



# **ELM Ethiopia Afar & South Omo Endline 2017**

**November 2017**

**Dagnachew Melese and Jonathan Seiden**

Special thanks to our team of data collectors:

Bereket Mayle, Dereje Solomon, Memiru Mathewos, Endashaw Solomon, Anteneh Getachew, Hailu Arballo, Daniel Gezahegne, Alo Gnakay, Baje Shiferaw, Zahara Mohammed, Mussa eleyita, Saulih Mohammed, Eyob Asefa Aminate Mohammed, and Mohammed Hassen

## Table of Contents

Executive summary .....	3
Introduction .....	3
Background .....	3
Context.....	4
Research questions .....	4
Methods.....	5
Sample.....	5
Measurement.....	6
Data collection .....	7
Analysis .....	8
Impact evaluation .....	9
Balance tests between intervention and comparison groups .....	9
Change in IDELA scores over time.....	10
Estimates of impact.....	11
Differences in Afar and South Omo .....	14
Equity and background factors .....	17
Home Learning Environment and Activities .....	17
Background characteristics .....	18
Conclusion.....	20
Limitations .....	21
Appendix A1: Background characteristics & SES in intervention and comparison areas.....	22
Appendix A2: Background characteristics & SES in intervention and comparison areas.....	23
Appendix B1: Simplified multivariate model building process predicting Total IDELA (standard errors clustered at community level) .....	24
Appendix B2: Multivariate models predicting IDELA Domains (complex model, standard errors clustered at the community level).....	25
Appendix B3: Multivariate models predicting IDELA Domains (parsimonious model, standard errors clustered at the community level).....	26
Appendix C1: Estimates of impact on IDELA using multivariate regression at midline.....	27

## **Executive summary**

This report reviews the results from an endline quasi-experimental assessment conducted to estimate the impact of Save the Children’s Early Literacy and Mathematics (ELM) programming and look at trends in the growth and development of children in Afar and South Omo. We compare the developmental status and background factors of children benefiting from Save the Children’s programs (in SC-funded ECD centers and SC-supported government ECD centers) to that of children in unsupported government ECD centers. We assess early learning and development with Save the Children’s International Early Learning and Development Assessment (IDELA).

We attempt to answer two primary research questions: 1) what is the evidence that SC support impacted children’s development, and 2) what are other factors that are related to children’s growth and development in intervention areas.

We find that children in intervention areas had, without fail, better outcomes than those in comparison areas. An overall effect size of over one standard deviation was observed for the overall IDELA score. These findings are robust to a variety of estimation techniques and suggest that that program was highly successful. However, we must caveat these results. The comparison group was only added at the midline data collection point, so we have no baseline to compare developmental status with, and were numerous differences in the background characteristics between comparison and treatment areas. While estimated impact is very large, the overall strength of the evidence is somewhat weak.

Examining the internal relationships within the intervention sample, we present the results of growth over time in South Omo and Afar. We find large differences by the province of children that warrant further investigation into programing. Children in Afar, after starting out below children in South Omo, caught up by midline, and by endline had outpaced children in South Omo. We also find encouraging trends in the literacy environments for children in Afar; parents reported a greater number of home learning activities and reading materials from baseline to midline to endline.

Regarding equity factors, we find that children with more access to reading materials have higher overall development, and that this is concentrated in early literacy skills. Amharic speaking children at baseline had better developmental status than non-Amharic speaking children. This gap began to close at midline, but the trend was reversed at endline with the gap again widening. Other factors, such as SES, gender, and parental education level, did not appear to be significant predictors of children’s developmental status.

Overall, the evaluation presents a cautiously positive tale about implementation in Afar and South Omo.

## **Introduction**

### **Background**

One of Save the Children’s programmatic priorities is supporting 4-6 year-old children around the world with quality early childhood care and development (ECCD) programs. Our focus is on strengthening school readiness skills so that children are ready to enter Grade 1 and succeed in school. Emergent literacy and math (ELM) skills developed in these preschool years are crucial for later reading and math

outcomes. Since 2013, Save the Children has implemented an innovative ELM toolkit in Ethiopia with an aim to provide targeted training to ECCD and government “O” class<sup>1</sup> facilitators on how best to support these skills through play and joyful learning, in both pastoralist communities and school-based early childhood centers.

Save the Children supports the Government of Ethiopia to strengthen Early Childhood Care and Development in Tigray, Oromia, Amhara, the Southern Nations Nationalities and Peoples Region (South Omo), Afar, Benishangul Gumuz, Somali, and Gambella. The longitudinal data described in this study are from three yearly data collections from 2015-2017 in South Omo and Afar pastoralist contexts. The baseline assessment was conducted only in the intervention area before the introduction of an ELM toolkit. The midline data collection added a comparison group which was followed along with intervention children to the endline data collection at the conclusion of the project in 2017.

## **Context**

The intervention was implemented in Afar region and South Omo Zone of SNNPR, where SC supports 16 ECCD centers and 6 government “O” classes in South Omo Zone and 6 ECCD centers at Assayita District of Afar region. All of the intervention ECCD centers and government “O” classes are located on the grounds of primary schools. At least one trained facilitator at each of these centers engages children in structured and unstructured play, using learning and teaching materials which are provided by the program. The children attend the centers five days a week, for 3 hours per day on average. All facilitators receive a foundational skills training on ELM at home and center components at the beginning of the program. These ELM interventions began in 2015 with an aim to improve school readiness by foundational emergent literacy and math skills development. The facilitators of the ECCD centers were provided with a five day interactive training on ELM, focused on simple ways to incorporate ELM in their daily classroom routine through playful games and materials. Centers were also provided with a resource bank of 50 early literacy games and 50 math games whose use was the focus on the training. Parents were also engaged in the intervention, with parent facilitators trained on ELM at Home components and then leading sessions in intervention areas.

Data for this study was collected using the International Developmental and Early Learning Assessment (IDELA) tool, a play-based assessment tool designed for children in the 4-6 age group. The tool has domains that assess emergent literacy and language, emergent numeracy, motor development and social-emotional development of the child. The baseline assessment was conducted in October 2015 in both the Afar and South Omo intervention areas. A total of 120 children with their caregivers in Afar and 380 children with their caregivers in South Omo were assessed.

## **Research questions**

The key research questions explored in this report include:

- 1) How did children’s learning and development differ at endline for children who did and did not receive Save the Children ELM programming in Afar and South Omo?

---

<sup>1</sup> “O” class is one year of government-provided pre-primary education

- 2) How did learning and development change over time for children who received ELM programming?
- 3) What are the gaps in learning and development due to gender and ethnicity and how did they change over time?
- 4) How did caregiver’s attitudes and practices relating to childhood learning and development change over time?
- 5) What are the relationships between learning and development and parental background factors, interactions with children, etc.?

