Scaling up Children' School Readiness in The Gambia: Lessons from an Experimental Study*

Moussa P. Blimpo
World Bank

Todd Pugatch
Oregon State University & IZA

March 2017

Abstract: Early childhood experiences lay the foundation for outcomes later in life. Large shares of children in Africa enter formal education without prior exposure to any structured pre-school program, including contact with, and practice of, the instructional language. Policymakers face a dual challenge of promoting access and quality in pre-school services, but evidence on how to manage this tradeoff is scarce. In early 2010's, The Gambia government developed a comprehensive curriculum and decided to experimentally test it with two approaches to delivering pre-school services nationally. In the first experiment, new community-based centers were introduced to randomly chosen villages that had no pre-existing structured services. Another group of communities, which did not receive the program, served as a comparison group. In the second experiment, existing kindergartens tied to primary schools, known as Annexes, were randomly split in two groups. One group received the new curriculum along with a comprehensive training for an effective implementation, while the other group received the curriculum only and served as control group. We found evidence that both programs show significant heterogeneous impact, while not raising significantly the overall average levels of school readiness measured by a standardized assessment of language and fine motor skills. Children from more advantaged households improved less when exposed to community-based ECD centers, while more disadvantaged children benefitted from provider training in existing Annexes. Taking into account additional implementation-related considerations, we argue that on both the equity and efficiency grounds that the expansion of formal public kindergarten tied to primary schools would be more effective than the initiation of a community-based approach.

JEL Codes: I25, I38, O15, O22

Keywords: early childhood development; cognitive stimulation; teacher training; Gambia; randomized control trials; Malawi Developmental Assessment Tool

* Author contact information: mblimpo@worldbank.org, todd.pugatch@oregonstate.edu. We thank the World Bank and the Japan Social Development Fund (JSDF) and the Strategic Impact Evaluation Fund (SIEF) for financial support. We also thank the support of the ESRC through a grant (RES-589-28-0001) to the Centre for Microdata Methods and Practice. We thank the Gambian Ministry of Basic and Secondary Education (MoBSE) for their active support and engagement; the National Nutrition Agency (NaNA) for the Implementation of the intervention; The Gambia Bureau of Statistics (GBOS) for organizing the data collection. Finally, we thank Yi-Kyoung Lee, Elizabeth Joof, Hervé Akinocho**,** for their guidance and support for our research. Roberta Nilson, Ruifan Shi, and Yabin Xu provided excellent research assistance. All errors are ours. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

1. Introduction

Early childhood experiences lay the foundation for outcomes later in life. Extensive research has documented links between early life circumstances and adult outcomes (see Currie and Almond 2011 for a review of the U.S. literature; Nores and Barnett 2010 for evidence outside the U.S.; and Tanner, Candland, and Odden 2015 for developing countries). Due to adverse early life circumstances, in developing countries almost 220 million children under age 5 fail to reach their development potential (S. Grantham-McGregor et al. 2007).

Nearly one-third of these children live in Sub-Saharan Africa, representing more than 60% of children in the region. The region faces two main challenges to promoting children's development. The first challenge is lack of access to early childhood development (ECD) services. Enrollment in pre-primary schooling is 22% in Sub-Saharan Africa, compared to 79% in OECD countries (World Bank 2015).

The second challenge is ensuring the quality of existing ECD services. Even where access exists, program quality can be low and uneven due to resource constraints and low skills of ECD providers. This low quality can result in inadequate cognitive stimulation for children, identified as one of four key risk factors hindering childhood development in developing countries (Walker et al. 2007; Walker et al. 2011).[1] Programs to stimulate cognitive abilities are among the most effective ECD interventions, in terms of both the proportion of interventions showing positive results (Engle et al. 2007; Nores and Barnett 2010; S. M. Grantham-McGregor et al. 2014) and the range of outcomes affected (Tanner, Candland, and Odden 2015). Although program quality is a basic determinant of the success of ECD services (Elango et al. 2015; Engle et al. 2007; Engle et al. 2011; Araujo et al. 2016), evidence on the effectiveness of efforts to improve program quality is scarce, particularly in developing countries.

This paper evaluates two experiments to improve access and quality in early childhood development services in The Gambia. In the first experiment, new community-based early childhood development centers were introduced to randomly chosen villages that had no pre-existing structured ECD services. We assess how such increased access to an ECD center affected children's development. In the second experiment, a randomly assigned subset of existing ECD centers received intensive provider training to implement a new curriculum. The

[1] The other risk factors are stunting, iodine deficiency, and iron deficiency anemia.

curriculum, known as Gambia Open and Active Learning Spaces (GOALS), aims to cognitively stimulate children aged 3-6 through structured play. A control group of ECD centers received the new curriculum without the provider training.

We find no evidence that either experimental intervention improved average levels of fine motor skills or language and hearing, our key child development measures. We also find no evidence of program effects at other points in the outcome distribution. Despite these results for the full sample, the interventions had differential effects according to baseline characteristics. Children from more advantaged households scored about 0.3 standard deviations lower in language skills when assigned to the community-based ECD treatment. Though surprising, these results are sensible if better-off households can provide a higher quality home environment than the community-based alternative. Similarly, the provider training treatment in ECD Annex centers benefitted disadvantaged children. Children with low initial development or whose mothers experienced greater mental distress at baseline increased their fine motor skills by 0.3-0.4 standard deviations due to treatment. These results underscore the importance of provider quality in ECD services in developing countries, particularly for disadvantaged children.

Separately, these two experiments explore the dual challenge of increasing access and quality in ECD services, but for which there is little evidence in Sub-Saharan Africa. Although the variation *between* the two experiments is not random due to differences in the underlying populations, we nonetheless attempt to compare the efficacy of the two treatments by reweighting the observed characteristics of participants to increase their comparability. This exercise allows us to generate additional evidence on the marginal benefit to expanding ECD services to new areas versus attempting to improve the quality of existing ECD programs, a tradeoff faced by policymakers in many countries in Sub-Saharan Africa and elsewhere in the developing world.

When reweighting the sample to compare children with similar baseline characteristics across experiments, we find that children in the ECD Annex experiment gained 0.4 standard deviations in language skills relative to the community-based ECD treatment. This effect holds regardless of whether a child was assigned to the provider training treatment or control group in ECD Annex centers. These results suggest that attaching ECD services to primary schools is

more effective at promoting language development than community-based ECD services, at least following initial construction of community-based centers.

Childhood development, or "the ordered emergence of interdependent skills of sensori-motor, cognitive-language, and social-emotional functioning" (Engle et al. 2007, p. 229), influences the ability of adolescents and adults to acquire skills (Cunha and Heckman 2007; Cunha, Heckman, and Schennach 2010; Heckman and Mosso 2014). By facilitating skill acquisition later in life, ECD programs tend to yield higher returns than alternative human capital investments (Heckman 2006). As Gertler et al. (2014) note, the returns to early childhood investment are likely highest in developing countries, where household resources and skills are more scarce. The years immediately preceding school entry, when children are ages 3-6, are particularly critical for public investment. During these years, children are more likely to transition into pre-school or other forms of center-based care and instruction, offering policymakers greater opportunities to influence early childhood development. Interventions that are high in quality and well-targeted to disadvantaged children can help overcome developmental deficits faced earlier in life or in utero.

We contribute to two strands of the literature on pre-school interventions (i.e., center-based programs for 3-6 year olds) in developing countries (for a review, see J. Behrman, Fernald, and Engle 2013). The first strand, corresponding to our evaluation of community-based ECD centers, compares children exposed to an ECD program with children exposed for shorter durations or to home-based care (e.g., J. R. Behrman, Cheng, and Todd 2004 for Bolivia; Aboud 2006 for Bangladesh; Berlinski and Galiani 2007, Berlinski, Galiani, and Gertler 2009 for Argentina; Berlinski, Galiani, and Manacorda 2008 for Uruguay; José Rosero and Oosterbeek 2011, Jose Rosero 2012 for Ecuador; Rao et al. 2012, Bouguen et al. 2014 for Cambodia; Martinez, Naudeau, and Pereira 2013 for Mozambique; Bernal and Fernández 2013 for Colombia; Krafft 2015 for Egypt; Pinto, Santos, and Guimarães 2016 for Brazil; Bastos, Bottan, and Cristia 2017 for Guatemala). Of these, only Martinez, Naudeau, and Pereira (2013) and Bouguen et al. (2014) are randomized control trials, and only Martinez, Naudeau, and Pereira (2013) study a program in Sub-Saharan Africa.[2]

[2] Several other studies have evaluated pre-school attendance in Sub-Saharan Africa (Taiwo and Tyolo 2002 for Botswana; Mwaura, Sylva, and Malmberg 2008; Malmberg, Mwaura, and Sylva 2011 for Kenya, Tanzania, and

Most of the aforementioned studies find positive effects on children's development and subsequent school performance, with some exceptions. Consistent with evidence from developed countries, the counterfactual environment for children in ECD programs is often decisive for impact estimates. Where the alternative to ECD programs is a lower quality home environment, estimated effects will be positive. Alternately, when ECD participation substitutes for other forms of schooling, such as early primary school attendance (Bouguen et al. 2014), or reduces parenting quality by increasing maternal labor force attachment (José Rosero and Oosterbeek 2011; Jose Rosero 2012), impact estimates can be zero or negative. These impacts can also vary within a program according to household characteristics (Pinto, Santos, and Guimarães 2016).

A second, smaller strand of the literature evaluates attempts to upgrade the quality of existing ECD programs via teacher training, as in the ECD Annex experiment in this paper.[3] Despite a few studies using credible identification strategies (Baker-Henningham et al. 2012 for Jamaica, Yoshikawa et al. 2015 and Bowne, Yoshikawa, and Snow 2016 for Chile, and Bernal 2015 for Colombia), there is scant evidence on provider training in Sub-Saharan Africa. Ozler et al. (2016) find that ECD providers in Malawi who were randomly assigned to receive in-service training improved classroom organization and teaching quality, but these behavioral changes increased child development only when paired with parent education. Wolf, Aber, and Behrman (2017) evaluate an experiment with a similar design in Ghana, but find gains in child development only when provider training was *not* paired with parent education. These contrasting findings suggest that program content and implementation can be crucial for efforts to improve ECD program quality. We build on the important contributions of these studies, adding to the thin evidence base on ECD provider training in Sub-Saharan Africa.

We also contribute to the broader evidence on how implementation quality influences the results of development programs (e.g., Bold et al. 2013). The success of a new curriculum like GOALS depends crucially on its delivery by teachers. Yet teachers' ability to implement curricular and pedagogical changes rests on their skills, which in developing countries are often inadequately suited to curricula (Pritchett 2013; Murnane and Ganimian 2014). Comprehensive

Uganda; Zuilkowski et al. 2012 for Zambia), but these studies lack credible identification strategies to distinguish the role of ECD exposure from unobserved child or parent attributes.

[3] Studies of efforts other than provider training intended to upgrade ECD quality in developing countries include Armecin et al. (2006) for the Philippines; He, Linden, and MacLeod (2009) for India; and the many follow-ups to the Mauritius Child Health Project (e.g., Raine, Venables, and Mednick 1997; Raine et al. 2001; Raine et al. 2003; Raine et al. 2010).

reviews find that pedagogical change is a leading mechanism behind successful education interventions in developing countries (Evans and Popova 2015), and that in-service training can improve teaching (Glewwe et al. 2011). This study contributes to the relatively small literature that uses experiments or other credible identification strategies to evaluate in-service teacher training in developing countries, which includes Nitsaisook and Anderson (1989), Angrist and Lavy (2001), Duflo et al. (2006), Piper and Korda (2011), and Zhang et al. (2013), in addition to the studies on ECD programs cited earlier.

Our results on the heterogeneous impacts of the program are consistent with evidence from both the U.S. (Elango et al. 2015; Tough 2016) and developing countries (Engle et al. 2007; Engle et al. 2011) that the largest benefits of ECD programs accrue to disadvantaged children. Because disadvantaged children tend to experience less supportive home environments, ECD services more easily stimulate their development, whereas ECD programs are closer substitutes for the higher quality home environment experienced by more advantaged children. Our results also demonstrate the potential for ECD services to mitigate a particular form of disadvantage, maternal depression, which has been highlighted as a risk of major concern for children in developing countries due to its infrequent diagnosis and treatment (Walker et al. 2007; Engle et al. 2007; Hadley et al. 2008; Walker et al. 2011).[4]

In the next section, we describe the program and data. Section 3 describes our empirical methodology. Section 4 presents experimental results, while Section 5 compares outcomes across experiments. Section 6 discusses lessons drawn from the study about implementation of different approaches to improving ECD services, and Section 7 concludes.

2. **Program Description and Data**

The Gambia is a small West African country with population 1.9 million and per capita income of 427 USD (World Bank 2016). Its education system is divided into six numbered administrative regions. Region 1 is the capital, Banjul, on the Atlantic coast, with Regions 2-6 located at increasing remove heading east. Regions become more rural and poor as distance from the capital increases.

[4] An important caveat is that we measure mother's mental distress; we lack measures of clinical depression.

Enrollment in early childhood development (ECD) programs among Gambian children aged 3-6 was 22% in 2007 (Zoyem 2010). The Gambian government has sought to increase this proportion, with a goal of integrating ECD programs into the standard primary school sequence. Other than informal home care, ECD services exist in three forms in The Gambia: 1) private centers, located mostly in relatively urban areas and serving richer children; 2) public centers which are built as annexes to primary schools (hereafter referred to as ECD Annexes); and 3) community-based centers, which are publicly run, stand-alone facilities located in communities without primary schools. ECD Annexes and community-based centers do not charge fees.

To increase ECD access and improve quality, the Gambian Ministry of Basic and Secondary Education (MoBSE, hereafter "the ministry") implemented a new early childhood curriculum in 2012. The curriculum, known as Gambia Open and Active Learning Spaces (GOALS), aims to promote creativity, problem solving, and confidence through structured play. Each day includes time for group activities, games to promote critical thinking, physical development, and music/singing. The curriculum runs for 40 weeks annually, concurrent with the academic calendar. Activities run for four hours daily Monday through Thursday and three hours on Friday. Teachers are expected to spend an additional 1.5 hours each day to prepare for the next day's session. Additionally, a monthly meeting is held with parents to discuss the program and children's progress.

To assess the effectiveness of the new curriculum, the ministry conducted two experiments in parallel.[5] Both experiments occurred in Regions 2 and 6, with treatments assigned at the village level. In the first experiment, the ministry built new community-based early childhood development centers in randomly chosen villages that had no pre-existing structured ECD services. Each community-based center delivered the new GOALS curriculum. Management committees, comprised of parents and community leaders, were formed to oversee the centers. We refer to this as the community-based treatment, and the corresponding control group of eligible villages without ECD services as the pure control.

