–41. 94. “Women's Income and Household Expenditure Patterns: Gender or Flow? Evidence from Niger.” American Journal of Agricultural Economics - AMER J AGR ECON 76 (5). doi:10.2307/1243421.

Lai, F., L. Zhang, X. Hu, Q. Qu, Y. Shi, Y. Qiao, M. Boswell and S. Rozelle (2013). "Computer assisted learning as extracurricular tutor? Evidence from a randomised experiment in rural boarding schools in Shaanxi.” Journal of Development Effectiveness 5(2): 208-231.

Linden, Leigh L. 2008. Complement Or Substitute?: The Effect of Technology on Student Achievement in India. InfoDev. http://www.povertyactionlab.org/sites/default/files/publications/Complement%20of%20Substitute-%20The%20Effect%20of%20Technology%20on%20Student%20Achievement%20in%20India.pdf.

Maddox, Bryan. 2007. “Worlds Apart? Ethnographic Reflections on ‘Effective Literacy’ and Intrahousehold Externalities.” World Development 35 (3): 532–41.

Mo, Di, Johan Swinnen, Linxiu Zhang, Hongmei Yi, Qinghe Qu, Matthew Boswell, and Scott Rozelle. 2013. “Can One-to-One Computing Narrow the Digital Divide and the Educational Gap in China? The Case of Beijing Migrant Schools.” World Development 46 (June): 14–29.

Oxenham, John. 2002. “Skills and Literacy Training for Better Livelihoods : A Review of Approaches and Experiences,” June. https://openknowledge.worldbank.org/handle/10986/9767.

Paxson, Christina, and Norbert Schady. 2005. “Cognitive Development among Young Children in Ecuador : The Roles of Wealth, Health and Parenting.” Policy Research Working Paper Series 3605. The World Bank. https://ideas.repec.org/p/wbk/wbrwps/3605.html.

Senauer, Benjamin, Marito Garcia, and Elizabeth Jacinto. 1988. “Determinants of the Intrahousehold Allocation of Food in the Rural Philippines.” American Journal of Agricultural Economics 70 (1): 170–80.

Strauss, John, and Duncan Thomas. 1995. “Human Resources: Empirical Modeling of Household and Family Decisions.” Handbook of Development Economics. Elsevier. http://econpapers.repec.org/bookchap/eeedevchp/3-34.htm.

Thomas, Duncan. 1990. “Intra-Household Resource Allocation: An Inferential Approach.” Journal of Human Resources 25 (4): 635–64.

Vagh, ShaherBanu, The Abdul LatifJameel Poverty Action Lab (J-PAL), Pratham and ASER (2012). Evaluating READ INDIA: The Development of Tools for Assessing Hindi Reading and Writing Ability and Math Skills of Rural Indian Children in Grades 1-51. Mimeo.

Vagh, ShaherBanu (2012) Validating the ASER Testing Tools: Comparisons with Reading Fluency Measures and the Read India Measures. Mimeo.

White, Howard, and Edoardo Masset. 2003. “The Importance of Household Size and Composition in Constructing Poverty Profiles: An Illustration from Vietnam.” Development and Change 34 (1): 105–26. doi:10.1111/1467-7660.00298.

Appendix A: Reading and Numeracy

For the TA+ outcome evaluation, literacy performance indicators were required. This set of indicators was complemented with indicators of numeracy performance, which often develop with literacy, and other cognitive indicators, which help in the interpretation of literacy and numeracy outcomes.

The specific tests we use were developed by the ASER Centre to study how well the ASER tests compare with EGRA/EGMA type tests. We chose them for a variety of reasons.

- These measures were modeled after the EGRA/EGMA indicators developed by RTI International (an international non-profit organization headquartered in North Carolina, USA) and the Dynamic Measurement Group (developers of the well-known DIBELS indicators, headquartered in Oregon, USA). These indicators have been tested extensively, successfully adapted in many languages, and used in many literacy/numeracy projects around the world.
- The comparisons with ASER tests are well documented, and they have demonstrated psychometric properties. The reliability and validity of these measures are now well established.
- The measures we have adopted have been used extensively in the Reading India project. They thus permit us the comparison with other literacy-numeracy program evaluation studies.
- These tests are designed to assess the mastery of different segmental units of the Hindi language, including *akshar*, *barakhadi*, real words, invented words, and sentences. This feature makes it possible to differentiate the impact foci of the intervention program. This level of detail in the range of indicators is particularly useful for a) extracting a differentiated view of an intervention program impact and b) tweaking the intervention program's components to improve its effectiveness in the future.

Reading test:

The Fluency Battery was adapted by ASER from the Early Grade Reading Assessment (USAID) and the Dynamic Indicators of Basic Early Literacy Skills (University of Oregon Center on Teaching and Learning). The material was extensively evaluated and piloted to ensure its grade and content appropriateness for the population of interest. Scores for the fluency reading subtests represent the number of units (*akshar*s/words/non-words) read accurately in one minute, and scores for the reading comprehension subtest represent the number of questions correctly answered. The total administration time for the Fluency Battery is about 10 minutes.

The assessment of fluency is based on the premise that the ability to read fluently, i.e., with sufficient speed and accuracy, is important to read well and to comprehend text. The fluent decoding of letters, letter combinations, words in list form, and words in connected text are important and robust correlates of early reading ability and comprehension. The automaticity of these lower-level skills ensures that limited cognitive resources, such as attention and memory, can be freed and allocated to the higher-level skills of meaning making. Hence, fluency measures, which are orally administered tests, are widely used to assess children's early reading ability in English and several other languages.

As in the numeracy test, the test is structured to increase in difficulty. The objective is to gauge the comfort level of the women in the village with regard to the recognition of different letters, words, and reading them to form sentences. This test also includes two exit points. First, if a respondent fails to read a single word in the third round, the test stops. Second, if a respondent fails to read more than three sentences in the fifth round, the test stops.