## Methods

### Sample

A total of 780 children were involved in the data collection for this project. At baseline, data was only collected from 19 sites in “intervention” areas, meaning that they would receive Save the Children programming (Early Literacy and Math) at ECCD sites constructed by Save the Children Norway. There were two types of sites: Save the Children constructed ECCD centers (referred to as “intervention ECCD Centers” hereafter), and government-run “O”-classes (referred to as “intervention ‘O’-classes” hereafter). Both types of sites received similar programming.

At midline, we expanded the sample to include a comparison group of 12 sites not receiving Save the Children programming. In Afar, four comparison “O” classes were added. In South Omo, eight comparison “O” classes were added to the sample by randomly selecting communities adjacent to the intervention communities. **Error! Reference source not found.** summarizes the number of sites in each area at baseline and midline/endline.

**Table 1. Intervention and comparison sites**

	Baseline	Midline/Endline
<b>Afar</b>		
SC ECCD Centers	6	6
Intervention Government “O” classes	0	0
Comparison Government “O” classes	0	4
<b>South Omo</b>		
SC ECCD Centers	6	6
Intervention Government “O” classes	7	7
Comparison Government “O” classes	0	8

Sites were selected for assessment by the district education office or the local government authorities. At midline and endline, children were interviewed at 10 sites in Afar and 21 sites in South Omo. By using attendance lists, data enumerators randomly sampled children at each site and attempted to interview 20 children.

At endline, we restricted the data collection to two groups of children: in intervention areas, we only interviewed children for whom we had baseline and midline data. In comparison areas, we attempted to re-interview children whom were interviewed at midline. While attrition rates were high between baseline and midline in the intervention area (as discussed in the midline report), attrition rates from midline to endline were encouragingly low. Just 5 (2.3%) children from the comparison area and 7 (2.1%) children from the intervention area attrited between midline and endline, and this attrition did not differ between the intervention and comparison groups.

We offer two types of analysis in this report. We first estimate the midline and endline impact of the program by examining IDELA scores in intervention and comparison areas with complete midline and endline data as expressed in **Table 2**. For our examination of equity issues, we include only children from intervention areas for whom we have baseline, midline, and endline data as shown in **Table 3**.

**Table 2 Impact evaluation sample**

	Comparison	Intervention	Total
Afar	55	85	<b>140</b>
South Omo	153	242	<b>395</b>
<b>Total</b>	<b>208</b>	<b>327</b>	<b>535</b>

**Table 3 Equity analysis sample**

	Baseline	Midline	Endline	Total
Afar	83	83	83	249
South Omo	169	169	169	507
<b>Total</b>	<b>252</b>	<b>252</b>	<b>252</b>	<b>756</b>

## Measurement

In this study two main tools were used: the IDELA Child Assessment with children and the IDELA Caregiver Survey. The IDELA Child Assessment was used to measure children’s early learning and development and was administered by data collectors with direct child observation. Items included in IDELA are listed in **Table 4**. The same versions of both tools were used at baseline and midline. The IDELA Caregiver Survey is detailed in **Table 5**.

**Table 4. IDELA Child Assessment**

Gross and Fine Motor Skills	Emergent Literacy	Emergent Math	Social-emotional Development	Other items
Hopping	Print awareness	Size/length identification	Friends	Approaches to learning
Copying a shape	Expressive vocabulary	Sorting	Recognizing emotions in self	Inhibitory control

Drawing a human figure	Letter identification	Number identification	Recognizing emotions in others	Short term memory
Folding paper	Emergent writing	Shape identification	Conflict resolution	
	Phonemic awareness	One-to-one correspondence	Personal information	
	Oral comprehension	Simple operations		
		Puzzle completion		

**Table 5. IDELA Caregiver Questionnaire**

Section	Description
<b>1. General family information</b>	Sex of child, child age, number of children at home, parental literacy, parental education, languages spoken at home
<b>2. ECCD experience and educational expectations</b>	Child participation in ECCD programs, details of participation, parental expectation and aspirations of child's educational attainment.
<b>3. Access to early learning materials and resources at home</b>	Types of reading materials at home, types of toys at home
<b>4. Parenting practices and support for learning and development</b>	Adults in the home engaging with children to promote learning and development.
<b>5. Inadequate care</b>	Children left alone or in the care of another young child
<b>6. Caregiver self-efficacy</b>	Attitudes about parent's role in child's development.
<b>7. Socioeconomic status</b>	Housing materials, objects/appliances owned, land/animals owned.

## Data collection

Cluster supervisors for collecting IDELA data were identified by district education offices, based on criteria set by Save the Children (local language speaking ability, previous experiences in data collection from children and parents and with BA Degree holder from that specific area).

Supervisors for data collection were trained on the IDELA data collection tool including the questionnaires prepared for caregivers. And the data collectors also trained in such a way that collecting

data should be based on local pastoralist contexts. A total of 16 data collectors (10 from South Omo and 6 from Afar) were selected and trained.

From each ECCD center and government “O” class, children were randomly selected for data collection using the attendance list as the sample frame. After completing data collection with children, data collectors reached caregivers in one place around the school compound in some of the sites and for other site, data collectors by themselves went to the caregiver’s home and collected the data. The endline data collection was conducted from 15 May to 30 June 2017 and was encoded from 15 July to August 5 2017.

## Analysis

This report covers two main research questions that correspond to the two samples for analysis:

- 1) What evidence, if any, do we have that the ELM program impacted children’s early learning and development scores during the period from midline to endline?
- 2) Within the intervention group, what factors were related both to children’s baseline early learning and development, and also changes over time?

We attempt to answer our first research question using the midline-endline sample of children in ECCD Centers and “O” Classes who benefited from the ELM intervention and comparing outcomes to the comparison sample that did not benefit from the intervention. For the second research question, we restrict our analyses to only the children in the intervention group who were interviewed at baseline, midline, and endline.

For the first research question, we review the midline and endline scores and background characteristics of children in the intervention and comparison groups and highlight significant differences at midline. After noting differences between them, we attempt to control for these differences and estimate the impact of the program. As discussed here and in the midline report, there were large and significant differences in background characteristics that make it difficult to isolate the impact of the program by examining IDELA scores. To account for this fact, we attempt a variety of analytical methods, including multivariate regression analysis and Propensity Score Matching. **While we do provide an estimate of impact in this report, we include a long list of limitations, and caution against making strong conclusions about the success of the program according to this analysis.**

With the second research question, we take an extended look at the background and demographic factors associated with both different development status at the beginning of the study and the relationships that these factors had with growth over time. We primarily rely on multivariate regressions including interaction terms to answer this research question. Through this process, we attempt to build a model that best explains children’s growth and development over time and then conduct an equity analysis to examine children in our intervention groups benefited from our programs equitably.