In the second experiment, a teacher training program was delivered to a randomly chosen subset of ECD Annex centers. Teachers received intensive training in the new curriculum in

[5] A third experiment was conducted for children ages 0-3 (Akinocho et al. 2014), who were too young to be eligible for the services studied in this paper. This experiment occurred in a separate set of communities than the experiments described here.

three sessions between September 2012-September 2013. The trainings lasted five, eight, and eight days, respectively. The Gambia office of ChildFund, an international NGO, conducted the trainings, which were financed by the Japan Social Development Fund and World Bank. We refer to the experimental groups in this case as ECD Annex treated and ECD Annex control.

Additional program details help to clarify the experimental treatments. Teachers in the community-based ECD centers of the first experiment also received the training provided to the ECD Annex treatment group. Compared to the pure control group, the community-based treatment therefore represents the provision of new ECD services, including site construction, formation of management committees, curriculum, and teacher training. All ECD Annexes in the program regions received the new curriculum, but only the ECD Annex treatment group received the teacher training. The ECD Annex treatment therefore represents the teacher training, because the curriculum was delivered to all annexes. Table 1 summarizes the research design.

Although treatment status is randomly assigned to all sites, comparisons between ECD Annexes and community-based centers or the pure control cannot proceed as in an experiment, because the presence of an existing annex (and attached primary school) makes ECD Annex communities different from those eligible for community-based centers. We therefore rely on the random variation *within* each experiment separately, and use non-experimental techniques to compare outcomes *across* experiments. Importantly, communities do not overlap between the two experiments, mitigating concerns about self-selection of children into either experiment.

Figure 1 shows a map of sites included in the sample. The bulk of the sample sites are in Region 6, the poorest and most remote region of the country. Treatment was stratified by region to ensure sample balance. We include a dummy variable for Region 2 to account for this stratification throughout all analyses. The sample for the ECD Annex experiment includes 26 treatment sites and 27 control, while the sample for the community-based ECD experiment includes 40 treatment sites and 51 control.[6]

Table A1 uses the 2003 Census, the most recent conducted before the program began, to compare communities in program Regions 2 and 6 with the rest of the country (Regions 1 and 3-

[6] A larger number of control sites were sampled in the community-based ECD experiment because these were also sites for a separate experiment with children ages 0-3 (Akinocho et al. 2014), allowing for economies of scale in data collection. We dropped one ECD Annex treatment site because none of the sampled children met the age eligibility criteria according to their birth certificates. There are two fewer community-based ECD and pure control sites in the baseline data than in the full sample for the same reason.

5). Regions 2 and 6 differ from each other on many dimensions, with Region 2 more populous, educated, and developed. These differences reflect the more urban character of Region 2 compared with rural and remote Region 6. Stratifying the sample by region ensures that these differences are not spuriously correlated with treatment. Regions 2 and 6 also differ from the rest of the country, as shown by the many significant differences reported in columns (4)-(5) in the table. These differences suggest that the treatment effects reported in this study may not generalize to all regions of the country. However, the heterogeneity between Regions 2 and 6 ensures that the program occurred in a broad range of contexts found within the Gambia.

The new curriculum and initial teacher training began in September 2012, the start of the 2012-2013 academic year. A baseline survey was conducted before the academic year began, in May-July 2012. A representative sample of 16 households with children in the relevant age group (3-6 years) was taken within each community eligible for treatment. In households with multiple eligible children, one was randomly designated for sampling purposes. The household head and main caregiver for the eligible child were interviewed, with modules on household assets, expenditures, employment, demand for ECD services, health (own and the child's), parenting activities with the child, and attitudes towards disciplining the child. Anthropometric measurements (height, weight, brachial circumference) were taken from children.

An endline survey was conducted in November-December 2013, or 14-15 months after treatment began. The endline survey asked similar questions as the baseline, with an additional module to assess caregiver knowledge of childhood nutrition and health. Endline participants included baseline households and newly sampled households, in order to increase the sample size. We analyze attrition from the survey in the Data section. Figure 2 shows a timeline of project and research milestones.

In addition to the survey modules previously mentioned, children in sampled households were given the Malawi Developmental Assessment Tool (MDAT), a test of child development designed for rural Africa (Gladstone et al. 2010). The test consists of two modules: 1) fine motor skills, and 2) language and hearing. Each module has multiple versions tailored to different age ranges. The fine motor skills tool asks children to complete tasks such as stacking blocks in various configurations, placing pegs in a board, and determining the relative weight of objects. The language and hearing tests requires children to point to body parts by name, identify the

names or uses of objects, identify the letters in one's name, and similar tasks. Each item on the tests is marked as complete or incomplete, with the overall score determined by the total number of completed items. Enumerators received five days of training on the tests before administering them in the field. Not all children completed the MDAT due to interviewer time constraints.

Appendix A presents the MDAT versions used in the surveys. The baseline and endline used the MDAT versions intended for children aged 36-59 and 53-76 months, respectively (though children completed the tests regardless of age at the time of the surveys). We assessed the reliability of the MDAT by calculating Cronbach's α separately by module (fine motor and language and hearing) and survey wave (baseline and endline), using all available MDAT scores in the survey (including children from the pure control group and community-based ECD treatment). The Cronbach α values ranged from .82-.88, indicating high reliability to measure the underlying constructs.

Program implementation was largely successful. Although 14 of 40 community-based ECD treatment sites reported an implementation problem, most of the issues reported—such as absent or sick teachers or lack of materials—are common in developing country schools. Two sites reported that their facility had not been constructed, forcing teachers to provide services outdoors. Our results should therefore be interpreted as intent-to-treat effects, though we later check robustness by dropping sites reporting implementation problems. No ECD Annex treatment sites reported implementation problems.

3. Methodology

The research design allows us to use standard techniques in the analysis of cluster-randomized experiments. The unit of analysis is the child, with treatment assignment at the community level. Accordingly, we can analyze the effect of each experimental treatment through a comparison of mean outcomes between children in treated and control communities. We cluster all standard errors at the community level, in order to adjust for correlated outcomes among units exposed to the same treatment. To test for differences between treatment and control groups, we further adjust for the stratification of treatment status within regions. We make this adjustment by regressing the outcome on an indicator for treatment and a dummy for whether the community is in Region 2:

$$y_{ic} = \alpha + \beta D_c + \gamma Region2_c + \varepsilon_{ic} \quad (1)$$

where *i* indexes the child, *c* indexes the community, *y* is the outcome, *D* and *Region2* are indicators for treatment and Region 2, respectively, and ε is an error term. We estimate this equation separately for the community-based ECD and ECD Annex experiments, with the definition of treatment changing accordingly. In each case, the coefficient of interest is β, which measures the difference in mean outcomes between children exposed to each treatment compared to the corresponding control group. We will also run versions of equation (1) that include the baseline outcome y_{0ic} on the right-hand side in order to improve the precision of our estimates.

Our main outcomes of interest are scores for the two MDAT modules, fine motor skills and language and hearing. We normalize each score by child gender and age, using children from both experiments as the underlying population.[7] The outcome units are therefore measured in standard deviations from the relevant mean. Additionally, we analyze subsets of MDAT items which are most closely aligned with the structured play format of the GOALS curriculum, such as stacking and counting blocks. For each subset, we measure the percentage of items successfully completed by the child.

As discussed in the introduction, meeting the twin goals of access and quality in early childhood development services forces governments to make difficult choices. Comparing the efficacy of the community-based ECD treatment, which attempts to expand access to ECD services, with that of the ECD Annex treatment, which aims to increase quality, matters greatly for policy. Differences in the populations eligible for the two experiments present challenges for this comparison, as noted earlier.

We attempt to overcome this challenge by comparing *ex ante* similar children across experiments. Specifically, we compare the experimental treatments against each other using an inverse propensity score weighting estimator, which combines matching and reweighting to improve comparability between treated and control units (Hirano, Imbens, and Ridder 2003; Abadie 2005).[8] We first match on the propensity score by estimating the probability that the child is in the ECD Annex experiment and not the community-based ECD experiment

[7] Specifically, we regress a child's raw score on his/her age (in months), age squared, and a female dummy, then divide the residual by its standard deviation. We run this procedure separately for the fine motor and language and hearing modules and for baseline and endline.

[8] The exposition in this section closely follows Giordono and Pugatch (2017).

(regardless of treatment status within each experiment).[9] We then remove children falling outside the common support of the propensity score distributions for each experiment.

Using this trimmed sample, we weight by the inverse propensity score and estimate:

$$y_{ic} = \alpha + \beta_1 community_c + \beta_2 Annex_c + \gamma Region2_c + \varepsilon_{ic} \quad (2)$$

where *community* is an indicator for being in the community-based ECD treatment group, *Annex* is an indicator for being in the ECD Annex experiment, and all else is as in equation (1). Because the community-based ECD control group is the omitted category, the coefficients β_1 and β_2 measure the treatment effect of being in each of the corresponding groups relative to the pure control.

The identifying assumption is that treatment assignment is orthogonal to unobserved characteristics that also affect the outcome. In the case of the community-based ECD treatment, this is ensured by random assignment, whereas in the case of the ECD Annex experiment we rely on the weights to balance average characteristics with the community-based ECD sample. Although the method requires that unobserved characteristics do not systematically differ among children in different experimental groups with the same propensity score, we rely on a rich set of predetermined characteristics, including baseline MDAT scores, to generate the weights. Section 5 presents details of propensity score estimation. Moreover, the method consistently estimates the average treatment effect even if the propensity score equation is misspecified (Robins and Rotnitzky 1995; Imbens and Wooldridge 2009). Under this identifying assumption, we can also compare the ECD Annex experiment to the community-based ECD treatment through a test of the null hypothesis $\beta_2 - \beta_1 = 0$.

[9] Section 5 presents details of propensity score estimation.

Finally, we can compare the ECD Annex control and treatment separately to the community-based ECD groups by disaggregating the *Annex* experiment dummy in equation (2) into separate indicators of treatment assignment:

$$y_{ic} = \alpha + \beta_1 community_c + \beta_2 Annex\_control_c + \beta_3 Annex\_treated_c + \gamma Region2_c + \varepsilon_{ic} \quad (3)$$

Estimates of β_2 and β_3 in equation (3) help to determine if the Annex setting without training is leads to differences with the community-based experiment, or whether the provider training is essential for implementing the new curriculum.

4. Experimental results

4.1. Attrition and baseline balance

Table 2 presents sample sizes and analyzes attrition. The sample includes all children aged 3-6 at the time of the baseline. Panel A, columns (1)-(2) show the number of children sampled in baseline and endline, respectively, separately by experiment and treatment group. The endline sample is split into two categories: those who appeared in the original (baseline) sample, and those newly sampled to increase the sample size. Columns (3)-(4) of Panel A present the same information but restrict the sample to those who completed the MDAT fine motor skills and language and hearing tests. Some children present in the baseline survey completed the MDAT at the endline but not the baseline, and therefore appear in the "original sample" group of column (4) although they lack a baseline score. For this reason, in column (5) we show the number of children who completed the MDAT in both baseline and endline.

Panel B of Table 2 analyzes sample attrition. We define two types of attrition: 1) attrition from the survey, in which a household that completes the baseline survey fails to complete the endline survey, and 2) attrition from the test, in which a child that completes the baseline MDAT does not complete the endline MDAT. Columns (1)-(3) of Panel B show the sample proportion of each type of attrition by treatment status. Attrition in the community-based ECD experiment exceeds that of the ECD Annex experiment. Within each experiment, however, neither type of attrition differs significantly between treatment and control groups, as shown by the p-values in column (5).

Table 3 presents baseline characteristics and tests for balance between treatment and control groups within each experiment. Columns (1)-(4) show the control group mean, treatment group mean, difference, and corresponding p-value for the community-based ECD experiment. Columns (5)-(8) repeat the exercise for the ECD Annex experiment. Baseline values across all groups demonstrate the economic and social disadvantages faced by Gambian children. Height-for-age, which proxies for early life nutrition, is more than 1.25 standard deviations below the international average for all groups. Average completed schooling of mothers is 2 years or less. Household expenditure per capita ranges from US$450-510, 14-18% of children were ill at the time of the survey, and mothers report high levels of mental distress (measured as the percentage of indicators of distress, such as feeling lonely, sad, or fearful, experienced "most of the time").

Looking across experiments, children in the community-based ECD experiment score lower in fine motor skills, language and hearing, and height-for-age than those in the ECD Annex experiment. These disadvantages are not surprising, as the presence of primary schools in ECD Annex communities suggest greater opportunities for investment in children. Nonetheless, household socioeconomic indicators are not uniformly higher in the ECD Annex experiment. We postpone formal comparison between experimental groups until Section 5.

Within experiments, treatment and control groups are broadly similar. Fine motor skills, language and hearing, and height-for-age are not significantly different between treatment and control in either experiment. Differences in most other characteristics are also not statistically significant, with a few exceptions. Children in the control group of the community-based ECD experiment are 16 percentage points more likely than the treatment group to be attending an ECD program at baseline, significant at 1%. To the extent that baseline ECD attendance reflects greater early childhood investment, it will bias us against finding an effect of the community-based ECD treatment. Beyond this, each experiment has one difference that is significant at 10% (household head's agricultural employment and schooling in the community-based ECD and ECD Annex experiment, respectively), which is about what we would expect by chance given the number of variables tested. We conclude that the treatment and control groups are broadly balanced on baseline characteristics within each experiment. Because MDAT scores will summarize the cumulative effect of early childhood investments, specifications that control for baseline scores will mitigate spurious treatment effects estimates.

4.2. Endline outcomes

We report endline MDAT scores and treatment effects estimates in Table 4. The first two columns show the control and treatment mean, respectively, for the community-based ECD experiment. Columns (3) and (4) show estimates of equation (1), with and without controlling for the baseline outcome. Columns (5)-(8) present the same information for the ECD Annex experiment. In the community-based ECD experiment, the treatment group scores below the control in both language and hearing and fine motor skills. Although this pattern is surprising, neither difference is statistically distinguishable from zero, regardless of whether we control for baseline score. We also fail to find significant treatment effects when disaggregating each MDAT module by subsets of items most closely associated with the ECD curriculum. Overall, we find no evidence of differences in average scores based on exposure to community-based ECD services.

When dropping sites that reported implementation issues, the treatment effect of community-based ECD on language and hearing falls to -0.24 standard deviations, significant at 10%. Including the baseline score increases the magnitude to -0.29 standard deviations, significant at 5%. The result is surprising, as we would expect that the estimated benefits of community-based ECD would increase when omitting sites with implementation issues. A potential explanation is reporting bias, in which the most conscientious administrators are more likely to report problems, leaving only the worst managed sites in the sample. We find no significant differences in fine motor skills when dropping sites with implementation issues. We omit results for brevity but present them in Table SA1 of the Supplemental Appendix.

In the ECD Annex experiment, the treatment group scores higher on both language and hearing and fine motor skills, as we would expect, but the differences are not precisely estimated. Among subsets of items, the only significant differences are for counting and ordering rows of items, though here the treatment group scores below the control, by 9 and 8 percentage points, respectively. These deficits are surprising, but not sufficiently large to lead to differences in overall scores.

The mean effects presented in Table 4 might mask changes in other features of the outcome distribution between children in the treatment and control groups of each experiment.