Numeracy test

This test involves counting the number of a particular object and recognizing the numbers in numeric form. In addition, respondents are asked to perform simple mathematical observations, e.g., single-digit recognition, two-digit recognition, addition, and subtraction. The test includes two exit points. First, if a respondent makes four errors in a row or answers fewer than four digits right in the second round (recognizing single-digit numbers), the test stops there. Second, in part three (recognition of double-digit numbers), if the respondent incorrectly answers or fails to answer four in a row, the question (not the test) is stopped and we move on to the next question.

Appendix Table B1. Overview: Reading

Akshar (letter) reading fluency	Learners are shown a randomly arranged set of characters (*akshar*s) from the Hindi alphasyllabary and asked to sound them out. The score indicates the number of characters (*akshar*s) correctly sounded out in 1 minute. The administration of the test stops if the respondent is not able to read more than three letters in one minute.
Barakhadi (consonant-vowel syllable) reading fluency	Learners are shown a randomly arranged set of consonant-vowel (CV) *akshar* units and asked to decode them orally. The score indicates the number of *barakhadi* units decoded correctly in 1 minute.
Word reading fluency	Learners are shown a list of one- or two-syllable words and asked to read them aloud. The score indicates the number of words read correctly in 1 minute. The administration of the test stops if the respondent is not able to read a single word in one minute.
Nonword reading fluency	Learners are shown a list of one- or two-syllable invented words (or nonwords) and asked to read them aloud. The score indicates the number of nonwords read correctly in 1 minute.
Grade 1 level passage reading fluency	Learners are asked to read aloud passages comprising 7 sentences and 64 words. The score indexes the number of words read correctly in 1 minute. The administration of the test stops if the respondent is not able to read more than 3 lines. Here the stop rule is not with respect to sentences, but lines.
Grade 2 level passage reading fluency	Learners are asked to read aloud passages comprising 9 sentences and 94 words. The score indexes the number of words read correctly in 1 minute.
Grade 1 level comprehension questions	Learners are asked to answer four comprehension questions on each passage of the Grade 1 level passage reading fluency sub-test. The score is the number of questions answered correctly.
Grade 2 level comprehension questions	Learners are asked to answer two comprehension questions on each passage of the Grade 2 level passage reading fluency sub-test. The score is the number of questions answered correctly.

Appendix Table B2. Overview: Math Assessment

Math Assessment	
Oral counting	Learners are shown sets of pencils and asked to count the number in each set and state this verbally.
Number identification: single digit	Learners are shown single- digit numbers and asked to name them.
Number identification: two digits	Learners are shown two-digit numbers and asked to name them.
Counting: one-to-one correspondence	Learners are shown sets of objects and asked to count the number of objects in each set, and then circle the correct number.
Counting: one-to-one correspondence with writing	Learners are shown sets of objects and asked to count the number in each set, and then write the correct number.
Missing item	Learners are shown series of numbers with one number missing in each series and asked to write the missing number. The score indicates the number of correct responses.
Addition problems	Learners are given addition problems (e.g., How much is 1 and 2 together?). The score indicates the number of correct responses.
Subtraction problems	Learners are given subtraction problems (e.g., How much is 1 and 2 together?). The score indicates the number of correct responses.

Table 1: Treatment-wise Summary statistics and Balance test of Individual and Household Characteristics of Respondents

Variables	Treatment group (s.d.)	Control Group (s.d.)	Difference (s.e.)
Age (in years)	33.55	34.53	-1.00
	(10.45)	(9.88)	(0.79)
Never Been to School (dummy)	0.93	0.94	-0.01
	(0.25)	(0.23)	(0.02)
Upper Caste (dummy)	0.05	0.02	0.02
	(0.23)	(0.15)	(0.01)
OBC (dummy)	0.52	0.49	0.00
	(0.50)	(0.50)	(0.03)
Married (dummy)	0.89	0.90	0.00
	(0.31)	(0.30)	(0.03)
Unmarried (dummy)	0.06	0.05	0.02
	(0.25)	(0.21)	(0.02)
Widowed (dummy)	0.04	0.04	-0.01
	(0.18)	(0.21)	(0.02)
No of children in the household	3.30	3.31	0.07
	(2.08)	(1.92)	(0.15)
Household is landless (dummy)	0.02	0.03	-0.01
	(0.15)	(0.18)	(0.01)
Household has BPLCard (dummy)	0.29	0.30	0.00
	(0.45)	(0.46)	(0.04)
Electrified House (dummy)	0.27	0.26	0.00
	(0.44)	(0.44)	(0.03)
Brick build house (dummy)	0.57	0.58	0.00
	(0.50)	(0.50)	(0.04)
House has Sanitary Toilet (dummy)	0.04	0.04	0.00
	(0.20)	(0.20)	(0.01)
Household has access to safe source of water (dummy)	0.36	0.40	-0.04
	(0.48)	(0.49)	(0.04)
Household uses less polutant fuel source (dummy)	0.02	0.01	0.01
	(0.15)	(0.12)	(0.01)
Household has heating equipents (dummy)	0.04	0.02	0.01
	(0.18)	(0.15)	(0.01)
Household has TV/DVD/VCR/VCD (dummy)	0.24	0.21	0.01
	(0.43)	(0.41)	(0.03)
Household has sewing machine (dummy)	0.12	0.07	0.05
	(0.33)	(0.25)	(.023)**
Participants main Occupation Agriculture (dummy)	0.13	0.12	0.00
	(0.34)	(0.33)	(0.02)
Participants main Occupation Household work (dummy)	0.73	0.72	0.02
	(0.45)	(0.45)	(0.03)
Participant owns a Mobile Phone	0.38	0.39	-0.01
	(0.49)	(0.49)	(0.04)
PPI Score	25.01	24.49	0.27
	(12.63)	(12.85)	(0.98)

Notes: Values reported in Columns 1 & 2 are the means (and in the row beneath within bracket are the standard deviation) of the respective variables. Column 3 provides an estimate of the difference (and in the row beneath is the standard deviation) from a regression that includes strata (sub-village) fixed effects with robust standard errors clustered at the village level. ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively.