## Impact evaluation

At the recommendation of the SCUS DECP Research Team, a comparison group was added at midline data collection. This was in hope of comparing the results in the intervention area to those in the comparison area to estimate the impact of the program. However, comparison groups created after the decision to implement a project are subject to substantial bias and can be a poor estimate of the counterfactual for the intervention group. Without a baseline measure of our outcome (in this case early learning and development), we cannot know if children all started at the same level in both intervention and comparison communities. **This severely limits the strength of our estimates of impact and estimates of impact should be considered as “suggestive” rather than strongly causal conclusions.**

### Balance tests between intervention and comparison groups

To adequately frame our impact estimates and demonstrate their limitations, we first examine differences in background and demographic characteristics. A full breakdown of average scores in intervention and comparison communities is listed in Appendices A1 and A2. **Table 6** presents those characteristics which demonstrated significant differences at midline or endline. As the table shows, we find large and significant differences at midline and endline in both demographic and situational characteristics. Some of these differences may be related to the program—for example, caregivers in intervention areas may have improved their home learning environments or engaged in more learning activities with their children. However other characteristics, such as the education level, size of family, or socioeconomic status of families, is unlikely to be affected by the intervention. Overall, we find that children in the intervention group have more well-educated parents, come from smaller families, spend more time with their parents, and have families with more resources and learning interactions.

**Table 6. Background characteristics with significant differences between ELM and Comparison areas at Midline and Endline (significant differences only)**

	Midline			Endline		
	ELM	Comparison	p-value	ELM	Comparison	p-value
<b>Father is literate</b>	59%	27%	0.00	60%	27%	0.00
<b>Mother is literate</b>	46%	12%	0.00	30%	12%	0.04
<b>Average number of children in household</b>	3.2	3.4	0.32	3.4	3.7	0.03
<b>Average hours child is with mother per day</b>	2.9	2.6	0.01	3.1	2.6	0.00
<b>Average hours child is with father per day</b>	2.2	1.9	0.02	2.3	1.9	0.00
<b>Average number of hours child spends alone per day</b>	1.4	0.9	0.00	1.3	1.0	0.01
<b>Home has a bedroom</b>	39%	22%	0.06	40%	21%	0.04
<b>Home has a living room</b>	39%	23%	0.06	63%	35%	0.00
<b>Family has radio</b>	41%	19%	0.01	35%	20%	0.09
<b>Family has electricity</b>	43%	16%	0.00	41%	15%	0.00
<b>Family owns land</b>	62%	87%	0.00	66%	89%	0.01
<b>Total number of types of reading materials</b>	1.47	0.91	0.00	1.46	1.03	0.00

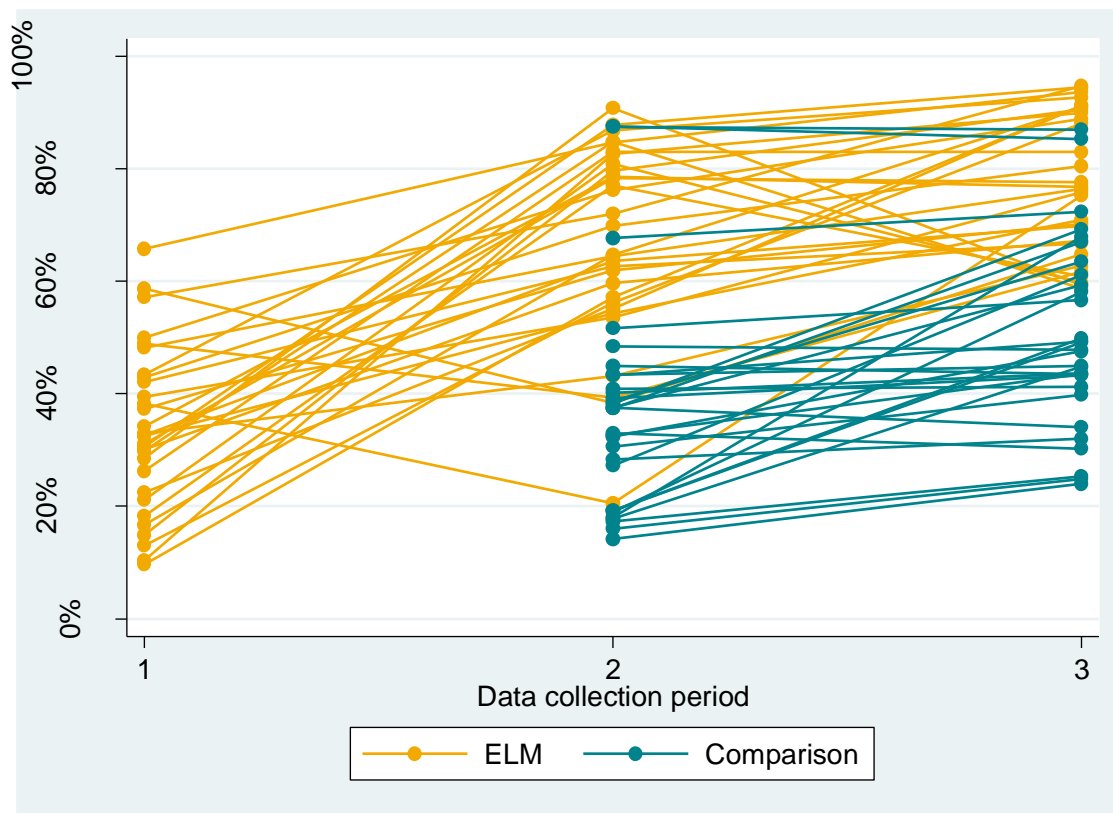
<b>Child has drawing toys</b>	39%	22%	0.05	50%	29%	0.02
<b>Total number of types of toys</b>	3.40	2.53	0.00	3.23	2.42	0.00
<b>Family reads to child</b>	50%	24%	0.01	57%	28%	0.00
<b>Family sing to child</b>	77%	59%	0.08	86%	63%	0.02
<b>Family goes out with child</b>	76%	56%	0.03	79%	59%	0.04
<b>Family draws with child</b>	60%	40%	0.04	68%	44%	0.01
<b>Family teaches child something new</b>	51%	27%	0.01	59%	36%	0.01
<b>Family teaches letters</b>	53%	24%	0.00	61%	30%	0.00
<b>Family teaches numbers</b>	52%	27%	0.01	60%	34%	0.01
<b>Average number of types of home learning activities</b>	4.62	2.85	0.00	5.05	3.41	0.00
<b>Average number of types of negative discipline</b>	1.54	1.23	0.01	1.32	1.11	0.06

While we will use analytical methods to control for many of these factors, there may be omitted variables which we cannot control for. Overall, the strength of our results is low and findings should not be overstated.