We explore this possibility in Figure 3, which plots the distributions of MDAT scores. Figure 3(a) shows kernel density estimates for the community-based ECD experiment, with baseline scores plotted in the first row and endline scores in the second row. Figure 3(b) shows the corresponding density estimates for the ECD Annex experiment. Comparing densities within a column shows how the distributions of each MDAT module shift between baseline and endline.[10]

Beginning with fine motor skills for the community-based ECD experiment in the first column of Figure 3(a), the treatment group lies slightly to the left of the control distribution at baseline. At endline, the mode of the treatment group distribution is to the right of the control group mode. The treatment group's thicker left tail and thinner right tail make the overall change unclear, however. For language skills in the second column, again the treatment group distribution lies slightly to the left of the control group, with only minor differences apparent at endline. In short, we fail to find strong evidence of relative shifts in the MDAT score distributions in the community-based ECD experiment.

Turning to the ECD Annex experiment in Figure 3(b), at baseline we see similar central locations for fine motor skills between the treatment and control groups, though the treatment group has a thicker right tail. By the endline, the treatment group distribution has shifted right relative to the control. In language and hearing, the treatment group begins to the right of the control group distribution at baseline, with the difference somewhat more pronounced at endline. These differences are consistent with the positive point estimates for the ECD Annex treatment coefficient in Table 4.

Although suggestive, the preceding visual inspection of unconditional distributions should not substitute for formal analysis. Even formal tests for equality of unconditional distributions would fail to account for the stratification of treatment assignment by region, or for differences in baseline outcomes. To overcome these limitations, we estimate equation (1) using quantile regressions. For each MDAT module, we estimate the coefficient on treatment status for the 5^{th}-95^{th} (conditional) quantiles, in increments of 5. We continue to cluster standard errors by community, the unit of treatment.

[10] Because we rely on different versions of the MDAT in baseline and endline, the densities are informative about relative changes in the treatment and control distributions, but not of absolute changes in child development.

Figure 4(a) presents results for the community-based ECD experiment, with results for the ECD Annex experiment in Figure 4(b). In each graph, the thick black line plots our estimates of the quantile treatment effects. The gray shaded area shows the 95% confidence interval around these estimates. For comparison, we also plot the OLS estimate of the average treatment effect in the long-dashed line, while the short-dashed line shows its 95% confidence interval. The red line shows the gridline for zero treatment effect.

For the community-based ECD experiment in Figure 4(a), a downward slope appears in quantile treatment effect estimates of fine motor skills beginning around the 40^{th} percentile. For language and hearing, estimates remain near the OLS coefficient across all quantiles. For the ECD Annex experiment in Figure 4(b), a downward slope appears in quantile treatment effect estimates around the 40^{th} percentile for fine motor skills and the 20^{th} percentile for language and hearing. These downward slopes suggest that treatment benefitted the middle of the outcome distributions most, with smaller benefits at the top of the distribution. Across both experiments and outcomes, however, estimates of quantile treatment effects are noisy. In fact, confidence intervals include zero around all the quantile effects estimated, suggesting that any differences between treatment and control distributions implied by Figure 3 are spurious. Results are similar when augmenting the quantile regression specification with baseline outcomes, or when using the nonparametric quantile treatment effect estimator of Firpo (2007).

4.3 Heterogeneous treatment effects

Although we have found no evidence of significant treatment effects for our main measures of childhood development in the analysis thus far, it is possible that null effects for each experimental sample mask significant effects for subgroups of children. Table 5 presents estimates of treatment effects for several subgroups, divided according to baseline characteristics. Each cell in the table reports a treatment effect estimate from a separate regression, with the rows indicating the outcome and columns indicating the subsample. We examine the following subgroups: male and female, to check for differences by sex; below and above median household assets, and whether the child's mother ever attended school, to look for differential effects by socioeconomic status; whether the child was sick in the past 3 days, to determine if a proxy for health correlates with outcomes; above and below median mother

mental distress, as a proxy for differences in the household environment; below or above median quantity of stimulating objects in the home, to determine if the curriculum's focus on structured play substitutes or complements similar household efforts; and whether the baseline outcome z-score is positive or negative, to determine whether children with different levels of baseline development respond differently to treatment. Subgroups defined as below/above median use both experimental samples to determine the threshold; results are similar when defining the median within each experimental sample. Panels A and B shows estimates of equation (1) for the community-based ECD and ECD Annex experiments, respectively, for specifications with and without the baseline outcome.

In the community-based ECD experiment, several subgroups experienced negative and significant treatment effects for language and hearing, ranging from -0.25 to -0.35 standard deviations. These subgroups are females, households with above median assets or stimulating objects in the home, and children with mothers who attended school or who were not sick in the past 3 days. Most of these estimates remain similar in magnitude and significance when controlling for the baseline outcome. With the exception of females, the pattern suggests that children from more advantaged households experienced slower language and hearing development via exposure to the community-based ECD program. Although surprising, these results are plausible if better-off households might be able to provide a higher quality home environment than the community-based alternative.[13]

For the ECD Annex experiment (Panel B), most point estimates are positive, but few subgroups of children have treatment effects significantly different from zero. Notable exceptions include the fine motor skills of children of mothers with mental distress above the median, and children with a negative baseline score. Effect sizes for these groups are sizable, with magnitudes of 0.36 and 0.34 standard deviations, respectively. The estimates remain significant and of similar magnitude when controlling for baseline outcomes. The treatment effect on language and hearing for children of mothers with high mental distress is also positive and significant when including the baseline score, albeit at the 10% level. These results provide some evidence that some groups of relatively disadvantaged children benefitted most from the ECD Annex treatment. The results

[13] We find no significant treatment effects for the community-based ECD treatment when splitting the sample by whether the child attended ECD at baseline. These results suggest that the baseline imbalance according to this characteristic is not driving results. Results not shown but available upon request.

are also consistent with the downward-sloping pattern of quantile treatment effects estimates, in which we found that the highest-achieving children benefitted least from treatment.

5. School Annexes versus Community-Based ECD: Between-experiment evaluation

[Intro here]

5.1

The previous section evaluated two experiments to improve early childhood development, using the random assignment within each experimental sample to identify treatment effects. We now compare these experimental interventions against each other, following the propensity score reweighting strategy described in Section 3. Because the community-based ECD experiment sought to increase ECD access, while the ECD Annex experiment sought to improve quality, this between-experiment comparison is important to understand the tradeoffs that governments face when allocating scarce ECD resources.

The populations eligible for each experiment live in distinct communities. Villages in the community-based ECD experiment lacked access to structured ECD services prior to the treatment, whereas villages in the ECD Annex experiment already had an ECD facility operating in conjunction with a primary school. These conditions reflect broader differences between households in the two experiments. Baseline balance tests similar to Table 3 between the two experimental samples reveal several significant differences. Children in the ECD Annex have higher baseline MDAT scores than their counterparts in the community-based ECD experiments, and their households are wealthier and more educated.[14]

To ensure comparability between the two experiments, we first estimate the probability that a child is in the ECD Annex sample by running a logit regression on a host of baseline characteristics. These characteristics are the child's age (in months); dummies for female, region, and previous ECD attendance; baseline z-scores in fine motor skills, language and hearing, and

[14] The table is omitted for brevity but appears in Supplemental Appendix Table SA2.

height-for-age; mother's schooling (in years); household expenditure per capita and willingness to pay for ECD services as a percentage of this expenditure; the proportion of 17 recommended vaccines received by the child; and an index of mother's mental distress, measured as the proportion of 11 mental health issues she reports experiencing "most of the time." We then exclude children whose propensity score falls outside the common support of the distributions from each experiment, leaving an estimation sample that includes 648 of the 844 children (77%) with an endline MDAT score.[15]

In Table 6, we check for balance in baseline fine motors skills and language and hearing using the trimmed sample. MDAT scores differ significantly between experiments when the data are unweighted, for both overall scores on each module and for most subsets of test items (columns 1-4). When weighting observations by the inverse of their propensity score (columns 5-8), no score differs at conventional significance levels.[16] These results give us confidence that reweighting generates an appropriate sample with which to compare experimental treatments.

In Table 7, we present results from inverse propensity score-weighted estimation of the effect of each experimental treatment. Column (1) shows estimates of equation (2) for language and hearing skills. The coefficient on the community-based ECD treatment indicator corresponds to a decrease of 0.64 standard deviations in language skills relative to the pure control group (the omitted category), significant at 5%. The effect attenuates somewhat to -0.37 when controlling for baseline score in column (2), but remains significant. These results are striking, given that the pure control group lacks ready access to structured ECD services and the variation between these groups is experimental. The estimates also contrast with the previous estimates of the community-based ECD treatment effect of -0.16 and -0.17 standard deviations, which were not statistically distinguishable from zero (Table 4, columns 3-4).

A clue to this discrepancy lies in the pure control group mean language score of 0.49 (reported halfway down Table 7), which is considerably higher than the unweighted mean of -0.07 (Table 4, column 1). In other words, the reweighting used to estimate equation (2) gives greater influence to relatively better-off children, who fare better at home than in community-

[15] Figure SA2 of the Supplemental Appendix shows the propensity score distributions.
[16] The result is not simply a mechanical consequence of reweighting. Although the overall score on each module enters the model for the propensity score, it is not obvious that the procedure would also succeed in balancing the means of each subset of items.

based ECD. This interpretation also echoes the subgroup analysis of Table 5, in which we found the same pattern among more advantaged children in the community-based ECD experiment. The results also match broader lessons from the literature on the heterogeneous effects of pre-school by socioeconomic status.

Returning to column (1) of Table 7, the coefficient on the pooled ECD Annex group is -0.17, but not statistically significant, indicating no distinguishable difference in outcomes between observationally equivalent children in pure control and ECD Annex communities. However, a test of the difference between the ECD Annex and community-based ECD coefficients, reported at the bottom of the table, shows an advantage of 0.47 standard deviations for the ECD Annex group, significant at 1%. The Lee (2009) bounds for this estimate (reported in brackets) do not include zero, alleviating concerns that differential attrition between experiments drives the result.[17] This effect falls to 0.40 standard deviations and remains significant at 10% when controlling for baseline score in column (2). These results suggest that providing ECD services through annexes attached to primary schools is more effective at promoting language development than community-based ECD services, at least during our sample period.[18]

Table 7, columns (3)-(4) disaggregate the ECD Annex group by treatment and control, corresponding to equation (3). The ECD Annex treatment and control have nearly identical effects relative to both the pure control group (as shown by their coefficients) and to the community-based ECD treatment (as shown by the estimates reported at the bottom of the table). Bounds for these estimates are also nearly identical. These results suggest that the teacher training provided in the ECD Annex treatment added little value relative to the control, at least by the end of the sample period. Instead, it appears that other features of the ECD Annex environment explain its differences with the community-based ECD program.

Columns (5)-(8) of Table 7 repeat the exercise for fine motor skills. For this outcome, we find no statistically significant differences between the pure control group and the community-

[17] The Lee (2009) bounds apply to a pairwise comparison between groups, but continue to weight by the inverse propensity score and adjust for regional stratification as in the regressions. We cannot bound the effect while controlling for the baseline outcome because there must be variation in treatment within covariate cells, whereas normalized MDAT scores are continuous.

[18] The results are similar when removing sites that reported implementation problems, suggesting that implementation issues in community-based ECD centers do not drive this result. See Supplemental Appendix Table SA6.

based ECD treatment. Nor are there significant differences between the pure control and the ECD Annex sample, either pooled or separately by treatment status. Children in the ECD Annex group do score significant higher than those in the community-based treatment, however, with magnitudes ranging from 0.35-0.42 standard deviations, though the effects lose significance when controlling for baseline outcome. As with language and hearing, for fine motor skills we again see no discernible benefit to the ECD Annex treatment relative to the control.

6. Implementation and Qualitative comparison

As stated earlier, while the outcome of the two impact evaluations in the short term help inform on quality, they cannot serve as the sole basis for decisions on appropriate scale-up. In this section, we bring together the implementation challenges and several qualitative lessons from the study to support the decision.

6.1 **Take up**

The demand for professional early childhood development program is high in The Gambia. By its nature, community-based ECD is built in the community, which provides local residents with higher accessibility. Parents can send their kids to the center next to their house and go to work on chores or on farms. Accordingly, we observed very high take-up in the community-based ECD. The two key driving factors are the presence of the center within the community and the facilitator being also from within the community. Most primary school children walk to schools by themselves and often in small groups. Since younger children will require parental drop off and pick up, it limits the take-up of annexes. When the services were offered within the community, it was oversubscribed almost systematically. This suggests that the community-based approach would be more effective for broader access. However, it has several limitations all of which pertained to quality.

6.2 **Staffing challenges**

Although community based ECD has the advantage of mitigating the communication barriers and issues of trust, in that facilitators are able to speak local language and live in the community, most communities lack qualified staff to play this role. Whereas a minimum of high school

degree was required from the staff, that requirement needed to be relaxed in numerous occasions, as one could not find candidates who met the criteria. By design, the community-based ECD is prone to introduce more inequality along the quality of the services provided as more well off and educated communities would be more likely to find resolve the staffing issue. The community based ECD has higher probability, compared with annex ECD center, to hire facilitators who have lower qualification.

6.3 **Quality insurance & monitoring**

While the governance structure of primary schools would directly apply to annexes, the same cannot be said of the community-based approach. The monitoring and quality control, salary payment, schedule management would require a new administrative layer for the community-based. The management cost would be significantly lower for the annexes while minimizing the equity issues. Given the differences in populations served by community-based ECD versus Annexes, expanding the community-based approach risks providing second-tier pre-school services to the most disadvantaged children.

6.4 **Peer effect and enabling environment**

The annexes offer an environment with large and diverse pool of peers for both the children and their teacher. Younger kids going to annex ECD centers may have interactions with older kids in the school and get exposure to school setting, which would also make them comfortable during the transition (from pre-school to primary school).

7. Conclusion

This paper evaluated two experiments to improve early childhood development services in The Gambia. The first experiment focused on increasing access to ECD services by constructing community-based centers in communities where no structured ECD program existed. The second experiment focused on improving quality of existing ECD centers by providing training teachers to deliver a new curriculum. We found no evidence that either intervention affected average levels of child development, as measured by tests of fine motor

skills or language and hearing. We find heterogeneous effects according to baseline household characteristics, however. Children from more advantaged households developed language skills more slowly when exposed to the community-based ECD treatment. In the ECD Annex experiment, children with low initial skill levels or whose mothers experienced greater mental distress at baseline improved their fine motor skills in response to the program.

This pattern of heterogeneous effects by socioeconomic status is consistent with the literature for both developed and developing countries, which tend to find that disadvantaged children benefit most from quality early childhood development programs. The explanation likely lies in the substitutability between endowments and investments for children at early ages (Heckman and Mosso 2014). This substitutability implies that increasing human capital investments for children from disadvantaged homes, i.e., those with low initial endowments, is more productive than for children from more advantaged backgrounds. More formally, at early ages:

$$\frac{\partial^2 \theta_{t+1}}{\partial \theta_t I_t} \leq 0$$

where θ represents endowments, *I* represents investment, and *t* indexes time (Heckman and Mosso 2014, p. 697). The negative sign on this cross-partial derivative indicates substitutability between θ and *I*. Our findings are consistent with this property of early life investments.