Table 2: Occupation of Survey participants

Occupation	Percent
Farm Activities	**18.71**
Agricultural work on own farm	17.66
Agricultural work (Off Farm)	1.05
Non-Farm Activities	**5.1**
Industrial Enterprise	0.45
Employee in manufacturing/mechanics	2.4
Trade	0.3
Construction work	0.3
Self Employed Profession	1.05
Miscellaneous Services	0.6
Activities that do not generate income	**76.2**
Observation	668

Note: Authors' Calculation.

Table 3: Balance test for baseline literacy and numeracy tests

Variables	Treatment group	Control Group	Difference
Letter reading fluency (letters per minute)	0.12	0.10	0.00
	(0.66)	(0.70)	(0.05)
Syllable reading fluency (syllables per minute)	0.36	0.51	-0.14
	(4.23)	(4.79)	(0.36)
Word reading fluency (words per minute)	0.36	0.28	0.17
	(4.26)	(3.60)	(0.31)
Non-word reading fluency (non-words per minute)	0.13	0.03	0.11
	(1.95)	(0.58)	(0.12)
Words per minute- grade 1 level	0.27	0.11	0.22
	(4.51)	(1.87)	(0.29)
Words per minute - grade 2 level	0.27	0.00	0.31
	(4.28)	0.00	(0.26)
Counting objects - name orally (out of 3)	2.86	2.86	0.00
	(0.52)	(0.50)	(0.04)
Number identification - single digit (out of 10)	2.06	2.16	-0.16
	(3.85)	(3.90)	(0.30)
Number identification - two digits (out of 10)	0.42	0.43	0.03
	(1.59)	(1.66)	(0.13)
Counting objects – circle a number (out of 2)	0.34	0.36	-0.03
	(0.75)	(0.76)	(0.06)
Counting objects – write number (out of 2)	0.18	0.21	-0.04
	(0.54)	(0.60)	(0.05)
Find the missing number (out of 2)	0.12	0.10	0.01
	(0.41)	(0.34)	(0.03)
Addition problems (out of 4)	0.16	0.14	0.01
	(0.68)	(0.56)	(0.05)
Subtraction problems (out of 2)	0.06	0.03	0.04
	(0.30)	(0.16)	(.019)*
Total Math Score (out of 35)	6.20	6.28	-0.14
	(6.94)	(6.71)	(0.54)

*Notes: Values reported in Columns 1 & 2 are the means (and in the row beneath in the bracket are the standard deviations) of the respective variables. Column 3 provides an estimate of the difference (and in the row beneath in the bracket is the standard deviation) from a regression that includes strata (sub-village) fixed effects with robust standard errors clustered at the village level. ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively.*

Table 4: Impact of Tara Akshar Plus Program on literacy outcomes (Combined Endline data)

VARIABLES	(1) Letter reading fluency (letters per minute)	(2) Syllable reading fluency (syllables per minute)	(3) Word reading fluency (words per minute)	(4) Non-word reading fluency (non-words per minute)	(5) Words per minute- grade 1 level	(6) Words per minute - grade 2 level
Mean of the control group	3.405	1.697	1.303	0.445	1.686	1.011
S.D. of control group	9.177	7.069	6.916	2.144	9.754	6.062
Simple Impact (No household and individual controls)						
Tara Akshar Plus	9.465***	4.422***	2.829***	1.428***	2.737**	2.095**
	(1.296)	(0.927)	(0.655)	(0.394)	(1.090)	(0.836)
Observations	621	621	621	621	621	621
R-squared	0.236	0.146	0.118	0.104	0.101	0.078
Simple Impact (with household and individual controls)						
Tara Akshar Plus	9.293***	4.241***	2.789***	1.263***	2.335**	1.583*
	(1.341)	(0.993)	(0.681)	(0.380)	(1.047)	(0.790)
Observations	598	598	598	598	598	598
R-squared	0.318	0.216	0.166	0.159	0.167	0.125
IV Regression of TA+ participation instrumented with TA+ assignment (with household and individual controls)						
Participation in Tara Akshar Plus	12.563***	5.688***	3.732***	1.687***	3.098**	2.103**
	(1.417)	(1.226)	(0.837)	(0.465)	(1.285)	(0.983)
Observations	596	596	596	596	596	596
R-squared	0.383	0.237	0.180	0.166	0.170	0.129

*Notes: Values in the parenthesis are the robust standard errors of the estimations clustered at the village level. ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively. Household and individual controls include all the variables depicted in Table 1. All regressions include strata (sub-village) fixed effects.*