### Change in IDELA scores over time

**Figure 1** presents a random subset of 70 children’s Total IDELA scores at baseline, midline, and endline (for children from intervention areas) and at midline and endline for children in comparison areas. It’s clear from the graph that the children in intervention areas, on average, made large gains between baseline and midline (nearly 28 percentage points on average). These children then continued to gain, on average 11 percentage points, between midline and endline. This was a slower rate, but there was comparatively little room to grow as IDELA scores cannot exceed 100%. In comparison areas, children had much lower scores at midline than children in the intervention group, but gained, on average, 15 percentage points between midline and endline.

**Figure 1. Early learning and development trajectories in intervention and comparison areas (n=70)**



It's clear from **Figure 1** that, at both midline and endline, children benefiting from ELM were scoring far higher on average than children in comparison areas. As mentioned previously, we know that there are many significant differences between children in intervention and comparison areas.

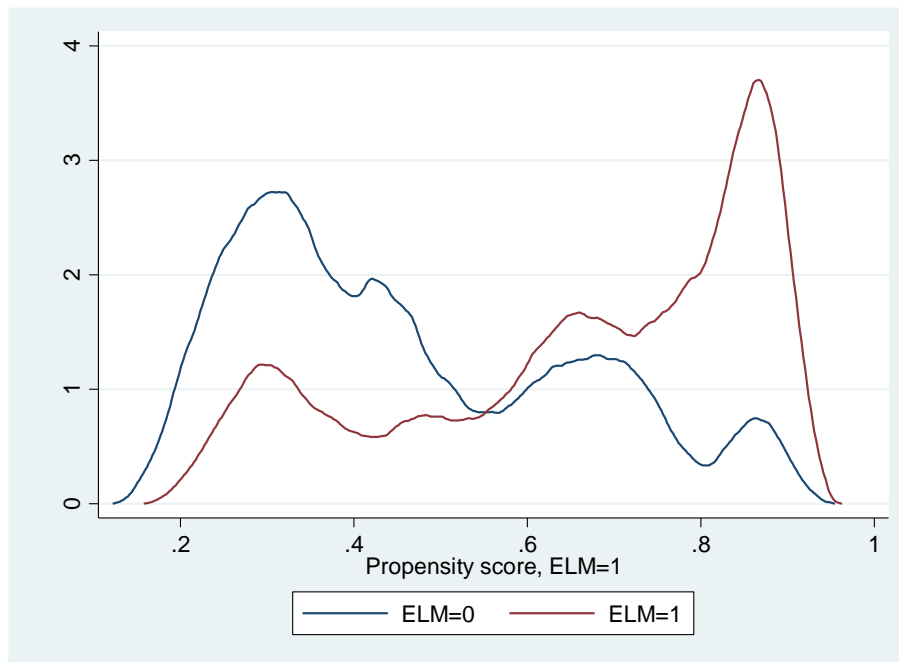
### Estimates of impact

We now attempt to control for those factors and isolate an estimate of the impact of the program. To estimate the impact of the program, we attempt two methods: Propensity Score Matching (PSM) and multivariate regression. Propensity Score Matching attempts to account for imbalance in the intervention and comparison groups by only comparing observations that have a similar “likelihood of receiving the intervention”.

We first create a vector predicting the likelihood of receiving the intervention by running a logistic regression on background characteristics and predicting for each observation. In our case, we include area, parental literacy and educational status, parental age, number of children in the family, socio-economic status, midline home learning environment (number of types of toys and books) and midline home learning activities. We then discard from our sample children who have no good match and only compare outcomes between children who we believe are similarly likely to receive the intervention. An overlap plot of predicted likelihood of receiving ELM programming is presented in **Figure 2**. While, as we would expect, most children who received ELM had a high propensity score, there is substantial overlap between the comparison and intervention groups that allows us to compare outcomes between

the two. Children on the far left-hand of the distribution (where there is no overlap) are excluded from this analysis.

**Figure 2. Overlap plot of likelihood of receiving ELM program based on background characteristics**



After checking our propensity score, we then compare outcomes for children who have similar scores. **Table 7** presents the PSM estimates of programmatic impact (in units of effect-size). The table presents the estimated impact in units of effect size (Cohen’s  $d$ )<sup>2</sup> and percentage points. For example, we estimate that the program led to, on average, a 29 percentage point increase in IDELA scores at midline, which had faded slightly to 22 percentage points at endline. These estimates of impact, if accurate, are very large and would indicate a highly successful program. In order to be accurate, children with similar Propensity Scores should be similar with the exception of the program. Similar to multivariate regression, PSM is vulnerable to “omitted variable bias,” meaning that there may be important but unobserved differences between children in ELM areas and those in the comparison group.

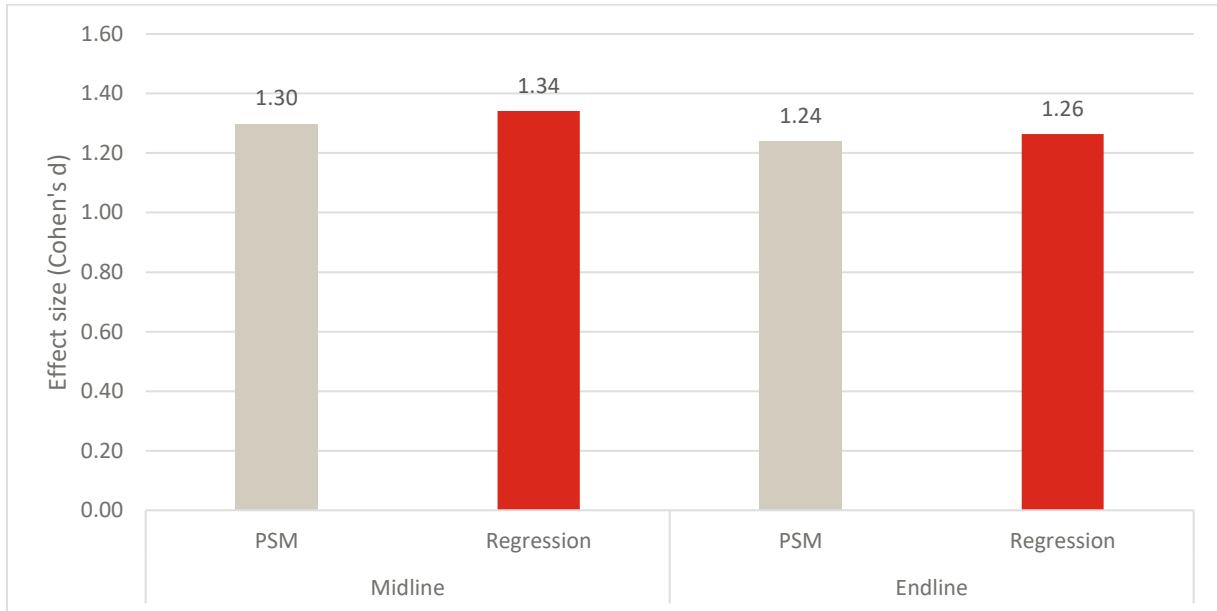
**Table 7. Estimates of impact on IDELA domains using Propensity Score Matching (PSM)**