Comparing outcomes across experiments among observationally similar children reveals that providing ECD services in existing facilities attached to primary schools is more effective than expanding access through community-based centers. This result has important implications for policymakers facing a tradeoff between promoting ECD access and quality using limited funds. A caveat to our results, however, is that they apply to newly constructed community-based centers in their first year of operation. The gap between community-based ECD centers and primary school annexes might close over time if the community-based centers improve with experience.

An additional caveat to our results is that they are short-term, as they focus on young children less than two years after treatment assignment. Given the importance of early childhood circumstances for adult outcomes, longer-term measures are required to account fully for the effects of the interventions studied here.

References

Abadie, Alberto. 2005. “Semiparametric Difference-in-Differences Estimators.” *The Review of Economic Studies* 72 (1): 1–19.

Aboud, Frances E. 2006. “Evaluation of an Early Childhood Preschool Program in Rural Bangladesh.” *Early Childhood Research Quarterly* 21 (1): 46–60.

Akinocho, Herve, Moussa Blimpo, Pedro Carneiro, and Nathalie Lahire. 2014. “Improving Parental Investments in Children: Experimental Evidence from The Gambia.”

Angrist, Joshua D., and Victor Lavy. 2001. “Does Teacher Training Affect Pupil Learning? Evidence from Matched Comparisons in Jerusalem Public Schools.” *Journal of Labor Economics* 19 (2): 343–69. doi:http://www.jstor.org/action/showPublication?journalCode=jlaboreconomics.

Araujo, M. Caridad, Pedro Carneiro, Yyannú Cruz-Aguayo, and Norbert Schady. 2016. “Teacher Quality and Learning Outcomes in Kindergarten.” *The Quarterly Journal of Economics* 131 (3): 1415–1453.

Armecin, Graeme, Jere Behrman, Paulita Duazo, Sharon Ghuman, Socorro Gultiano, Elizabeth King, and Nannette Lee. 2006. “Early Childhood Development through an Integrated Program: Evidence from the Philippines.” *World Bank Policy Research Working Paper*, no. 3922. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=917493.

Baker-Henningham, Helen, Stephen Scott, Kelvyn Jones, and Susan Walker. 2012. “Reducing Child Conduct Problems and Promoting Social Skills in a Middle-Income Country: Cluster Randomised Controlled Trial.” *The British Journal of Psychiatry* 201 (2): 101–108.

Bastos, Paulo, Nicolas L. Bottan, and Julian Cristia. 2017. “Access to Preprimary Education and Progression in Primary School: Evidence from Rural Guatemala.” *Economic Development and Cultural Change*, February, 000–000. doi:10.1086/691090.

Behrman, Jere, Lia Fernald, and Patrice Engle. 2013. “Preschool Programs in Developing Countries.” *Education Policy in Developing Countries*, 65–105.

Behrman, Jere R., Yingmei Cheng, and Petra E. Todd. 2004. “Evaluating Preschool Programs When Length of Exposure to the Program Varies: A Nonparametric Approach.” *Review of Economics and Statistics* 86 (1): 108–132.

Berlinski, Samuel, and Sebastian Galiani. 2007. “The Effect of a Large Expansion of Pre-Primary School Facilities on Preschool Attendance and Maternal Employment.” *Labour Economics* 14 (3): 665–680.

Berlinski, Samuel, Sebastian Galiani, and Paul Gertler. 2009. “The Effect of Pre-Primary Education on Primary School Performance.” *Journal of Public Economics* 93 (1–2): 219–34. doi:http://www.sciencedirect.com/science/journal/00472727.

Berlinski, Samuel, Sebastian Galiani, and Marco Manacorda. 2008. “Giving Children a Better Start: Preschool Attendance and School-Age Profiles.” *Journal of Public Economics* 92 (5–6): 1416–40. doi:http://www.sciencedirect.com/science/journal/00472727.

Bernal, Raquel. 2015. “The Impact of a Vocational Education Program for Childcare Providers on Children’s Well-Being.” *Economics of Education Review* 48 (October): 165–83. doi:http://www.sciencedirect.com/science/journal/02727757.

Bernal, Raquel, and Camila Fernández. 2013. “Subsidized Childcare and Child Development in Colombia: Effects of Hogares Comunitarios de Bienestar as a Function of Timing and Length of Exposure.” *Social Science & Medicine* 97: 241–249.

Bold, Tessa, Mwangi Kimenyi, Germano Mwabu, Alice Ng'ang'a, and Justin Sandefur. 2013. "Scaling up What Works: Experimental Evidence on External Validity in Kenyan Education." *Center for Global Development Working Paper*, no. 321. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2241240.

Bouguen, Adrien, Deon Filmer, Karen Macours, and Sophie Naudeau. 2014. "Preschools and Early Childhood Development in a Second Best World: Evidence from a Scaled-up Experiment in Cambodia." doi:http://www.cepr.org/active/publications/discussion_papers/dp.php?dpno=10170.

Bowne, Jocelyn Bonnes, Hirokazu Yoshikawa, and Catherine E. Snow. 2016. "Experimental Impacts of a Teacher Professional Development Program in Early Childhood on Explicit Vocabulary Instruction across the Curriculum." *Early Childhood Research Quarterly* 34: 27–39.

Cunha, Flavio, and James Heckman. 2007. "The Technology of Skill Formation." *The American Economic Review* 97 (2): 31–47.

Cunha, Flavio, James J. Heckman, and Susanne M. Schennach. 2010. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica* 78 (3): 883–931.

Currie, Janet, and Douglas Almond. 2011. "Human Capital Development before Age Five." *Handbook of Labor Economics* 4: 1315–1486.

Duflo, Esther, Pascaline Dupas, Michael Kremer, and Samuel Sinei. 2006. "Education and HIV/AIDS Prevention: Evidence from a Randomized Evaluation in Western Kenya." *World Bank Policy Research Working Paper*, no. 4024. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=935173.

Elango, Sneha, Jorge Luis García, James J. Heckman, and Andrés Hojman. 2015. "Early Childhood Education." SSRN Scholarly Paper ID 2696852. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2696852.

Engle, Patrice L., Maureen M. Black, Jere R. Behrman, Meena Cabral De Mello, Paul J. Gertler, Lydia Kapiriri, Reynaldo Martorell, Mary Eming Young, and International Child Development Steering Group. 2007. "Strategies to Avoid the Loss of Developmental Potential in More than 200 Million Children in the Developing World." *The Lancet* 369 (9557): 229–242.

Engle, Patrice L., Lia CH Fernald, Harold Alderman, Jere Behrman, Chloe O'Gara, Aisha Yousafzai, Meena Cabral de Mello, et al. 2011. "Strategies for Reducing Inequalities and Improving Developmental Outcomes for Young Children in Low-Income and Middle-Income Countries." *The Lancet* 378 (9799): 1339–1353.

Evans, David, and Anna Popova. 2015. "What Really Works to Improve Learning in Developing Countries? An Analysis of Divergent Findings in Systematic Reviews." *An Analysis of Divergent Findings in Systematic Reviews (February 26, 2015). World Bank Policy Research Working Paper*, no. 7203. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2571082.

Firpo, Sergio. 2007. "Efficient Semiparametric Estimation of Quantile Treatment Effects." *Econometrica* 75 (1): 259–276.

Gertler, Paul, James Heckman, Rodrigo Pinto, Arianna Zanolini, Christel Vermeersch, Susan Walker, Susan M. Chang, and Sally Grantham-McGregor. 2014. "Labor Market Returns to an Early Childhood Stimulation Intervention in Jamaica." *Science* 344 (6187): 998–1001.

Giordono, Leanne, and Todd Pugatch. 2017. “Non-Tuition Costs, School Access and Student Performance: Evidence from the Gambia.” *Journal of African Economies*. http://jae.oxfordjournals.org/content/early/2017/01/05/jae.ejw033.abstract.

Gladstone, Melissa, Gillian A. Lancaster, Eric Umar, Maggie Nyirenda, Edith Kayira, Nynke R. van den Broek, and Rosalind L. Smyth. 2010. “The Malawi Developmental Assessment Tool (MDAT): The Creation, Validation, and Reliability of a Tool to Assess Child Development in Rural African Settings.” *PLOS Med* 7 (5): e1000273. doi:10.1371/journal.pmed.1000273.

Glewwe, Paul W., Eric A. Hanushek, Sarah D. Humpage, and Renato Ravina. 2011. “School Resources and Educational Outcomes in Developing Countries: A Review of the Literature from 1990 to 2010.” Working Paper 17554. National Bureau of Economic Research. http://www.nber.org/papers/w17554.

Grantham-McGregor, Sally, Yin Bun Cheung, Santiago Cueto, Paul Glewwe, Linda Richter, Barbara Strupp, and International Child Development Steering Group. 2007. “Developmental Potential in the First 5 Years for Children in Developing Countries.” *The Lancet* 369 (9555): 60–70.

Grantham-McGregor, Sally M., Lia CH Fernald, Rose Kagawa, and Susan Walker. 2014. “Effects of Integrated Child Development and Nutrition Interventions on Child Development and Nutritional Status.” *Annals of the New York Academy of Sciences* 1308 (1): 11–32.

Hadley, Craig, Ayalew Tegegn, Fasil Tessema, Makonnen Asefa, and Sandro Galea. 2008. “Parental Symptoms of Common Mental Disorders and Children’s Social, Motor, and Language Development in Sub-Saharan Africa.” *Annals of Human Biology* 35 (3): 259–275.

He, Fang, Leigh L. Linden, and Margaret MacLeod. 2009. “A Better Way to Teach Children to Read? Evidence from a Randomized Controlled Trial.” *Unpublished Manuscript. New York, NY: Columbia University*. http://leighlinden.com/Teach%20Children%20to%20Read.pdf.

Heckman, James J. 2006. “Skill Formation and the Economics of Investing in Disadvantaged Children.” *Science* 312 (5782): 1900–1902.

Heckman, James J., and Stefano Mosso. 2014. “The Economics of Human Development and Social Mobility.” *Annu. Rev. Econ* 6: 689–733.

Hirano, Keisuke, Guido W. Imbens, and Geert Ridder. 2003. “Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score.” *Econometrica* 71 (4): 1161–1189.

Imbens, Guido W., and Jeffrey M. Wooldridge. 2009. “Recent Developments in the Econometrics of Program Evaluation.” *Journal of Economic Literature* 47 (1): 5–86. doi:http://www.aeaweb.org/jel/index.php.

Krafft, Caroline. 2015. “Increasing Educational Attainment in Egypt: The Impact of Early Childhood Care and Education.” *Economics of Education Review* 46: 127–143.

Lee, David S. 2009. “Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects.” *The Review of Economic Studies* 76 (3): 1071–1102.

Malmberg, Lars-Erik, Peter Mwaura, and Kathy Sylva. 2011. “Effects of a Preschool Intervention on Cognitive Development among East-African Preschool Children: A Flexibly Time-Coded Growth Model.” *Early Childhood Research Quarterly* 26 (1): 124–33. doi:10.1016/j.ecresq.2010.04.003.

Martinez, Sebastian, Sophie Naudeau, and Vitor Pereira. 2013. “The Promise of Preschool in Africa: A Randomized Impact Evaluation of Early Childhood Development in Rural Mozambique,” November. http://www.fucape.br/downloads/Artigo%20Vitor%20Azevedo.pdf.

Murnane, Richard J., and Alejandro J. Ganimian. 2014. “Improving Educational Outcomes in Developing Countries: Lessons from Rigorous Evaluations.” Working Paper 20284. National Bureau of Economic Research. http://www.nber.org/papers/w20284.

Mwaura, Peter A.M., Kathy Sylva, and Lars-Erik Malmberg. 2008. “Evaluating the Madrasa Preschool Programme in East Africa: A Quasi-experimental Study.” *International Journal of Early Years Education* 16 (3): 237–55. doi:10.1080/09669760802357121.

Nitsaisook, Malee, and Lorin W. Anderson. 1989. “An Experimental Investigation of the Effectiveness of Inservice Teacher Education in Thailand.” *Teaching and Teacher Education* 5 (4): 287–302.

Nores, Milagros, and W. Steven Barnett. 2010. “Benefits of Early Childhood Interventions across the World: (Under) Investing in the Very Young.” *Economics of Education Review*, Special Issue in Honor of Henry M. Levin, 29 (2): 271–82. doi:10.1016/j.econedurev.2009.09.001.

Ozler, Berk, Lia C. H. Fernald, Patricia Karol Kariger, Christin Mcconnell, Michelle J. Neuman, and Eduardo Fraga. 2016. “Combining Preschool Teacher Training with Parenting Education: A Cluster-Randomized Controlled Trial.” SSRN Scholarly Paper ID 2849117. Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=2849117.

Pinto, Cristine Campos de Xavier, Daniel Santos, and Clarissa Guimarães. 2016. “The Impact of Daycare Attendance on Math Test Scores for a Cohort of Fourth Graders in Brazil.” *The Journal of Development Studies*, 1–23.

Piper, Benjamin, and Medina Korda. 2011. “EGRA Plus: Liberia. Program Evaluation Report.” *RTI International*. http://eric.ed.gov/?id=ED516080.

Pritchett, Lant. 2013. *The Rebirth of Education: Schooling Ain’t Learning*. Washington, D.C: Center for Global Development.

Raine, Adrian, Jianghong Liu, Peter H. Venables, Sarnoff A. Mednick, and C. Dalais. 2010. “Cohort Profile: The Mauritius Child Health Project.” *International Journal of Epidemiology* 39 (6): 1441–1451.

Raine, Adrian, Kjetil Mellingen, Jianghong Liu, Peter Venables, and Sarnoff A. Mednick. 2003. “Effects of Environmental Enrichment at Ages 3-5 Years on Schizotypal Personality and Antisocial Behavior at Ages 17 and 23 Years.” *American Journal of Psychiatry* 160 (9): 1627–1635.

Raine, Adrian, Peter H. Venables, Cyril Dalais, Kjetil Mellingen, Chandra Reynolds, and Sarnoff A. Mednick. 2001. “Early Educational and Health Enrichment at Age 3–5 Years Is Associated with Increased Autonomic and Central Nervous System Arousal and Orienting at Age 11 Years: Evidence from the Mauritius Child Health Project.” *Psychophysiology* 38 (02): 254–266.

Raine, Adrian, Peter H. Venables, and Sarnoff A. Mednick. 1997. “Low Resting Heart Rate at Age 3 Years Predisposes to Aggression at Age 11 Years: Evidence from the Mauritius Child Health Project.” *Journal of the American Academy of Child & Adolescent Psychiatry* 36 (10): 1457–1464.

Rao, Nirmala, Jin Sun, Veronica Pearson, Emma Pearson, Hongyun Liu, Mark A. Constas, and Patrice L. Engle. 2012. “Is Something Better than Nothing? An Evaluation of Early Childhood Programs in Cambodia.” *Child Development* 83 (3): 864–876.