Table 5: Impact of Tara Akshar Plus Program on numeracy outcomes (Combined Endline data)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	**Counting objects (out of 3)**	**Number identification - single digit (out of 10)**	**Number identification - two digits (out of 10)**	**Counting objects – circle a number (out of 2)**	**Count objects – write number (out of 2)**	**Find the missing number (out of 2)**	**Addition problems (out of 4)**	**Subtraction problems (out of 2)**	**Total Math Score (out of 35)**
Mean of the control group	2.879	3.353	0.971	0.662	0.313	0.165	0.250	0.074	8.665
S.D. of control group	0.467	4.413	2.537	0.927	0.683	0.484	0.799	0.324	8.537
Simple Impact (No household and individual controls)									
Tara Akshar Plus	0.063*	3.090***	0.881***	0.619***	0.651***	0.332***	0.662***	0.169***	6.466***
	(0.035)	(0.358)	(0.271)	(0.081)	(0.080)	(0.058)	(0.095)	(0.031)	(0.802)
Observations	621	621	621	621	621	621	621	621	621
R-squared	0.057	0.168	0.125	0.160	0.203	0.167	0.163	0.097	0.186
Simple Impact (with household and individual controls)									
Tara Akshar Plus	0.056	3.069***	0.783**	0.604***	0.654***	0.305***	0.654***	0.156***	6.281***
	(0.038)	(0.403)	(0.292)	(0.088)	(0.092)	(0.063)	(0.110)	(0.038)	(0.905)
Observations	598	598	598	598	598	598	598	598	598
R-squared	0.078	0.256	0.223	0.242	0.274	0.235	0.238	0.179	0.288
IV Regression of TA+ participation instrumented with TA+ assignment (with household and individual controls)									
Participation in Tara Akshar Plus	0.077	4.122***	1.041***	0.810***	0.882***	0.419***	0.891***	0.215***	8.456***
	(0.048)	(0.407)	(0.346)	(0.097)	(0.105)	(0.069)	(0.122)	(0.042)	(0.935)
Observations	596	596	596	596	596	596	596	596	596
R-squared	0.068	0.323	0.246	0.305	0.354	0.276	0.277	0.203	0.356

*Notes: Values in the parenthesis are the robust standard errors of the estimations clustered at the village level. ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively. Household and individual controls include all the variables depicted in Table 1. All regressions include strata (sub-village) fixed effects.*

Table 6: Impact of Tara Akshar Plus Program on literacy (using pre and post data)

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	**Letter reading fluency (letters per minute)**	**Syllable reading fluency (syllables per minute)**	**Word reading fluency (words per minute)**	**Non-word reading fluency (non-words per minute)**	**Words per minute- grade 1 level**	**Words per minute - grade 2 level**
Panel 1: Difference-in-difference Regression						
Tara Akshar Plus	9.390***	4.699***	2.834***	1.344***	2.674**	1.876**
	(1.328)	(0.903)	(0.838)	(0.398)	(1.105)	(0.832)
Observations	1,242	1,242	1,242	1,242	1,242	1,242
R-squared	0.330	0.113	0.095	0.082	0.075	0.057
Panel 2: Difference-in-difference Regression (using IV and Fixed effects)						
Participation in Tara Akshar Plus	12.521***	6.207***	3.740***	1.774***	3.524***	2.473**
	(1.400)	(1.120)	(1.038)	(0.478)	(1.350)	(1.029)
Observations	1,238	1,238	1,238	1,238	1,238	1,238
R-squared	0.456	0.141	0.101	0.085	0.066	0.051

*Notes: The sample used in these regressions comprises particpants that were tested in both baseline and endline survey. Values in the parenthesis are the robust standard errors of the estimations, clustered at the village level. ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively. The instrumental variable used at the first stage for the variable "Participated in Tara Akshar Plus" is the randomized treatment allocation at the individual level during the baseline.*

Table 7: Impact of Tara Akshar Plus Program on Numeracy (using pre and post data)

VARIABLES	(1) Counting objects (out of 3)	(2) Number identification - single digit (out of 10)	(3) Number identification - two digits (out of 10)	(4) Counting objects – circle a number (out of 2)	(5) Count objects – write number (out of 2)	(6) Find the missing number (out of 2)	(7) Addition problems (out of 4)	(8) Subtraction problems (out of 2)	(9) Total Math Score (out of 35)
Panel 1: Difference-in-difference Regression									
Tara Akshar Plus	0.069	3.155***	0.869***	0.644***	0.676***	0.318***	0.656***	0.142***	6.529***
	(0.056)	(0.442)	(0.217)	(0.092)	(0.078)	(0.057)	(0.090)	(0.034)	(0.864)
Observations	1,242	1,242	1,242	1,242	1,242	1,242	1,242	1,242	1,242
R-squared	0.075	0.210	0.119	0.219	0.214	0.148	0.157	0.088	0.226
Panel 2: Difference-in-difference Regression (using IV and Fixed effects)									
Participation in Tara Akshar Plus	0.091	4.163***	1.144***	0.849***	0.894***	0.425***	0.877***	0.193***	8.636***
	(0.070)	(0.459)	(0.262)	(0.107)	(0.104)	(0.072)	(0.123)	(0.045)	(0.970)
Observations	1,238	1,238	1,238	1,238	1,238	1,238	1,238	1,238	1,238
R-squared	0.003	0.426	0.189	0.413	0.352	0.201	0.227	0.105	0.440

*Notes: The sample used in these regressions comprises particpants that were tested in both baseline and endline survey. Values in the parenthesis are the robust standard errors of the estimations, clustered at the village level. ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively. The instrumental variable used at the first stage for the variable "Participated in Tara Akshar Plus" is the randomized treatment allocation at the individual level during the baseline.*

Table 8: Long-term Impact of Tara Akshar Plus Program on literacy outcomes (Phase 2, 12 months Follow-up data)