Variable	Effect at Midline	Effect at Midline (p.p.)	P-value at Midline	Effect at Endline	Effect at Endline (p.p.)	P-value at Endline
Motor	1.18	27	0.000	0.82	20	0.000
Early Literacy	1.23	28	0.000	1.10	22	0.000
Early Numeracy	1.33	29	0.000	1.35	26	0.000
Social-Emotional	1.42	31	0.000	1.01	22	0.000
Total IDELA	1.43	29	0.000	1.19	23	0.000

<sup>2</sup> Cohen’s  $D$  is calculated by the following equation:  $d = \frac{(\text{intervention mean}) - (\text{comparison mean})}{\text{pooled standard deviation}}$

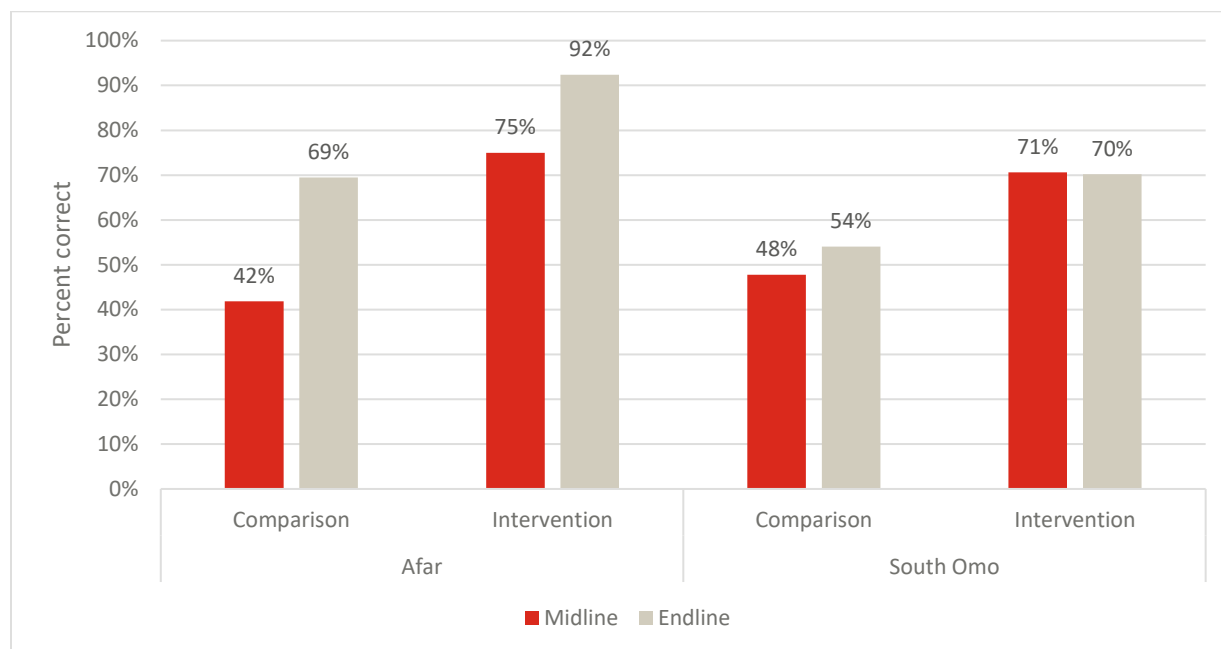
When comparing these results to a different methodology—namely using multivariate regression to control for the background characteristics—we find similarly large and striking results. As Figure 3 shows, using multivariate regression results in slightly larger estimates of impact at both midline and endline

**Figure 3. Estimates of impact on IDELA total at midline and endline using PSM and multivariate regression**



The multivariate regression also allows us to (begin to) examine differences in impact between areas. As **Figure 4** shows, children in intervention areas of Afar and South Omo had higher average IDELA scores at both midline and endline. However, children in Afar gained much more between midline and endline, regardless of intervention status. In South Omo, while the intervention group was higher than the comparison group at midline and endline, scores in the intervention group did not change significantly between midline and endline compared to the comparison group, after controlling for age and background factors.

**Figure 4. Predicted midline and endline scores in Afar and South Omo by intervention status, controlling for age and background characteristics**



By any measure, children in intervention areas far outperformed children in comparison areas. The effect sizes (ranging between 1.18-1.43 at midline and 0.82-1.35 at endline) are very, very large. If these estimates are accurate, they are very impressive and would indicate a highly successful program. However, these estimates of impact rest on assumptions that children in the intervention and comparison group are from comparable sub-populations, other than the factors controlled for in the multivariate regression and PSM models. While these are encouraging results, without baseline scores in both areas to establish the comparability of the treatment groups, we may be greatly overestimating the impact of the program. **As such, we can conclude that while we have evidence of a large and significant impact of the program, the strength of this evidence is fairly weak.**