Robins, James M., and Andrea Rotnitzky. 1995. “Semiparametric Efficiency in Multivariate Regression Models with Missing Data.” *Journal of the American Statistical Association* 90 (429): 122–129.

Rosero, Jose. 2012. “On the Effectiveness of Child Care Centers in Promoting Child Development in Ecuador.” Tinbergen Institute. http://papers.tinbergen.nl/12075.pdf.

Rosero, José, and Hessel Oosterbeek. 2011. “Trade-Offs between Different Early Childhood Interventions: Evidence from Ecuador.” http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1898566.

Taiwo, A. A., and J. B. Tyolo. 2002. “The Effect of Pre-School Education on Academic Performance in Primary School: A Case Study of Grade One Pupils in Botswana.” *International Journal of Educational Development* 22 (2): 169–180.

Tanner, Jeffrey C., Tara Candland, and Whitney S. Odden. 2015. “Later Impacts of Early Childhood Interventions: A Systematic Review,” World Bank Independent Evaluation Group Working Paper, , no. 3. https://ieg.worldbankgroup.org/Data/reports/ImpactsofInterventions_EarlyChildhoodonLaterOutcomes.pdf.

Tough, Paul. 2016. *Helping Children Succeed: What Works and Why*. Houghton Mifflin Harcourt.

Walker, Susan P., Theodore D. Wachs, Julie Meeks Gardner, Betsy Lozoff, Gail A. Wasserman, Ernesto Pollitt, Julie A. Carter, and International Child Development Steering Group. 2007. “Child Development: Risk Factors for Adverse Outcomes in Developing Countries.” *The Lancet* 369 (9556): 145–157.

Walker, Susan P., Theodore D. Wachs, Sally Grantham-McGregor, Maureen M. Black, Charles A. Nelson, Sandra L. Huffman, Helen Baker-Henningham, et al. 2011. “Inequality in Early Childhood: Risk and Protective Factors for Early Child Development.” *The Lancet* 378 (9799): 1325–1338.

Wolf, Sharon, J. Lawrence Aber, and Jere R. Behrman. 2017. “The Impacts of Teacher Training and Parental Education on Kindergarten Quality in Ghana.” Innovations for Poverty Action Study Summary. https://www.poverty-action.org/printpdf/20016.

World Bank. 2015. “World Development Indicators.” World Bank. http://data.worldbank.org/data-catalog/world-development-indicators.

———. 2016. “The Gambia: Country at a Glance.” http://www.worldbank.org/en/country/gambia.

Yoshikawa, Hirokazu, Diana Leyva, Catherine E. Snow, Ernesto Treviño, M. Barata, Christina Weiland, Celia J. Gomez, et al. 2015. “Experimental Impacts of a Teacher Professional Development Program in Chile on Preschool Classroom Quality and Child Outcomes.” *Developmental Psychology* 51 (3): 309.

Zhang, Linxiu, Fang Lai, Xiaopeng Pang, Hongmei Yi, and Scott Rozelle. 2013. “The Impact of Teacher Training on Teacher and Student Outcomes: Evidence from a Randomised Experiment in Beijing Migrant Schools.” *Journal of Development Effectiveness* 5 (3): 339–358.

Zoyem, Jean-Paul. 2010. “Design and Support to the Impact Evaluation for the Early Childhood Care Development (ECCD) in Gambia: Inception Report.”

Zuilkowski, Stephanie Simmons, Gunther Fink, Corrina Moucheraud, and Beatrice Matafwali. 2012. “Early Childhood Education, Child Development and School Readiness: Evidence from Zambia.” *South African Journal of Childhood Education* 2 (2): 20.

Figure 1: Sample locations

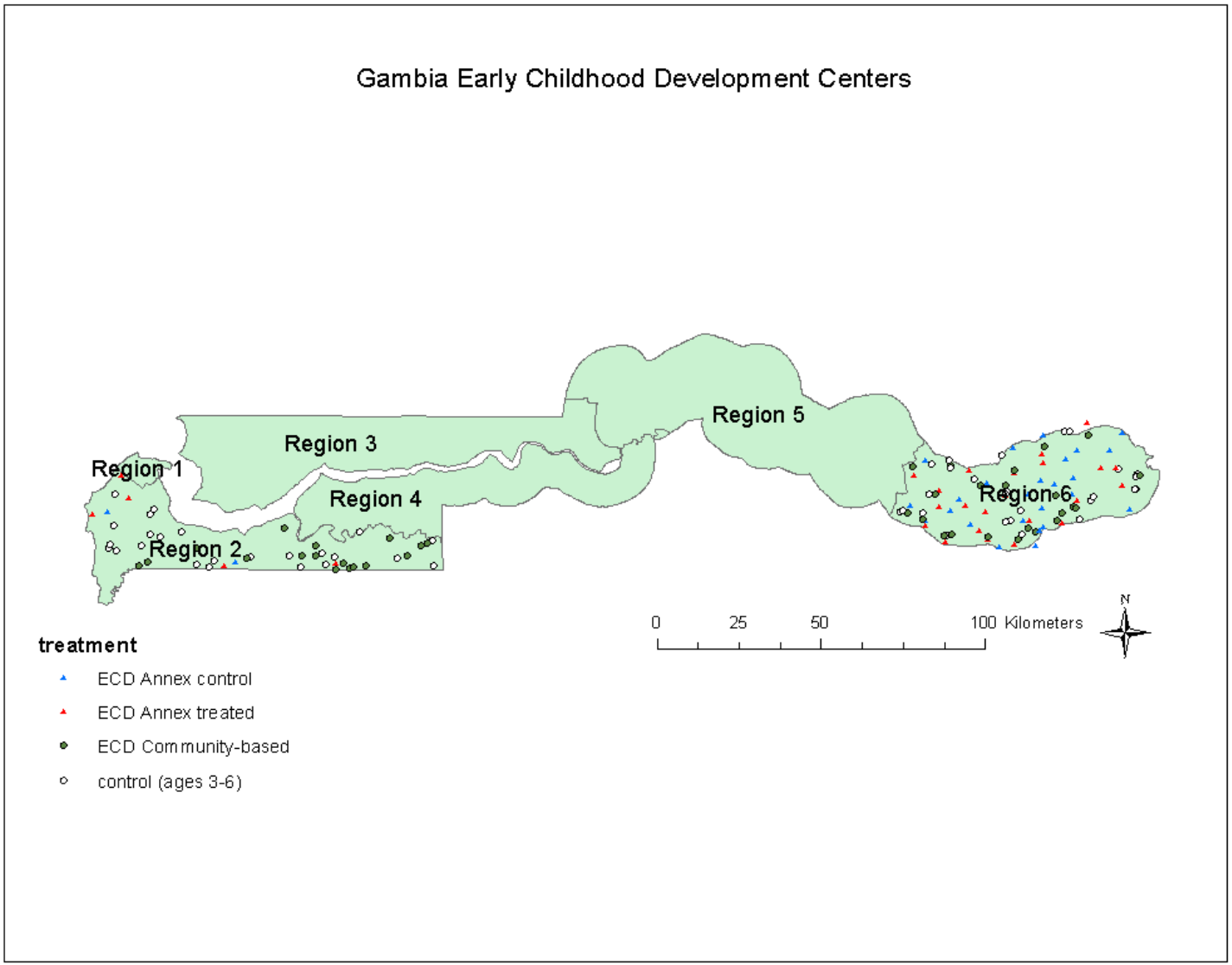


Figure 2: Project and research timeline

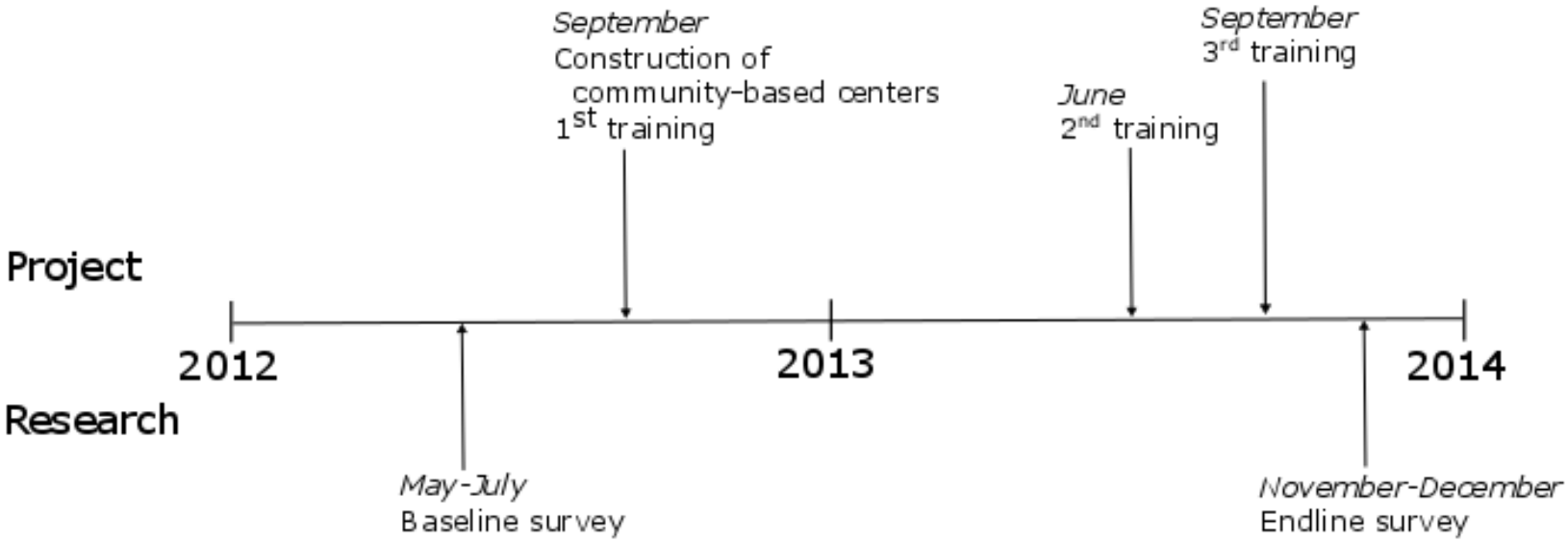
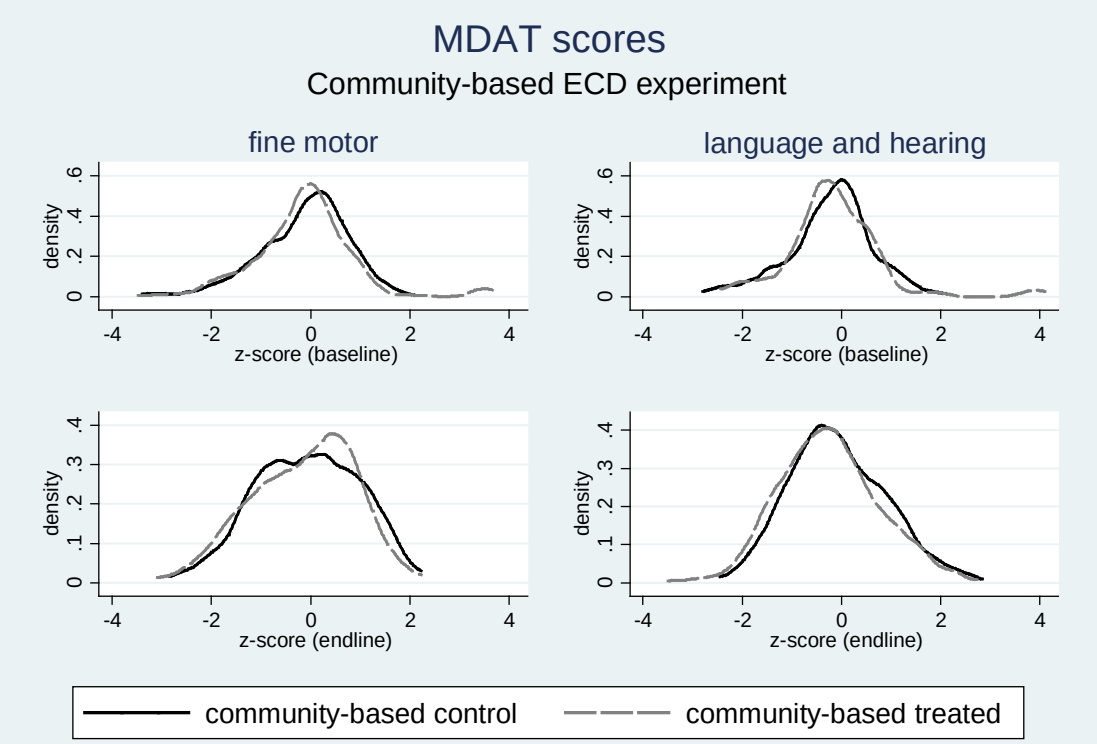
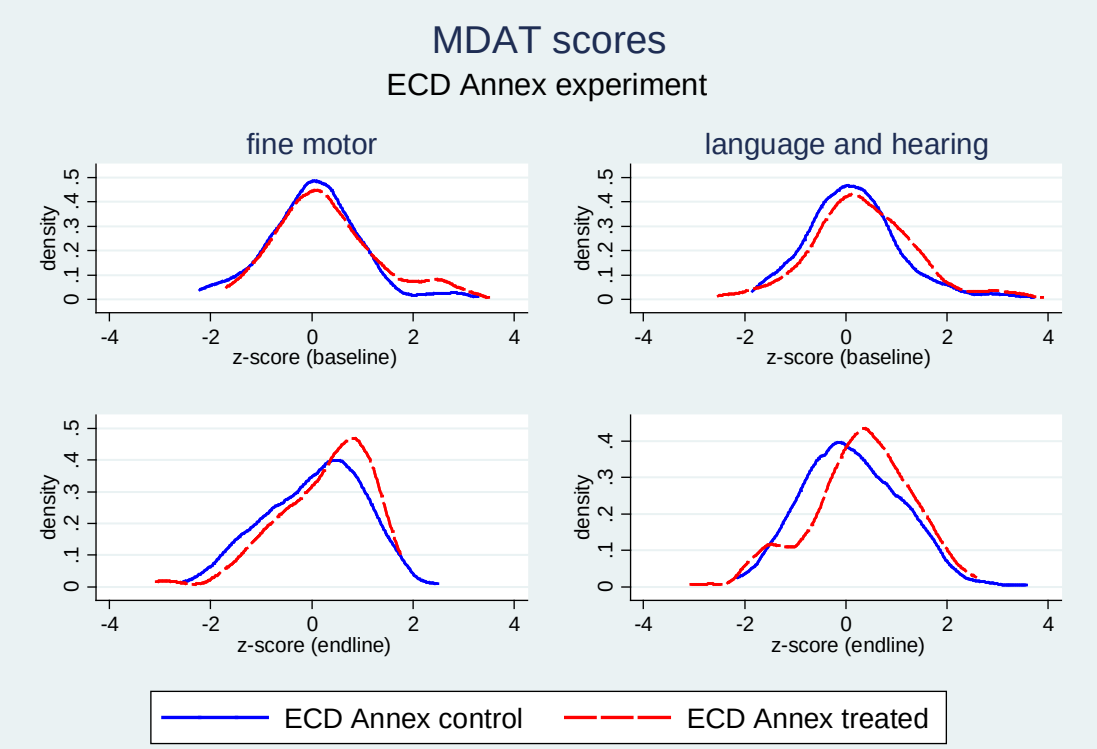


Figure 3(a): Outcome distributions, Community-based ECD experiment



Scores adjusted by age, age squared, and gender.