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	**Letter reading fluency (letters per minute)**	**Syllable reading fluency (syllables per minute)**	**Word reading fluency (words per minute)**	**Non-word reading fluency (non-words per minute)**	**Words per minute-grade 1 level**	**Words per minute - grade 2 level**
Mean of the control group	3.013	1.209	0.993	0.459	1.157	1.065
S.D. of control group	9.523	6.851	5.446	2.522	6.917	6.710
Simple Impact (No household and individual controls)						
Tara Akshar Plus	8.359***	4.536***	2.077***	1.039**	2.235**	1.587**
	(1.081)	(0.876)	(0.536)	(0.348)	(0.741)	(0.632)
Observations	407	407	407	407	407	407
R-squared	0.163	0.096	0.096	0.105	0.106	0.100
Simple Impact (with household and individual controls)						
Tara Akshar Plus	7.806***	3.973**	1.667*	0.800	1.811	1.142
	(1.675)	(1.378)	(0.795)	(0.458)	(1.088)	(0.838)
Observations	405	405	405	405	405	405
R-squared	0.253	0.187	0.184	0.201	0.196	0.197
IV Regression of TA+ participation instrumented with TA+ assignment (with household and individual controls)						
Participation in Tara Akshar Plus	11.488***	6.065***	2.504**	1.214*	2.859*	1.842
	(1.802)	(1.714)	(1.065)	(0.650)	(1.535)	(1.205)
Observations	377	377	377	377	377	377
R-squared	0.331	0.212	0.208	0.221	0.211	0.211

*Notes: Values in the parenthesis are the robust standard errors of the estimations clustered at the village level. ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively. Household and individual controls include all the variables depicted in Table 1. All regressions include strata (sub-village) fixed effects.*

Table 9: Long-term impact of Tara Akshar Plus Program on numeracy outcomes (Phase 2, 12 months Follow-up data)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	**Counting objects (out of 3)**	**Number identification - single digit (out of 10)**	**Number identification - two digits (out of 10)**	**Counting objects – circle a number (out of 2)**	**Count objects – write number (out of 2)**	**Find the missing number (out of 2)**	**Addition problems (out of 4)**	**Subtraction problems (out of 2)**	**Total Math Score (out of 35)**
Mean of the control group	2.854	3.086	0.903	0.643	0.292	0.173	0.195	0.054	8.200
S.D. of control group	0.472	4.369	2.434	0.928	0.652	0.503	0.784	0.270	8.521
Simple Impact (No household and individual controls)									
Tara Akshar Plus	0.080*	3.083***	0.745**	0.631***	0.600***	0.338***	0.483***	0.088*	6.048***
	(0.042)	(0.413)	(0.250)	(0.086)	(0.090)	(0.079)	(0.114)	(0.040)	(0.880)
Observations	407	407	407	407	407	407	407	407	407
R-squared	0.054	0.147	0.082	0.141	0.148	0.143	0.121	0.101	0.147
Simple Impact (with household and individual controls)									
Tara Akshar Plus	0.081	2.896***	0.542	0.590***	0.585***	0.317***	0.441**	0.077	5.529***
	(0.045)	(0.514)	(0.309)	(0.108)	(0.119)	(0.082)	(0.158)	(0.052)	(1.214)
Observations	405	405	405	405	405	405	405	405	405
R-squared	0.085	0.244	0.182	0.237	0.257	0.256	0.199	0.185	0.255
IV Regression of TA+ participation instrumented with TA+ assignment (with household and individual controls)									
Participation in Tara Akshar Plus	0.109*	4.585***	0.975***	0.928***	0.912***	0.501***	0.662***	0.117*	8.789***
	(0.065)	(0.514)	(0.375)	(0.123)	(0.095)	(0.089)	(0.164)	(0.067)	(1.115)
Observations	377	377	377	377	377	377	377	377	377
R-squared	0.099	0.312	0.226	0.296	0.363	0.336	0.266	0.228	0.344

*Notes: Values in the parenthesis are the robust standard errors of the estimations clustered at the village level. ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively. Household and individual controls include all the variables depicted in Table 1. All regressions include strata (sub-village) fixed effects.*

Table 10: Comparison of Tara Akshar results to learning levels of children in Uttar Pradesh

	Uttar Pradesh	Uttar Pradesh	India	Proportion of Tara Akshar learners at specific ASER levels (Conversion A)		Proportion of Tara Akshar learners at specific ASER levels (Conversion B)	
ASER-level	**grade 1**	**grade 2**	**grade 1**	**Control**	**Treatment**	**Control**	**Treatment**
Nothing	0.52	0.3	0.47	0.87	0.43	0.89	0.51
Akshar	0.31	0.36	0.32	0.039	0.23	0.072	0.34
Word	0.09	0.14	0.13	0.075	0.32	0.027	0.098
Para	0.04	0.09	0.04	0	0	0	0
Story	0.04	0.11	0.04	0.015*	0.028*	0.015*	0.046*

Notes: Approximate proportion of learners who score at specific ASER equivalent levels in letter recognition on the fluency battery. Own calculations for Tara Akshar learners. Conversion A uses information on the mean fluency battery scores of children at different ASER levels for children from Bihar based on Vagh (2012); Conversion B uses information on children from Uttarkhand, also from Vagh(2012). Information on ASER levels of children in Uttar Pradesh from ASER (2013). ASER tests are conducted during the middle of a school year. * denotes that these learners score above the paragraph level, but the average is lower than the story level in the surveys analyzed in Vagh(2012)

Table 11: Effect size in TA+ and two variants of another adult literacy program

Literacy	Impact of two variants of an adult education program reported in Banerji et al. (2013)					Tara Akshar results		
	Control sd	Treatment effects from two versions of the program		Average Implied effect size	Max. Implied effect size	Control sd	Treatment effect	Implied effect size
		ML	ML & CHAMP					
Identify digits 1-9	0.46	0.0661***	**0.109*** **	0.56	0.70	0.427	0.367***	0.86
		[0.0103]	[0.0108]				(0.046)	
Identify digits 11-20	0.36	0.0228***	**0.0399*** **	0.26	0.33	0.327	0.119***	0.36
		[0.00677]	[0.00752]				(0.035)	
Identify numbers 21-99	0.28	0.0140***	**0.0147*** **	0.15	0.15	0.229	0.0456*	0.20
		[0.00435]	[0.00439]				(0.025)	
Single digit addition	0.34	0.0124	**0.0321*** **	0.19	0.28	0.242	0.112***	0.46
		[0.00752]	[0.00762]				(0.029)	
Single digit subtraction	0.28	**0.00651**	0.00592	0.07	0.07	0.248	0.194***	0.78
		[0.00574]	[0.00570]				(0.032)	

Note: Author's calculations. Left panel based on results in Banerji et al. (2013), adjusted for 0.34 impact of treatment assignment on participation. ML denotes "Mother Literacy", CHAMP receipt of a "Child Home Activities and Materials Packet". Bold numbers are the larger of the two treatment effects. Tara Akshar results are made comparable to Banerji et al. results by focussing on comparable tasks. Sample includes high baseline learners. Results from instrumental variables estimation. ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively.