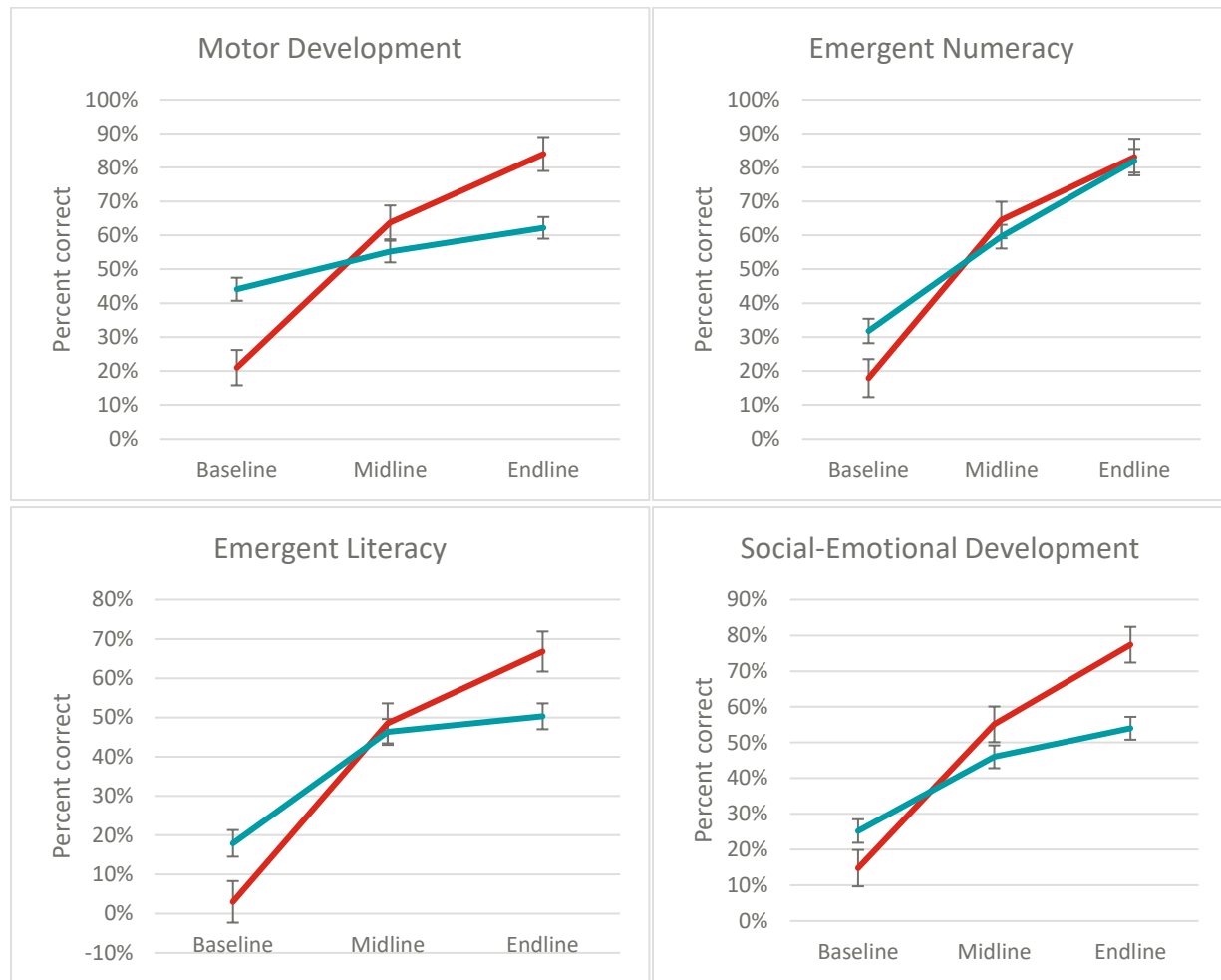
### **Differences in Afar and South Omo**

We now focus exclusively on children in the intervention area that were interviewed at baseline, midline, and endline (n=252) and examine differences in growth and development and background characteristics. When looking at intervention status, we control for intervention area due to the large differences noted between Afar and South Omo. We compare the growth of these children over the past two years across the IDELA core domains, and consider differences in their growth patterns by background characteristics as measured on the caregiver survey.

As discussed in the midline report, at baseline children in South Omo scored higher than children in Afar on IDELA in every domain. However, children in Afar gained more between baseline and midline, leaving children in Afar with similar or higher scores than children in South Omo at midline. This trend continued from midline to endline. As **Figure 5** shows, while children in Afar and South Omo scored similarly at midline, at endline children in Afar scored higher in every domain except for Emergent Numeracy. This

size of the gap between South Omo and Afar at endline was somewhat surprising, and we should examine program implementation to see if there were differences that may explain these surprising results.

**Figure 5. Growth in IDELA core domains in Afar (red) and South Omo (blue)**<sup>3</sup> — Afar — South Omo



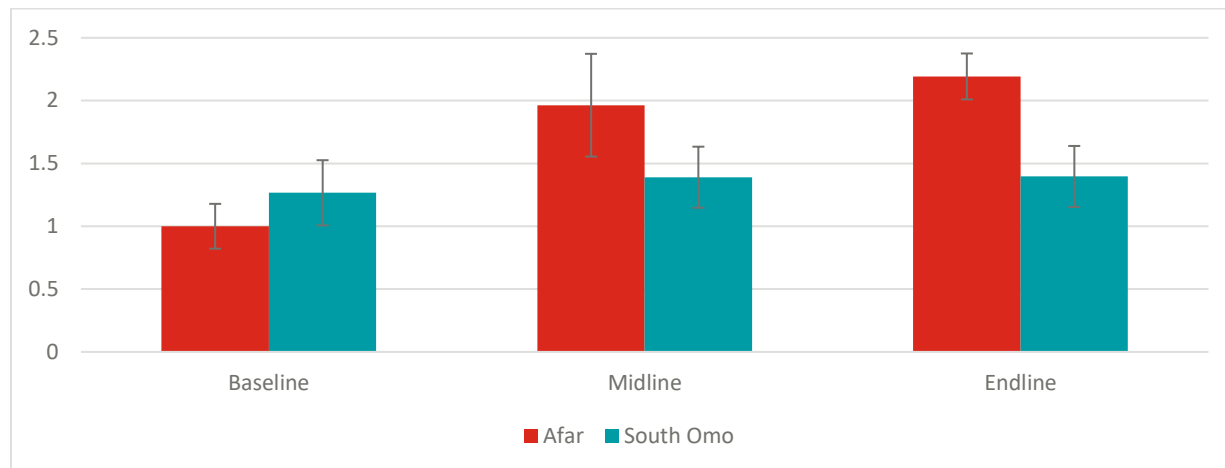
As noted above, the only domain in which we do not observe a significant gap between Afar and South Omo is in the Emergent Numeracy domain, where children in both areas demonstrate strong growth. Notably, the trajectory for students in South Omo fell off sharply in the Emergent Literacy and Social-Emotional Domains. Given that these domains rely strongly on language use, student may be struggling with language challenges in the classroom in South Omo.