Figure 3(b): Outcome distributions, ECD Annex experiment



Scores adjusted by age, age squared, and gender.

Figure 4(a): Quantile treatment effects, Community-based ECD experiment

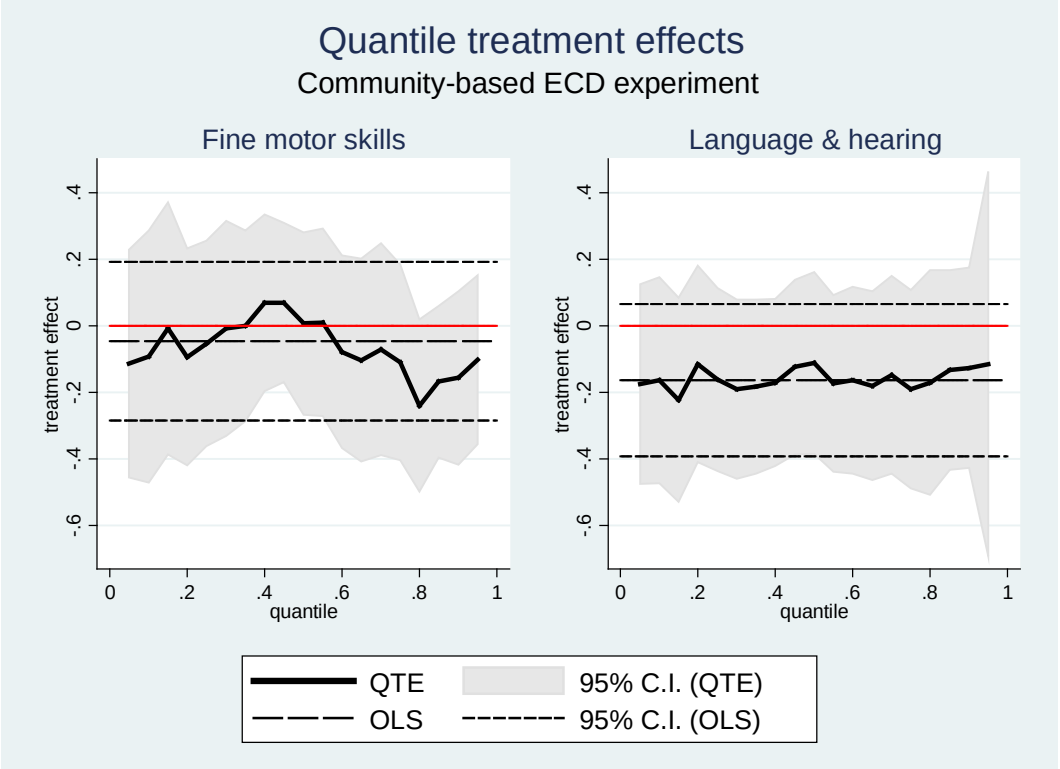


Figure 4(b): Quantile treatment effects, ECD Annex experiment

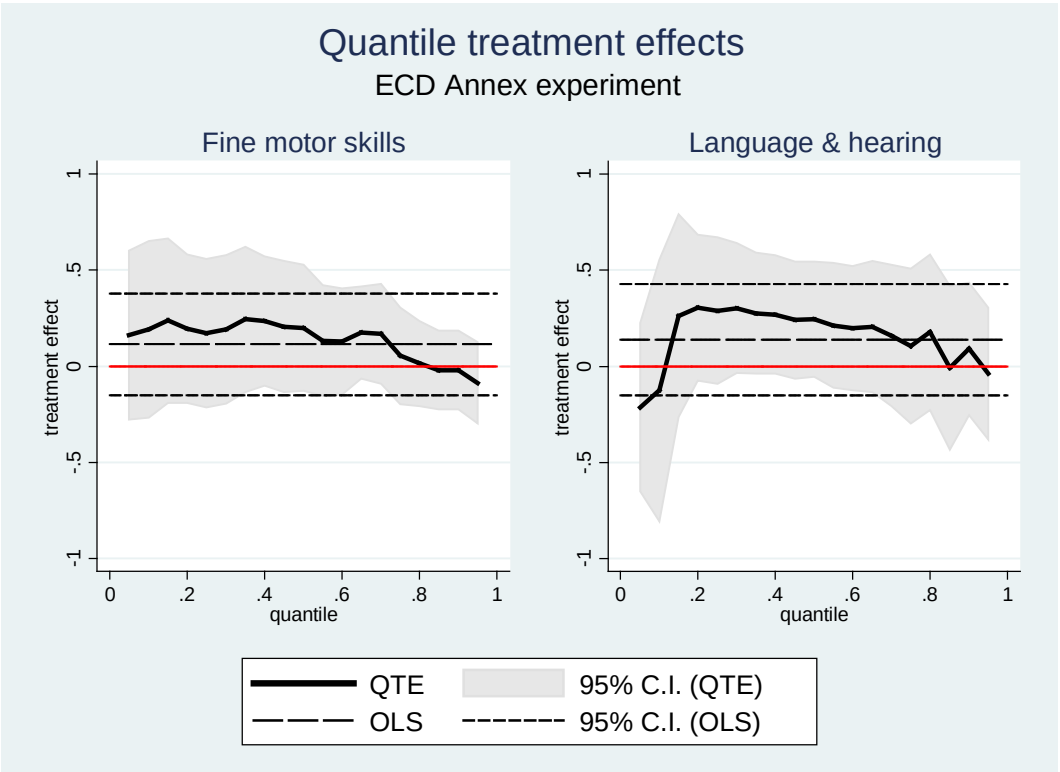


Table 1: Experimental design

Experiment	Sample	Control	Treatment
1	Villages without pre-existing ECD services, Regions 2/6	Pure control • No structured ECD services	Community-based • Construction of ECD center • GOALS curriculum • Teacher training • Management committee
2	Villages with ECD Annex centers, Regions 2/6	ECD Annex control • GOALS curriculum	ECD Annex treatment • GOALS curriculum • Teacher training

Table 2: Sample sizes and attrition

	(1)	(2)	(3)	(4)	(5)
Panel A: sample sizes	interview		test score		
	baseline	endline	baseline	endline	both
Community-based ECD experiment					
control					
original sample	606	481	319	270	243
added sample		133		75	
treatment					
original sample	441	356	267	226	204
added sample		87		55	
ECD Annex experiment					
control					
original sample	365	322	192	183	170
added sample		44		24	
treatment					
original sample	351	326	182	165	162
added sample		40		23	
Panel B: attrition	control	treatment	difference	p-value	
Community-based ECD experiment					
from endline survey	0.21	0.19	-0.01	0.99	
	(0.03)	(0.04)	(0.05)		
from endline test	0.24	0.24	0.00	0.82	
	(0.04)	(0.04)	(0.06)		
number of communities	51	40			
ECD Annex experiment					
from endline survey	0.12	0.07	-0.05	0.11	
	(0.02)	(0.01)	(0.03)		
from endline test	0.12	0.11	-0.01	0.83	
	(0.03)	(0.02)	(0.03)		
number of communities	27	26			

Panel A shows sample sizes by survey wave and treatment status. Original sample refers to those present at baseline. Added sample refers to new subjects added at endline who were not present at baseline. Interview refers to completed interview. Test score refers to completed MDAT test for fine motor skills and language/hearing. Panel B shows attrition rates by treatment group. Attrited from endline is indicator for not being present for endline interview, conditional on being present for baseline interview. Attrited from endline test is indicator for not being present for endline test, conditional on taking baseline test. Standard errors in parentheses, clustered by settlement. p-values adjusted for stratification of treatment by region.

Table 3: Baseline characteristics and balance tests

Experiment	Community-based ECD				ECD Annex			
Group	control	treatment	difference	p-value	control	treatment	difference	p-value
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
age (months)	47.2	47.3	0.1	0.97	59.2	57.9	-1.3	0.58
	(0.3)	(0.4)	(0.5)		(1.1)	(1.2)	(1.6)	
female	0.49	0.44	-0.05	0.14	0.45	0.49	0.05	0.29
	(0.02)	(0.02)	(0.03)		(0.03)	(0.03)	(0.04)	
attend ECD	0.39	0.23	-0.16	0.01	0.86	0.90	0.04	0.58
	(0.04)	(0.04)	(0.06)		(0.04)	(0.03)	(0.05)	
fine motor skills	-0.12	-0.13	0.00	0.78	0.04	0.29	0.25	0.33
	(0.08)	(0.13)	(0.15)		(0.12)	(0.16)	(0.20)	
language and hearing skills	-0.22	-0.15	0.07	0.47	0.17	0.36	0.19	0.45
	(0.08)	(0.13)	(0.15)		(0.12)	(0.16)	(0.19)	
height-for-age	-1.38	-1.55	-0.17	0.19	-1.37	-1.28	0.09	0.64
	(0.08)	(0.09)	(0.12)		(0.12)	(0.21)	(0.24)	
household size	7.5	7.9	0.4	0.31	7.8	8.0	0.2	0.53
	(0.2)	(0.3)	(0.4)		(0.2)	(0.3)	(0.4)	
mother's schooling	1.7	1.7	0.0	0.78	1.4	2.0	0.5	0.28
	(0.2)	(0.3)	(0.3)		(0.2)	(0.3)	(0.4)	
household head schooling	2.3	1.9	-0.4	0.74	2.1	1.9	-0.2	0.09
	(0.2)	(0.3)	(0.4)		(0.3)	(0.4)	(0.5)	
household head employed	0.83	0.84	0.01	0.67	0.83	0.76	-0.06	0.18
	(0.03)	(0.03)	(0.04)		(0.04)	(0.04)	(0.06)	
household head work hours	38.6	38.7	0.1	0.98	38.7	35.6	-3.1	0.36
	(2.3)	(2.3)	(3.3)		(2.3)	(3.0)	(3.8)	
household head in agriculture	0.55	0.68	0.13	0.08	0.59	0.56	-0.02	0.82
	(0.04)	(0.05)	(0.06)		(0.04)	(0.05)	(0.06)	
household expenditure per capita	483.4	510.4	27.1	0.67	505.2	450.6	-54.6	0.36
	(72.4)	(51.8)	(88.6)		(43.0)	(36.2)	(55.7)	
asset index	-0.08	-0.35	-0.27	0.29	0.08	0.66	0.58	0.15

	(0.20)	(0.17)	(0.26)		(0.24)	(0.30)	(0.38)	
ECD willingness to pay	0.10	0.17	0.07	0.10	0.14	0.16	0.02	0.91
(as % of household expenditure)	(0.01)	(0.05)	(0.05)		(0.07)	(0.05)	(0.08)	
vaccinations (% of 17)	0.54	0.58	0.04	0.24	0.52	0.54	0.03	0.67
	(0.02)	(0.02)	(0.03)		(0.03)	(0.03)	(0.04)	
child ill	0.14	0.17	0.03	0.36	0.18	0.16	-0.02	0.65
	(0.02)	(0.02)	(0.03)		(0.02)	(0.03)	(0.03)	
mother mental distress	44.8	42.2	-2.6	0.32	47.5	45.0	-2.5	0.55
(0-100 scale, where 100 is worst)	(1.7)	(2.3)	(2.8)		(2.8)	(3.0)	(4.1)	
stimulating objects (% of 10)	0.27	0.27	0.01	0.64	0.29	0.30	0.01	0.75
	(0.01)	(0.01)	(0.01)		(0.01)	(0.01)	(0.02)	
corporal punishment (% use)	0.69	0.67	-0.02	0.64	0.63	0.69	0.05	0.48
	(0.03)	(0.03)	(0.04)		(0.04)	(0.03)	(0.05)	
Observations								
children	606	441			365	351		
sites	49	38			27	26		

All variables are means from baseline survey. Standard errors in parentheses, clustered by settlement. p-values adjusted for stratification of treatment by region. Fine motor, language and hearing skills are z-scores from MDAT (adjusted scores based on standardized residuals from regression of raw score on child's age, age squared, and female dummy). Height-for-age z-score based on World Health Organization 2007 benchmark. Household expenditure per capita is annual value in USD, winsorized at 1st/99th percentiles. Asset index is first principal component of reported household assets. ECD willingness to pay is stated willingness to pay for early childhood development services as share of household per capital expenditure. Vaccinations is proportion of 17 vaccinations received by child. Mother mental distress is percentage of 11 mental health issues experienced "most of the time" by mother. Stimulating objects is proportion of 10 objects for stimulating play found in home. Corporal punishment is indicator for using corporal punishment as usual form of discipline for severe misbehavior.

Table 4: Endline outcomes

Experiment	Community-based ECD				ECD Annex			
Group	control	treatment	treatment effect without baseline outcome	treatment effect with baseline outcome	control	treatment	treatment effect without baseline outcome	treatment effect with baseline outcome
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Language and hearing								
overall score (z)	-0.07	-0.25	-0.16	-0.17	0.12	0.30	0.14	0.11
	(0.08)	(0.10)	(0.12)	(0.12)	(0.10)	(0.12)	(0.15)	(0.15)
knows own name & its letters (% of 4)	0.99	0.98	-0.01	-0.01	1.00	1.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
speaks in clear sentences	0.99	0.97	-0.02	-0.04	0.99	0.99	0.00	0.00
	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
counting (% of 3)	0.39	0.43	0.04	0.04	0.49	0.45	-0.04	-0.09**
	(0.02)	(0.03)	(0.04)	(0.04)	(0.02)	(0.02)	(0.03)	(0.03)
name colors (% of 4)	0.16	0.19	0.02	0.01	0.24	0.25	0.02	-0.01
	(0.01)	(0.02)	(0.03)	(0.03)	(0.01)	(0.02)	(0.02)	(0.03)
play with blocks (% of 3)	0.38	0.37	0.00	N/A	0.45	0.43	-0.02	N/A
	(0.02)	(0.02)	(0.03)		(0.02)	(0.02)	(0.03)	
open books (% of 3)	0.18	0.17	0.00	N/A	0.23	0.19	-0.04	N/A
	(0.01)	(0.02)	(0.02)		(0.02)	(0.02)	(0.03)	
Fine motor skills								
overall score (z)	-0.07	-0.15	-0.05	-0.04	0.06	0.24	0.11	0.16
	(0.10)	(0.12)	(0.12)	(0.14)	(0.10)	(0.12)	(0.14)	(0.14)
play with blocks (% of 6)	0.38	0.40	0.03	-0.01	0.42	0.40	-0.02	-0.02
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
draw lines & shapes (% of 6)	0.41	0.45	0.05	0.03	0.45	0.43	-0.02	-0.04
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
order rows of items (% of 2)	0.29	0.29	0.00	-0.01	0.39	0.34	-0.06	-0.08*
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.02)	(0.04)	(0.04)

fold paper	0.62	0.62	0.01	N/A	0.68	0.76	0.07	N/A
	(0.04)	(0.05)	(0.06)		(0.06)	(0.05)	(0.08)	
play with blocks	0.38	0.39	0.01	N/A	0.43	0.41	-0.02	N/A
(language and motors skills combined, % of 6)	(0.02)	(0.02)	(0.03)		(0.02)	(0.02)	(0.03)	
Observations								
children	345	281			207	188		
sites	50	40			27	26		

* significant at 10%, ** significant at 5%, *** significant at 1%. All variables are means from MDAT endline survey. Standard errors in parentheses, clustered by settlement. Treatment effects control for stratification of treatment assignment by region. Estimates in columns (4) and (8) augment regression with control for baseline outcome. z-scores are standardized residuals from regression of raw score on child's age, age squared, and female dummy. Other variables are subsets of items on MDAT test, measured as percent of items completed correctly. Speaks in clear sentences is just one item, while other categories have number of items indicated.