Table A1: Impact of Tara Akshar Plus Program on literacy, by phase

VARIABLES	(1) Letter reading fluency (letters per minute)	(2) Syllable reading fluency (syllables per minute)	(3) Word reading fluency (words per minute)	(4) Non-word reading fluency (non-words per minute)	(5) Words per minute- grade 1 level	(6) Words per minute - grade 2 level
Phase 1: IV Regression of TA+ participation instrumented with TA+ assignment (with household and individual controls)						
Participation in Tara Akshar Plus	16.105***	8.785***	5.591***	2.779***	6.894***	3.871**
	(1.571)	(1.523)	(1.086)	(0.726)	(1.845)	(1.727)
Mean of the control group	12.975	7.483	4.890	2.346	6.617	4.172
S.D. of control group	14.662	15.647	12.268	6.878	19.601	15.263
Observations	200	200	200	200	200	200
R-squared	0.453	0.275	0.264	0.227	0.254	0.188
Phase 2: IV Regression of TA+ participation instrumented with TA+ assignment (with household and individual controls)						
Participation in Tara Akshar Plus	9.793***	3.388***	2.473***	1.012**	0.807	0.979
	(1.572)	(1.022)	(0.842)	(0.427)	(1.191)	(0.899)
Mean of the control group	6.320	2.449	1.833	0.652	1.442	1.155
S.D. of control group	11.239	8.513	6.292	2.986	7.745	6.665
Observations	396	396	396	396	396	396
R-squared	0.346	0.212	0.159	0.143	0.123	0.124

*Notes: Values in the parenthesis are the robust standard errors of the estimations clustered at the village level. ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively. Household and individual controls include all the variables depicted in Table 1. All regressions include strata (sub-village) fixed effects. The instrumental variable used at the first stage for the variable "Participated in Tara Akshar Plus" is the randomized treatment allocation at the individual level during the baseline.*

Table A2: Impact of Tara Akshar Plus Program on numeracy, by phase

VARIABLES	(1) Counting objects (out of 3)	(2) Number identification - single digit (out of 10)	(3) Number identification - two digits (out of 10)	(4) Counting objects – circle a number (out of 2)	(5) Count objects – write number (out of 2)	(6) Find the missing number (out of 2)	(7) Addition problems (out of 4)	(8) Subtraction problems (out of 2)	(9) Total Math Score (out of 35)
Phase 1: IV Regression of TA+ participation instrumented with TA+ assignment (with household and individual controls)									
Tara Akshar Plus	0.080	4.221***	1.794***	0.810***	0.798***	0.513***	1.223***	0.377***	9.814***
	(0.071)	(0.723)	(0.368)	(0.185)	(0.162)	(0.093)	(0.089)	(0.070)	(1.337)
Mean of the control group	2.846	5.534	1.774	1.149	0.801	0.516	0.796	0.244	13.661
S.D. of control group	0.451	4.423	3.092	0.977	0.898	0.705	1.268	0.551	10.256
Observations	200	200	200	200	200	200	200	200	200
R-squared	0.091	0.402	0.312	0.344	0.419	0.326	0.329	0.299	0.422
Phase 2: IV Regression of TA+ participation instrumented with TA+ assignment (with household and individual controls)									
Tara Akshar Plus	0.080	4.031***	0.442	0.779***	0.914***	0.348***	0.678***	0.128***	7.401***
	(0.058)	(0.463)	(0.404)	(0.101)	(0.128)	(0.077)	(0.139)	(0.043)	(1.188)
Mean of the control group	2.949	5.026	1.558	0.977	0.651	0.312	0.605	0.160	12.237
S.D. of control group	0.316	4.783	2.975	0.992	0.898	0.630	1.147	0.463	9.971
Observations	396	396	396	396	396	396	396	396	396
R-squared	0.080	0.319	0.236	0.307	0.364	0.269	0.284	0.219	0.353

*Notes: Values in the parenthesis are the robust standard errors of the estimations clustered at the village level. ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively. Household and individual controls include all the variables depicted in Table 1. All regressions include strata (sub-village) fixed effects. The instrumental variable used at the first stage for the variable "Participated in Tara Akshar Plus" is the randomized treatment allocation at the individual level during the baseline.*

Appendix A3: Sample Size, Program and Survey Participation and Attrition, by Round and Treatment status

		Baseline survey	Post-survey	Any TA participation (for survey participants)	No participation in TA (for survey participants)	Missing information (for survey participants)
Phase I						
	Treatment	139	135	132	3	0
	Control	99	97	10	87	0
	Total	238	232	142	90	0
Phase II						
	Treatment	264	238	176	62	0
	Control	215	192	182	8	2
	Total	479	430	358	70	2
Total		717	662	500	160	2

Table A4: Phase-wise Summary statistics of Individual and Household Characteristics of Respondents