When trying to explain differences between these two areas, we can examine differences in background characteristics and how they changed over time. Perhaps most notable was the change in the number of types of reading materials between Afar and South Omo. Caregivers were asked if they had nine different types of reading materials in their home. As **Figure 6** shows, at baseline, caregivers reported owning about one type of reading material in both Afar and South Omo. However, at endline, while this

<sup>3</sup> Error bars around each point estimate represent the 95% confidence interval of the estimate

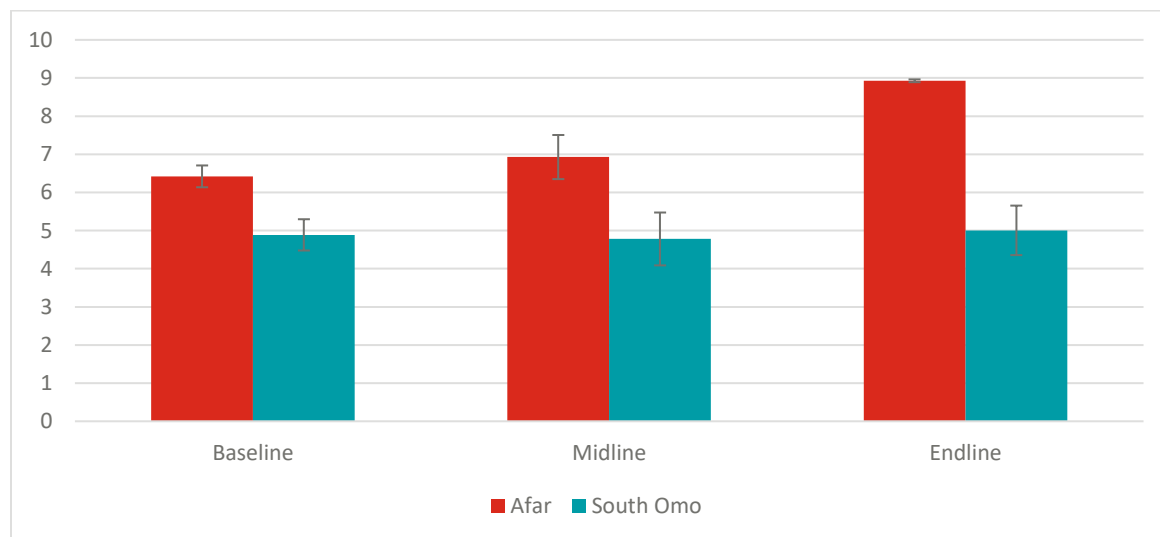
figure had not changed in South Omo, the number of types of reading materials in Afar had more than doubled.

**Figure 6. Number of types of reading materials in Afar and South Omo**



We find similar results when examining home learning activities. Caregivers were asked if they, or any other adult in their household engaged in learning activities with their child such as teaching them letters, playing games, or taking them out of the house. As **Figure 7** shows, children in Afar were already being exposed to more learning activities at baseline than children in South Omo. This gap only widened as nearly all caregivers in Afar reported engaging in all activities with their children by endline, while caregivers in South Omo had no change from baseline to endline.

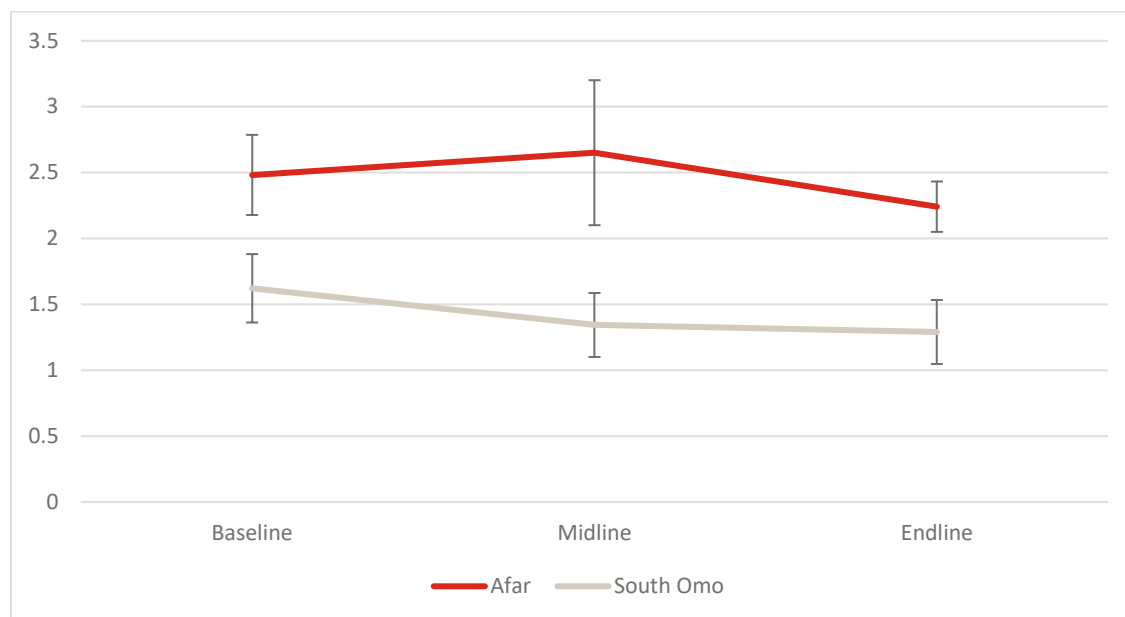
**Figure 7. Number of types of learning activities caregivers engage in with children in Afar and South Omo**



Trends in negative discipline reveal a persistent difference between Afar and South Omo. As **Figure 8** demonstrates, caregivers in Afar consistently reported a higher level of negative discipline than in South Omo (hitting, spanking, or yelling at a child). While we don't observe a significant difference between baseline and endline, the general trend in both areas was a slight decline.



**Figure 8. Number of types of Negative Discipline reported by caregivers**



### Equity and background factors

For the final portion of this analysis, we again focus on children in the intervention area that were interviewed at baseline, midline, and endline (n=252). We consider the relationships between the different domains of IDELA and how growth over the past two years varied by background characteristics as measured on the caregiver survey.