Table 5: Treatment effect heterogeneity

Subsample	sex		Household assets		Mother schooling		child sick last 3 days		mother mental distress		Stimulating objects in home		baseline z-score>0	
	male	female	below median	above median	did not attend	attended	no	yes	below median	above median	below median	above median	no	yes
Outcome	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A: Community-based ECD experiment														
without control for baseline outcome														
Fine motor skills	-0.01	-0.09	-0.02	-0.10	0.03	-0.30	-0.11	0.07	-0.02	-0.09	0.13	-0.25	0.02	-0.08
	(0.15)	(0.14)	(0.14)	(0.17)	(0.13)	(0.19)	(0.14)	(0.22)	(0.13)	(0.18)	(0.14)	(0.16)	(0.16)	(0.16)
Language and hearing	-0.09	-0.25	-0.06	-0.35	-0.11	-0.39	-0.30	0.03	-0.13	-0.20	-0.03	-0.36	-0.12	-0.17
	(0.14)	(0.14)*	(0.13)	(0.17)**	(0.12)	(0.20)*	(0.13)**	(0.25)	(0.14)	(0.15)	(0.13)	(0.17)**	(0.15)	(0.15)
with control for baseline outcome														
Fine motor skills	0.10	-0.20	-0.05	0.01	0.08	-0.33	-0.05	0.03	-0.01	-0.04	0.11	-0.14	0.02	-0.09
	(0.17)	(0.18)	(0.18)	(0.17)	(0.16)	(0.20)	(0.15)	(0.23)	(0.16)	(0.18)	(0.19)	(0.16)	(0.17)	(0.16)
Language and hearing	-0.05	-0.31	-0.14	-0.18	-0.10	-0.35	-0.23	0.05	-0.17	-0.15	-0.10	-0.24	-0.12	-0.17
	(0.14)	(0.16)**	(0.15)	(0.17)	(0.13)	(0.20)*	(0.13)*	(0.26)	(0.16)	(0.14)	(0.14)	(0.16)	(0.15)	(0.15)
Panel B: ECD Annex experiment														
without control for baseline outcome														
Fine motor skills	0.25	-0.04	0.02	0.22	0.09	0.17	0.22	-0.13	-0.04	0.36	0.06	0.20	0.34	0.06
	(0.17)	(0.15)	(0.17)	(0.19)	(0.15)	(0.21)	(0.15)	(0.23)	(0.17)	(0.17)**	(0.15)	(0.19)	(0.13)**	(0.21)
Language and hearing	0.26	0.01	0.12	0.16	0.12	0.18	0.17	-0.02	0.03	0.30	-0.05	0.39	0.01	0.16
	(0.17)	(0.20)	(0.16)	(0.20)	(0.16)	(0.25)	(0.15)	(0.31)	(0.20)	(0.19)	(0.14)	(0.22)*	(0.16)	(0.22)
with control for baseline outcome														
Fine motor skills	0.22	0.08	0.09	0.23	0.18	0.07	0.23	-0.10	-0.03	0.40	0.22	0.16	0.32	0.06
	(0.18)	(0.14)	(0.18)	(0.17)	(0.15)	(0.21)	(0.16)	(0.20)	(0.19)	(0.14)***	(0.13)*	(0.19)	(0.13)**	(0.20)
Language and hearing	0.16	0.07	0.09	0.14	0.13	0.09	0.15	-0.03	-0.04	0.31	-0.05	0.33	0.03	0.17
	(0.17)	(0.21)	(0.19)	(0.16)	(0.16)	(0.23)	(0.16)	(0.32)	(0.23)	(0.17)*	(0.13)	(0.21)	(0.16)	(0.21)

* significant at 10%, ** significant at 5%, *** significant at 1%. Each cell shows coefficient on treatment indicator from separate regression. Fine motor skills and language and hearing outcomes are MDAT z-scores, adjusted for age, age squared, and gender. All regressions include dummy for Region 2 to control for stratification. Baseline outcomes included in regressions reported where indicated. Sample split according to baseline characteristic indicated in column heading. Household expenditure measured per capita. Household assets based on first principal component of asset

ownership. Mother mental distress is proportion of responses to mental health questionnaire answering "3," which is "most of the time" for mental health issue. Stimulating objects in home based on proportion of 10 objects owned. Standard errors in parenthesis, clustered by settlement.

Table 6: Between-experiments baseline balance tests

	unweighted				weighted			
	community-based	ECD Annex	(1) vs. (2)	p-value	community-based	ECD Annex	(1) vs. (2)	p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Language and hearing								
overall score (z)	-0.19	0.27	0.58	0.00	0.33	0.09	-0.24	0.30
	(0.07)	(0.10)	(0.12)		(0.20)	(0.11)	(0.23)	
knows own name	0.13	0.19	0.06	0.00	0.18	0.17	-0.01	0.87
& its letters (% of 4)	(0.01)	(0.01)	(0.02)		(0.04)	(0.01)	(0.04)	
speaks in clear sentences	0.93	0.96	0.02	0.26	0.92	0.95	0.03	0.36
	(0.02)	(0.01)	(0.02)		(0.03)	(0.02)	(0.03)	
counting (% of 3)	0.09	0.25	0.17	0.00	0.23	0.19	-0.03	0.55
	(0.01)	(0.02)	(0.02)		(0.05)	(0.02)	(0.06)	
name colors (% of 4)	0.04	0.10	0.06	0.00	0.13	0.07	-0.07	0.15
	(0.01)	(0.02)	(0.02)		(0.04)	(0.01)	(0.05)	
Fine motor skills								
overall score (z)	-0.15	0.16	0.46	0.00	0.26	0.03	-0.23	0.33
	(0.08)	(0.10)	(0.13)		(0.22)	(0.11)	(0.24)	
play with blocks (% of 6)	0.35	0.36	0.03	0.11	0.41	0.36	-0.05	0.16
	(0.02)	(0.02)	(0.02)		(0.03)	(0.02)	(0.04)	
draw lines & shapes (% of 6)	0.35	0.40	0.07	0.01	0.44	0.39	-0.05	0.28
	(0.02)	(0.02)	(0.03)		(0.04)	(0.02)	(0.05)	
order rows of items (% of 2)	0.09	0.12	0.06	0.00	0.13	0.11	-0.02	0.67
	(0.01)	(0.02)	(0.02)		(0.06)	(0.02)	(0.06)	
Observations								
children	328	320			328	320		
sites	55	50			55	50		

All variables are means from baseline survey. Drops observations outside common support of propensity score distribution. Columns (5)-(8) weighted by inverse propensity score where indicated. Standard errors in parentheses, clustered by settlement. p-values obtained from regression of characteristic on community-based treatment and Region 2 dummy in order to adjust for stratification by region. Fine motor, language and hearing skills are z-scores from MDAT. Adjusted scores are standardized residuals from regression of raw score on child's age, age squared, and female dummy. Other variables are subsets of items on MDAT test, measured as percent of items completed correctly. Speaks in clear sentences is just one item, while other categories have number of items indicat

Table 7: Endline outcomes, combined experimental groups

MDAT module	Language and hearing				Fine motor			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Community-based ECD	-0.64	-0.37	-0.64	-0.36	-0.33	-0.14	-0.33	-0.14
	(0.28)**	(0.18)**	(0.28)**	(0.18)**	(0.25)	(0.21)	(0.25)	(0.21)
ECD Annex	-0.17	0.04			0.06	0.20		
	(0.29)	(0.22)			(0.21)	(0.16)		
ECD Annex control			-0.15	0.13			0.09	0.26
			(0.35)	(0.31)			(0.25)	(0.22)
ECD Annex treatment			-0.19	-0.07			0.02	0.12
			(0.27)	(0.20)			(0.21)	(0.15)
Observations								
Children	639	572	639	572	644	590	644	590
Sites	132	127	132	127	132	128	132	128
pure control group mean	0.49	0.49	0.49	0.49	0.23	0.23	0.23	0.23
	(1.14)	(1.14)	(1.14)	(1.14)	(0.99)	(0.99)	(0.99)	(0.99)
Includes baseline outcome		x		x		x		x
Difference from Community-based ECD								
ECD Annex	0.47***	0.40*			0.39**	0.34		
	(0.18)	(0.21)			(0.19)	(0.21)		
	[0.22,0.74]				[0.16,0.66]			
ECD Annex control			0.49*	0.49			0.42*	0.40
			(0.26)	(0.31)			(0.23)	(0.25)
			[0.15,0.74]				[0.42,0.76]	
ECD Annex treatment			0.45***	0.29*			0.35*	0.26
			(0.17)	(0.17)			(0.19)	(0.19)
			[0.15,0.68]				[0.11,0.58]	

* significant at 10%, ** significant at 5%, *** significant at 1%. Pure control group (i.e., control group from community-based ECD experiment) is omitted category. All regressions include region 2 dummy to adjust for stratification of treatment assignment and weight by inverse propensity score. Sample drops observations outside common support of propensity score distribution. Standard errors in parentheses, clustered by settlement. Outcomes are adjusted z-scores from MDAT modules for language and hearing and fine motor skills. Adjusted scores are standardized residuals from regression of raw score on child's

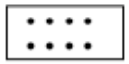
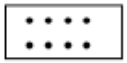
age, age squared, and female dummy. Regressions include baseline outcome where indicated. Differences with community-based ECD reported at bottom of table based on tests of indicated coefficient with community-based ECD. Lee bounds reported in brackets, based on pairwise comparison, but still reweighting by inverse propensity score and stratifying by region. Propensity score obtained from logit model of membership in ECD Annex sample regressed on baseline characteristics. Included baseline characteristics: age (exact based on DOB), female, Region 2, ECD attendance, fine motor skills (age-adjusted z-score), language and hearing (age-adjusted z-score), height-for-age, household size, mother's years of schooling, household expenditure per capita (winsorized at 1st/99th percentiles), willingness to pay for ECD as % of household expenditure per capita, % of vaccines received, mother mental distress (% of items reported as experiencing "most of the time"). Missing values imputed to zero, with dummies for imputed value included as additional covariates in regression.

Table A1: Regional characteristics, 2003 Census

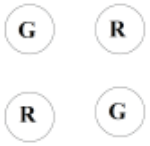
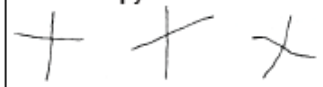
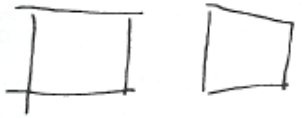
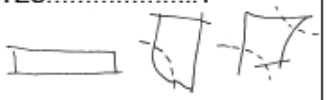
	Regions 1, 3-5	Region 2	Region 6	(1) vs. (2)	(1) vs. (3)	(1) vs. (2), p-value	(1) vs. (3), p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mandinka	0.21	0.17	0.29	0.04	-0.08	0.05	0.00
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)		
Fula	0.47	0.16	0.58	0.31	-0.11	0.00	0.00
	(0.01)	(0.01)	(0.02)	(0.02)	(0.03)		
Jola	0.01	0.44	0.00	-0.43	0.01	0.00	0.01
	(0.00)	(0.02)	(0.00)	(0.01)	(0.00)		
Islam	0.99	0.83	0.99	0.16	-0.01	0.00	0.22
	(0.00)	(0.02)	(0.00)	(0.01)	(0.01)		
Christianity	0.02	0.17	0.01	-0.15	0.01	0.00	0.21
	(0.00)	(0.02)	(0.00)	(0.01)	(0.01)		
never attended school	0.45	0.43	0.65	0.02	-0.21	0.10	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)		
currently attending school	0.25	0.38	0.20	-0.13	0.05	0.00	0.00
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)		
literate	0.39	0.49	0.25	-0.10	0.14	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)		
employed (18 and over)	0.78	0.63	0.78	0.15	0.00	0.00	0.77
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)		
children ever born	3.2	2.7	3.1	0.5	0.2	0.00	0.00
	(0.0)	(0.0)	(0.0)	(0.1)	(0.0)		
has electricity	0.02	0.03	0.01	-0.01	0.01	0.47	0.02
	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)		
piped water	0.10	0.12	0.14	-0.02	-0.03	0.27	0.04
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)		
population	648.6	1,132.5	496.2	-483.9	152.4	0.01	0.30
	(80.7)	(215.3)	(54.6)	(189.1)	(147.2)		
population 3-6 years old	82.2	140.7	72.3	-58.5	9.9	0.01	0.53
	(8.6)	(25.2)	(7.3)	(20.8)	(15.8)		
population 3-6 years old (%)	0.15	0.13	0.15	0.02	0.00	0.00	0.48
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		
Settlements	1,171	344	368	1,515	1,539		

Sample: settlements, grouped by regions as indicated. Data source: 2003 Census. Mandinka, Fula, and Jola are major ethnic groups. Standard errors in parenthesis.