Variables	Phase 1 (s.d.)	Phase 2 (s.d.)	Overall (s.d.)
Age (in years)	30.615	35.550	33.981
	(8.49)	(10.56)	(10.20)
Never Been to School (dummy)	0.980	0.920	0.940
	(0.14)	(0.27)	(0.24)
Upper Caste (dummy)	0.054	0.034	0.040
	(0.23)	(0.18)	(0.20)
OBC (dummy)	0.556	0.495	0.515
	(0.50)	(0.50)	(0.50)
Married (dummy)	0.912	0.886	0.895
	(0.28)	(0.32)	(0.31)
Unmarried (dummy)	0.059	0.059	0.059
	(0.24)	(0.24)	(0.24)
Widowed (dummy)	0.024	0.048	0.040
	(0.16)	(0.21)	(0.20)
No of children in the household	3.000	3.520	3.355
	(1.53)	(2.17)	(2.00)
Household is landless (dummy)	0.093	0.000	0.029
	(0.29)	0.00	(0.17)
Household has BPLCard (dummy)	0.249	0.316	0.295
	(0.43)	(0.47)	(0.46)
Electrified House (dummy)	0.444	0.170	0.257
	(0.50)	(0.38)	(0.44)
Brick build house (dummy)	0.439	0.636	0.574
	(0.50)	(0.48)	(0.50)
House has Sanitary Toilet (dummy)	0.034	0.045	0.042
	(0.18)	(0.21)	(0.20)
Household has access to safe source of water (dummy)	0.239	0.439	0.375
	(0.43)	(0.50)	(0.49)
Household uses less polutant fuel source (dummy)	0.039	0.009	0.019
	(0.19)	(0.10)	(0.14)
Household has heating equipents (dummy)	0.054	0.018	0.029
	(0.23)	(0.13)	(0.17)
Household has TV/DVD/VCR/VCD (dummy)	0.185	0.250	0.229
	(0.39)	(0.43)	(0.42)
Household has sewing machine (dummy)	0.122	0.082	0.095
	(0.33)	(0.27)	(0.29)
Participants main Occupation Agriculture (dummy)	0.278	0.041	0.116
	(0.45)	(0.20)	(0.32)
Participants main Occupation Household work (dummy)	0.493	0.848	0.735
	(0.50)	(0.36)	(0.44)
Participant owns a Mobile Phone	0.307	0.427	0.389
	(0.46)	(0.50)	(0.49)
PPI Score	22.566	25.466	24.544
	(13.19)	(12.04)	(12.48)

Notes: Values reported in Columns 1 & 2 are the means (and in the row beneath within bracket are the standard deviation) of the respective variables. Column 3 provides an estimate of the difference (and in the row beneath is the standard deviation) from a regression that includes strata (sub-village) fixed effects with robust standard errors clustered at the village level. ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively.

Table A5: Impact of Tara Akshar Plus Program on literacy (controling for baseline outcome)

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	**Letter reading fluency (letters per minute)**	**Syllable reading fluency (syllables per minute)**	**Word reading fluency (words per minute)**	**Non-word reading fluency (non-words per minute)**	**Words per minute-grade 1 level**	**Words per minute - grade 2 level**
Simple Impact with Endline Data (with household, individual control and baseline scores)						
Tara Akshar Plus	9.296***	4.248***	2.792***	1.267***	2.342**	1.596*
	(1.348)	(0.991)	(0.674)	(0.381)	(1.049)	(0.795)
Baseline Scores	0.479	0.036	-0.074***	-0.030**	-0.032***	-0.036***
	(0.595)	(0.043)	(0.016)	(0.011)	(0.009)	(0.011)
Observations	598	598	598	598	598	598
R-squared	0.319	0.216	0.167	0.159	0.167	0.125
IV Regression of TA+ participation instrumented with TA+ assignment (with household, individual control and base scores)						
Tara Akshar Plus	12.566***	5.700***	3.735***	1.687***	3.102**	2.112**
	(1.417)	(1.221)	(0.830)	(0.464)	(1.283)	(0.983)
Baseline Scores	0.246	0.046	-0.066***	-0.001	-0.011	-0.020*
	(0.576)	(0.034)	(0.017)	(0.009)	(0.009)	(0.010)
Observations	596	596	596	596	596	596
R-squared	0.383	0.237	0.181	0.166	0.170	0.129

*Notes: Values in the parenthesis are the robust standard errors of the estimations clustered at the village level. ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively. Household and individual controls include all the variables depicted in Table 1. All regressions include strata (sub-village) fixed effects. The instrumental variable used at the first stage for the variable "Participated in Tara Akshar Plus" is the randomized treatment allocation at the individual level during the baseline.*

Table A6: Impact of Tara Akshar Plus Program on numeracy (controling for baseline outcome)

VARIABLES	(1) Counting objects (out of 3)	(2) Number identification - single digit (out of 10)	(3) Number identification - two digits (out of 10)	(4) Counting objects – circle a number (out of 2)	(5) Count objects – write number (out of 2)	(6) Find the missing number (out of 2)	(7) Addition problems (out of 4)	(8) Subtraction problems (out of 2)	(9) Total Math Score (out of 35)
Simple Imapct with Endline Data (with household, individual control and baseline scores)									
Tara Akshar Plus	0.058	3.239***	0.795***	0.636***	0.681***	0.309***	0.663***	0.147***	6.608***
	(0.037)	(0.399)	(0.260)	(0.086)	(0.091)	(0.062)	(0.110)	(0.038)	(0.913)
Baseline Scores	0.072	0.492***	0.749***	0.463***	0.342***	0.249***	0.312***	0.349**	0.589***
	(0.062)	(0.056)	(0.105)	(0.046)	(0.072)	(0.080)	(0.094)	(0.145)	(0.079)
Observations	598	598	598	598	598	598	598	598	598
R-squared	0.086	0.391	0.386	0.349	0.315	0.255	0.265	0.214	0.428
IV Regression of TA+ participation instrumented with TA+ assignment (with household, individual control and base scores)									
Tara Akshar Plus	0.079*	4.327***	1.063***	0.849***	0.912***	0.423***	0.901***	0.203***	8.847***
	(0.047)	(0.420)	(0.303)	(0.100)	(0.106)	(0.070)	(0.126)	(0.044)	(0.993)
Baseline Scores	0.071	0.485***	0.738***	0.449***	0.319***	0.221***	0.291***	0.349***	0.572***
	(0.058)	(0.048)	(0.097)	(0.044)	(0.062)	(0.064)	(0.089)	(0.132)	(0.072)
Observations	596	596	596	596	596	596	596	596	596
R-squared	0.075	0.453	0.404	0.405	0.389	0.291	0.300	0.238	0.488