Appendix A.1: MDAT, baseline

MDAT FINE MOTOR 36-59 MONTHS: ITEMS 18-41				
23. USE 12 blocks. Make tower of 6 **BLOCKS** and then put 6 blocks in front of child and say "Make a tower just like mine." Allow 2 trials.  YES.....1 NO.....0	22. Child can stack 4 blocks. **SCORE FROM ITEM 23**  YES.....1 NO.....0	20. Child can stack 2 blocks. **SCORE FROM ITEM 23**  YES.....1 NO.....0	30. USE 6 BLOCKS. Child can make bridge of 3 blocks  YES.....1 NO.....0	35. USE 12 BLOCKS. Child can make bridge of 6 blocks  YES.....1 NO.....0
36. USE 12 BLOCKS. Child can make stairs with 6 blocks  YES.....1 NO.....0	24. Show child how fingers go into peg holes. Put in 2 pegs, explaining what you are doing. Remove pegs and ask child to put them all in. Child puts pegs into board in up <= 30 secs.  YES.....1 NO.....0	21. Child puts pegs into board in up to 2 minutes **SCORE FROM ITEM 24**  YES.....1 NO.....0	25. Screws and unscrews lid on jar (after demonstration).  YES.....1 NO.....0	26. Threads 6 beads (demonstrate with 2 beads).  YES.....1 NO.....0

MDAT FINE MOTOR 36-59 MONTHS: ITEMS 18-41

28. Put down 2 sticks of different length on the mat and say to the child, "Do not pick up the sticks. I want you to show me which one is longest." Score YES if child points to longest stick 3 times out of 3 tries  YES……..1 NO……..0	29. Ask child to put hands out. Put the weights in hands simultaneously. Say, "Which is heavier?" Score YES if child picks heaviest container 3 times out of 3 tries  YES……..1 NO……..0	37. USE 8 counters. Make pattern of 4 counters as below and then ask child to copy your pattern (can try 2 times).  YES……..1 NO……..0	38. USE 12 counters. Make pattern of 6 counters as below and ask child to copy your pattern (can try 2 times).  YES……..1 NO……..0	18. Scribbles on paper (straight scribble). Must be more than light marks or stabs to score YES  YES……..1 NO……..0
19. Scribbles on paper (circular scribble)  YES……..1 NO……..0	33. Can copy a cross  YES……..1  NO……..0	27. Copies a vertical line (as drawn by YOU) WITHIN 30 DEGREES:  YES……..1 NO……..0	32. Can copy a circle  YES……..1  NO……..0	34. Can copy a square  YES……..1  NO……..0
41. Can copy 5 letters. Write E C A M J H and ask child to copy each letter. Credit given for letters with all elements, even if upside down or backwards. YES……..1 NO……..0	40. Can copy 3 letters. Write E C A M J H and ask child to copy each letter. Credit given for letters with all elements, even if upside down or backwards. **SCORE FROM ITEM 41.** YES……..1 NO……..0	39. Can copy 1 letters. Write E C A M J H and ask child to copy each letter. Credit given for letters with all elements, even if upside down or backwards. **SCORE FROM ITEM 41.** YES……..1 NO……..0		

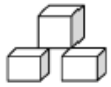
MDAT LANGUAGE AND HEARING 36-59 MONTHS: ITEMS 16-43

20. MARK 'YES' IF CHILD HAS TOLD YOU HIS/HER NAME. If not, ask child to tell you his or her first name.  YES……1 NO……0	18. Ask child to point to **eyes, ears, nose, mouth, hands**. SCORE YES IF CHILD POINTS TO >=2 PARTS.  YES……1 NO……0	23. Remove **12 objects.** Touch or point to each object, saying "Tell me what this is," or "What is this called?" Mark YES if child can **NAME 10** objects  YES……1 NO……0	19. Child can **NAME 5** objects in the basket. **SCORE FROM THE ADMINISTRATION OF 23.** YES……1 NO……0	22. With the **12 objects out,** ask child, "Where's the ____?" DO NOT point, touch or look at the objects. Score YES if child can **IDENTIFY 10** objects.  YES……1 NO……0
16. Child can **IDENTIFY** (point to or give you) **5** or more objects you name. **SCORE FROM THE ADMINISTRATION OF 22.** YES……1 NO……0	21. **USE cup, pencil, matchbox, car. Without gesturing, ask** "Which one is for drinking?" "Which one is for writing?" "Which one is for lighting fire?" "Which one is for driving?" Score YES if child knows **3 or more**. YES……1  NO……0	26. USE **bicycle, spoon, broom, bottle. Without naming, POINT** to each object and say, "What do you do with THIS?" Score YES if child correctly uses action words (verbs) for **3 or more**.  YES……1 NO……0	24. Say, "Tell me 3 foods you like to eat." Be sure to prompt if child says "tea." Ask for 3 animals if child says 2 or fewer foods. Score YES if >=3 foods OR animals.  YES……1 NO……0	28. Say, "What do you do when you are. . .hungry/tired/ cold?" Score YES if child can answer 2 of 3.  YES……1 NO……0
29. Say, "Which goes faster, a car or a bicycle?" Score YES if child says car.  YES……1 NO……0	30. Tell child, "Repeat what I say. Say 'Pa.' Say 'Pa-Chi.' Say 'Pa-Chi-Tu.' Say 'Pa-Chi-Tu-Go.'" Score YES if child repeats all stages correctly.  YES……1 NO……0	25. Say, "Listen carefully to me and do EXACTLY what I say. Ready? OK. I want you to stand up, then clap your hands and then turn around."  YES……1 NO……0	31. Use cup and counter. Ask child to put counter on top of, under, next to, in front of and behind cup. Must do 3.  YES……1 NO……0	32. OPPOSITES. "I want you to help me finish some sentences. If a boy is small, a man is… If the sun comes up in the day, the moon comes up … A baby is young, and a grandma is old. Brothers and fathers are boys, mothers and sisters are girls. You cry when you are sad, but you smile when you are happy. Score YES if **2 or more of these**. YES……1 NO……0

MDAT LANGUAGE AND HEARING 36-59 MONTHS: ITEMS 16-43

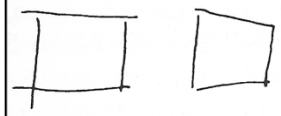
35. Remove 12 counters. Ask child to count them all. Score YES if child correctly counts at least 10 counters.  YES……1 NO……0	34. Child correctly counts at least 5 counters. SCORE FROM ITEM 35 ADMINISTRATION. YES……1 NO……0	33. Child correctly counts at least 3 counters. SCORE FROM ITEM 35 ADMINISTRATION. YES……1 NO……0	36. Child can answer the question "How old are you?" correctly.  YES……1 NO……0	37. USE PAGE WITH 4 COLORED CIRCLES. Point to the RED circle and ask, "What is the name of this color?" Score YES if child says RED.  YES……1 NO……0
38. USE PAGE WITH 4 COLORED CIRCLES. Point to the BLUE circle and ask, "What is the name of this color?" Score YES if child says BLUE.  YES……1 NO……0	39. USE PAGE WITH 4 COLORED CIRCLES. Point to the YELLOW circle and ask, "What is the name of this color?" Score YES if child says YELLOW.  YES……1 NO……0	40. USE PAGE WITH 4 COLORED CIRCLES. Point to the GREEN circle and ask, "What is the name of this color?" Score YES if child says GREEN.  YES……1 NO……0	43. Write child's name in capital letters. Ask child to name EACH letter. Score YES if child names 3 or more letters in first name.  YES……1 NO……0	42. Child names 2 letters in first name. SCORE FROM ITEM 43 ADMINISTRATION. YES……1 NO……0
41. Child names 1 letter in first name. SCORE FROM ITEM 43 ADMINISTRATION YES……1 NO……0	17. MARK 'YES' IF YOU HEARD CHILD speak clearly in sentences (more than just 2 words put together). IF UNCERTAIN, ask Mom if child uses sentences, even if not grammatically correct, e.g., "Mommy give cup," "Give me ball."   YES……1 NO……0			

Appendix A.2: MDAT, endline

ENDLINE MDAT FINE MOTOR: 53-76 MONTHS OF AGE v.5				
AFTER INTRODUCING YOURSELF TO CHILD, ADMINISTER AND SCORE LANGUAGE ITEMS 20, 36 AND 44				
30. USE 6 BLOCKS. Child can make bridge of 3 blocks. **CAN TRY 2 X**  YES.....1 NO.....0	35. USE 12 BLOCKS. Child can make bridge of 6 blocks. **CAN TRY 2 X**  YES.....1 NO.....0	36. USE 12 BLOCKS. Child can make stairs with 6 blocks. **CAN TRY 2 X**  YES.....1 NO.....0	***********OPTIONAL************ **ADMINISTER LANGUAGE ITEMS 33-35 (COUNTING), AND LANGUAGE ITEMS 46-48 NOW (HANDING BLOCKS). BE SURE TO RECORD PERFORMANCE ON LANGUAGE SCORE SHEET.**	28. Put down 2 sticks of different length on the mat and say to the child, "Do not pick up the sticks. I want you to show me which one is longest." **Score YES if child points to longest stick 3 X out of 3 tries**  YES.....1 NO.....0
29. Ask child to put hands out. Put the weights in hands simultaneously. Say, "Which is heavier?" **Score YES if child picks heaviest container 3 X out of 3 tries**  YES.....1 NO.....0	37. USE 8 counters. Make pattern of 4 counters as below (making each row from left to right) and then ask child to copy your pattern (**can try 2 X**). G R R G YES.....1 NO.....0	38. USE 12 counters. Make pattern of 6 counters as below (each row left to right) and ask child to copy your pattern (**can try 2 X**). **G G R** **R G G** YES.....1 NO.....0	33. Can copy a cross. **CAN TRY 2X**  YES.....1  NO.....0	27. Copies a vertical line (as drawn by YOU) WITHIN 30 DEGREES. **CAN TRY 2X**  Less than 30 Degrees YES.....1  More than 30 Degrees NO.....0

1

ENDLINE MDAT FINE MOTOR: 53-76 MONTHS OF AGE v.5

34. Can copy a square. **CAN TRY 2X**  YES……1  NO……0	32. Can copy a circle. **CAN TRY 2X**  YES……1  NO……0	43. Child is able to color within the lines of a square. Child should shade in the majority of the square. **SCORE YES if fewer than 5 marks outside the lines.** YES……1 NO……0	42. Child is able to fold a piece of paper neatly into quarters. Demonstrate, explaining what you are doing. **CAN TRY 2X. Score YES if child folds fairly evenly, with little mismatching (less than 1 inch all around).** YES……1 NO……0	41. Can copy **5** letters. Write **E C A M J H** and ask child to copy each letter. **Score YES for letters with all elements, even if upside down or backwards**. YES……1 NO……0
40. Can copy **3** letters. Write **E C A M J H** and ask child to copy each letter. Credit given for letters with all elements, even if upside down or backwards. **SCORE FROM ITEM 41.** YES……1 NO……0	39. Can copy **1** letter. Write **E C A M J H** and ask child to copy each letter. Credit given for letters with all elements, even if upside down or backwards. **SCORE FROM ITEM 41.** YES……1 NO……0			

ENDLINE MDAT LANGUAGE AND HEARING 53-76 MONTHS OF AGE v.5

43. Write child's name in capital letters. Ask child to name EACH letter, even if child says s/he recognizes name. **Score YES if child names 3 or more letters in first name.** YES……1 NO……0	42. Child names **2** letters in first name. **SCORE FROM ITEM 43 ADMINISTRATION.** YES……1 NO……0	41. Child names **1** letter in first name. **SCORE FROM ITEM 43 ADMINISTRATION** YES……1 NO……0	31. Use cup, counter/block and plate. Say, "Let's play a game. Listen carefully to me and do EXACTLY what I say. Ready?" Ask child to put counter on top of, under, between, in front of and behind cup or plate. **Score YES if child does 3 or more correctly.** YES……1 NO……0	25. **TURN THE CUP UP.** Say, "OK. Now I want you to do just as I say. Put the block in the cup, then put the cup in the basket and then come sit down." **OR "**I want you to stand up, then clap your hands and then turn around." **Score YES if child completes correctly. INSTRUCTIONS CAN BE REPEATED 2X BUT ONLY BEFORE CHILD BEGINS MOVING. Only one try OF EACH command.** YES……1 NO……0
INTRODUCTION TO ITEM 32 OPPOSITES. Say, "I want you to help me finish some sentences. A boy is small, and a man is _____." If the child says "big," "tall," continue to item 32. If child does not understand, say, "A MAN IS BIG. Let's try another one: An ant is small, but a donkey is _____." **If child still does not understand, score as "NO" and continue to item 35.**	32. OPPOSITES. The sun comes up in the day, and the stars come out at **NIGHT**. Tea is hot, but water is **COLD**. A baby is young, and a grandma is **OLD/BIG**. Daytime is bright, and nighttime is **DARK**. Fathers are men, mothers are **WOMEN**. You cry when you are sad, but you smile when you are **HAPPY**. **Score YES if 3 or more of these**. YES……1 NO……0	35. Use 12 blocks. Ask child to count ten blocks. **Score YES if child correctly counts to at least 10 in sequence. They do NOT need to show one to one correspondence for numbers and blocks.** YES……1 NO……0	34. Child correctly counts to at least **5** in sequence. **SCORE FROM ITEM 35 ADMINISTRATION.** YES……1 NO……0	33. Child correctly counts to at least **3** in sequence. **SCORE FROM ITEM 35 ADMINISTRATION.** YES……1 NO……0

ENDLINE MDAT LANGUAGE AND HEARING 53-76 MONTHS OF AGE v.5				
46. Use 12 blocks. Put them in a row. Ask child to hand you <u>one</u> block. YES……1 NO……0	47. Child can hand you <u>three</u> blocks. YES……1 NO……0	48. Child can hand you <u>five</u> blocks. YES……1 NO……0	37. USE PAGE WITH 4 COLORED CIRCLES. Point to the RED circle and ask, "What is the name of this color?" Score YES if child says RED. YES……1 NO……0	38. USE PAGE WITH 4 COLORED CIRCLES. Point to the BLUE circle and ask, "What is the name of this color?" Score YES if child says BLUE. YES……1 NO……0
39. USE PAGE WITH 4 COLORED CIRCLES. Point to the YELLOW circle and ask, "What is the name of this color?" Score YES if child says BLACK/YELLOW. YES……1 NO……0	40. USE PAGE WITH 4 COLORED CIRCLES. Point to the GREEN circle and ask, "What is the name of this color?" Score YES if child says GREEN. YES……1 NO……0	49. Present the book to child (sideways). Say, "Here's a fun book. Let's read the book together. Show me how we read the book (OR SAY "SHOW ME HOW YOU READ A BOOK"). Score YES if Child turns book to right orientation. YES……1 NO……0	50. Continued form item 49. Say, "Let's read this book together. Show me THE PAGE where we can start reading the book." Score YES if Child turns and points to first page where text is expected to begin. YES……1 NO……0	51. Continued from item 50. Say, "Great! Show me where we start reading. Put your finger on the first word where we start reading this story." Score YES if child puts finger on upper left word. YES……1 NO……0
***********CHECK********* HAVE THE FOLLOWING ITEMS ADMINISTERED AT THE START OF TEST (AFTER INTRODUCTIONS) BEEN SCORED?**	20. MARK 'YES' IF CHILD HAS TOLD YOU HIS/HER NAME. If not, ask child to tell you his or her first name. YES……1 NO……0	36. MARK 'YES' IF CHILD HAS TOLD YOU HOW OLD THEY ARE. If not, ask child to tell you his or her age. Verify. YES……1 NO……0	44. Ask the child the name of the village where s/he currently lives. Score YES if child answers correctly. Be sure to verify. YES……1 NO……0	17. MARK 'YES' IF YOU HEARD CHILD speak clearly in sentences (more than just 2 words put together). IF UNCERTAIN, ask Mom if child uses sentences, even if not grammatically correct, e.g., "Mommy give cup," "Give me ball."  YES……1 NO……0

4