*Notes: Values in the parenthesis are the robust standard errors of the estimations clustered at the village level. ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively. Household and individual controls include all the variables depicted in Table 1. All regressions include strata (sub-village) fixed effects. The instrumental variable used at the first stage for the variable "Participated in Tara Akshar Plus" is the randomized treatment allocation at the individual level during the baseline.*

Table A7: Impact of Tara Akshar Plus Program with High Baseline scores

Panel 1: Literacy Test with Household and Individual Controls

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	**Letter reading fluency (letters per minute)**	**Syllable reading fluency (syllables per minute)**	**Word reading fluency (words per minute)**	**Non-word reading fluency (non-words per minute)**	**Words per minute- grade 1 level**	**Words per minute - grade 2 level**
Tara Akshar Plus	9.114***	4.154***	2.763***	1.291***	2.286**	1.527**
	(1.338)	(0.976)	(0.649)	(0.349)	(0.986)	(0.717)
Highbase Score	16.309***	11.423**	7.969	4.170	6.373	7.899
	(4.468)	(5.181)	(4.636)	(2.863)	(3.787)	(7.919)
Highbase * TA+	4.929	10.816	6.547	4.756	19.052**	16.195
	(5.587)	(7.265)	(6.407)	(4.469)	(9.003)	(13.961)
Observations	639	639	639	639	639	639
R-squared	0.407	0.325	0.291	0.293	0.294	0.241

Panel 2: Numeracy Test with Household and Individual Controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	**Counting objects (out of 3)**	**Number identification - single digit (out of 10)**	**Number identification - two digits (out of 10)**	**Counting objects – circle a number (out of 2)**	**Count objects – write number (out of 2)**	**Find the missing number (out of 2)**	**Addition problems (out of 4)**	**Subtraction problems (out of 2)**	**Total Math Score (out of 35)**
Tara Akshar Plus	0.056	3.025***	0.760**	0.593***	0.641***	0.290***	0.641***	0.154***	6.161***
	(0.038)	(0.397)	(0.290)	(0.087)	(0.091)	(0.062)	(0.108)	(0.037)	(0.898)
Highbase Score	0.089	2.143	1.429*	0.437	0.336	0.393**	0.522*	0.241*	5.591*
	(0.074)	(1.268)	(0.756)	(0.271)	(0.257)	(0.173)	(0.268)	(0.124)	(2.769)
Highbase * TA+	-0.163	-0.032	2.104	0.034	0.272	0.283	0.598	0.290	3.387
	(0.095)	(1.481)	(1.305)	(0.308)	(0.268)	(0.256)	(0.403)	(0.168)	(3.586)
Observations	639	639	639	639	639	639	639	639	639
R-squared	0.074	0.261	0.280	0.248	0.289	0.291	0.271	0.230	0.321

*Notes: Values in the parenthesis are the robust standard errors of the estimations clustered at the village level. ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively. Household and individual controls include all the variables depicted in Table 1. All regressions include strata (sub-village) fixed effects.*

Table A8: Impact of Tara Akshar Plus Program using Completed Participation as Dependent Variable in the Instrumental Variables Estimation)

Panel 1: Literacy Test with Household and Individual Controls

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	**Letter reading fluency (letters per minute)**	**Syllable reading fluency (syllables per minute)**	**Word reading fluency (words per minute)**	**Non-word reading fluency (non-words per minute)**	**Words per minute- grade 1 level**	**Words per minute - grade 2 level**
Tara Akshar Plus	11.849***	6.583***	4.061***	1.890***	3.995***	3.032**
	(1.381)	(1.274)	(0.929)	(0.587)	(1.309)	(1.389)
Observations	639	639	639	639	639	639
R-squared	0.365	0.247	0.216	0.205	0.193	0.156

Panel 2: Numeracy Test with Household and Individual Controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	**Counting objects (out of 3)**	**Number identification - single digit (out of 10)**	**Number identification - two digits (out of 10)**	**Counting objects – circle a number (out of 2)**	**Count objects – write number (out of 2)**	**Find the missing number (out of 2)**	**Addition problems (out of 4)**	**Subtraction problems (out of 2)**	**Total Math Score (out of 35)**
Tara Akshar Plus	0.011	4.174***	1.422***	0.818***	0.849***	0.395***	0.831***	0.240***	8.739***
	(0.033)	(0.433)	(0.249)	(0.095)	(0.108)	(0.062)	(0.099)	(0.032)	(0.883)
Observations	639	639	639	639	639	639	639	639	639
R-squared	0.068	0.334	0.263	0.307	0.349	0.282	0.275	0.215	0.367

*Notes: Values in the parenthesis are the robust standard errors of the estimations clustered at the village level. ***, **, * denote statistical significance at the 1, 5 and 10 percent levels, respectively. Household and individual controls include all the variables depicted in Table 1. All regressions include strata (sub-village) fixed effects.*