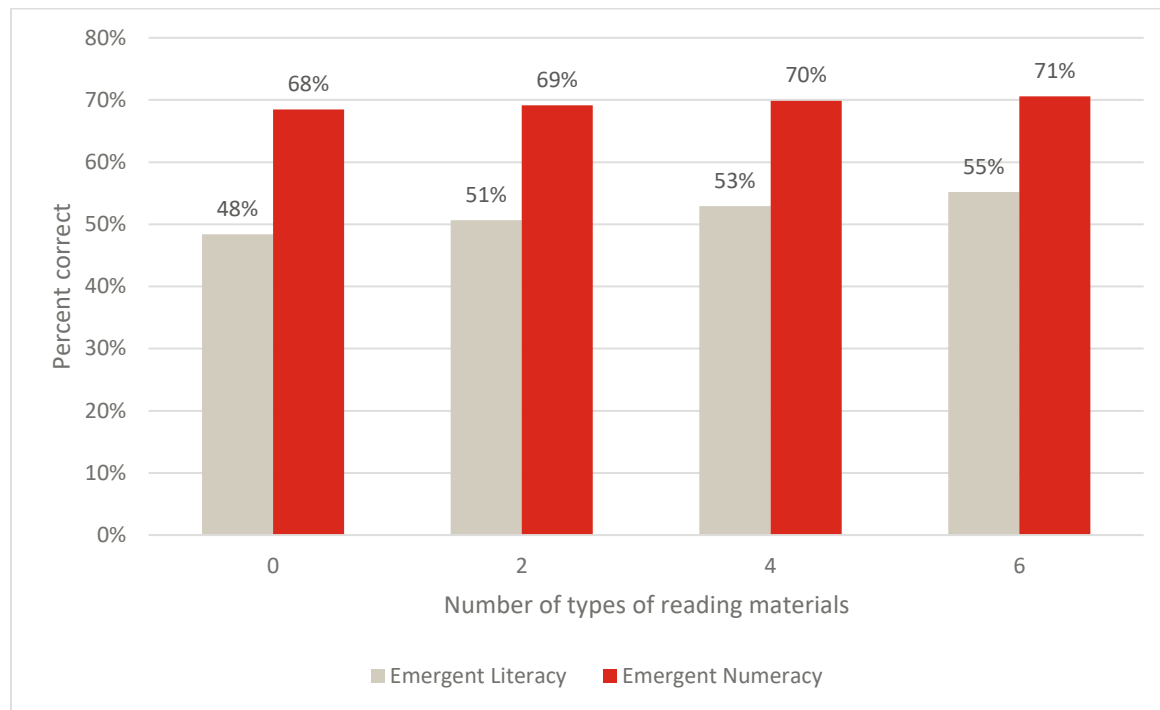
Estimates of the effects of different factors on IDELA scores comes as a result of a model building process described in appendices B1, B2, and B3. We first tested individual factors' relationships with children's total IDELA scores. If we found a significant relationship, we then tested to see if this relationship remained significant after controlling for the data collection period (baseline, midline, and endline) and area (South Omo and Afar). If the variable was significant in this regression, we report on it below. Variables that were found insignificant are not included in this report. We then included all significant variables in our "complex" model as shown in B2. Finally, B3 demonstrates the most "parsimonious" model that retains only those variables that remain significant when considered jointly.

### Home Learning Environment and Activities

Engaging in learning activities with children and fostering a learning environment have been shown to be strongly related to children's growth and development over time. Our findings reinforce these notions as we find significant positive relationships between children's learning environment and their IDELA scores. While we do not find any relationship between the number of types of toys a child has and their IDELA scores, we find significant relationships between the number of types of reading materials and children's Early Numeracy, Early Literacy, and Total IDELA scores. While the absolute magnitude of the relationship is not huge, it is statistically significant and consistent across time periods. It is notable, but not surprising, that the strongest relationship comes from the Early Literacy Domain. As **Figure 9** demonstrates, for one additional type of reading material, we predict a 1.3 percentage point increase in

the child's Emergent Literacy score. This is nearly double the 0.7 percentage point increase in their Emergent Numeracy score that we find.

**Figure 9. Predicted Emergent Literacy and Emergent Numeracy Scores by types of reading materials**



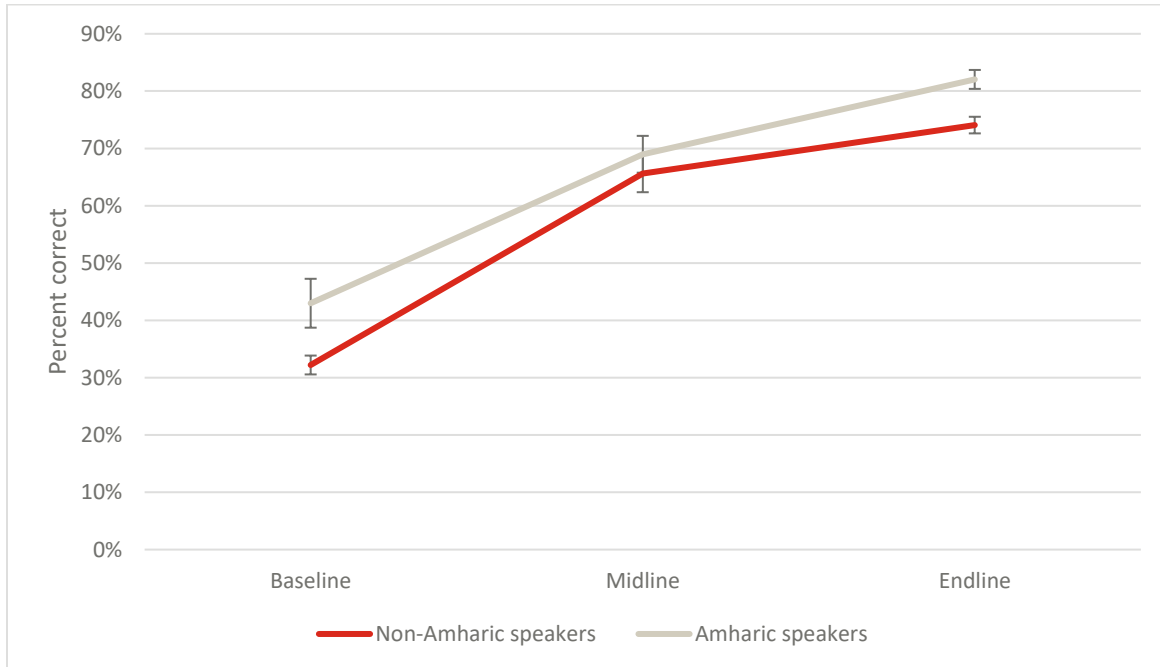
While we cannot assert a causal statement, that more reading materials cause children to learn more, the fact that relationship is observed to be strongest in the Emergent Literacy domain points to the importance of reading materials and the effect access has on children's early learning.

Interestingly, while we find a relationship in the learning environment and early development across time periods, we do not observe the relationship between the number of learning activities that a child engages in with their caregivers that we found at midline in Afar. This result may be slightly obscured by the fact that we are controlling for the time period and area in this analysis. As discussed above, HLA nearly doubled in Afar, while remaining flat in South Omo. It's possible that differences according to HLA are being masked by this control.

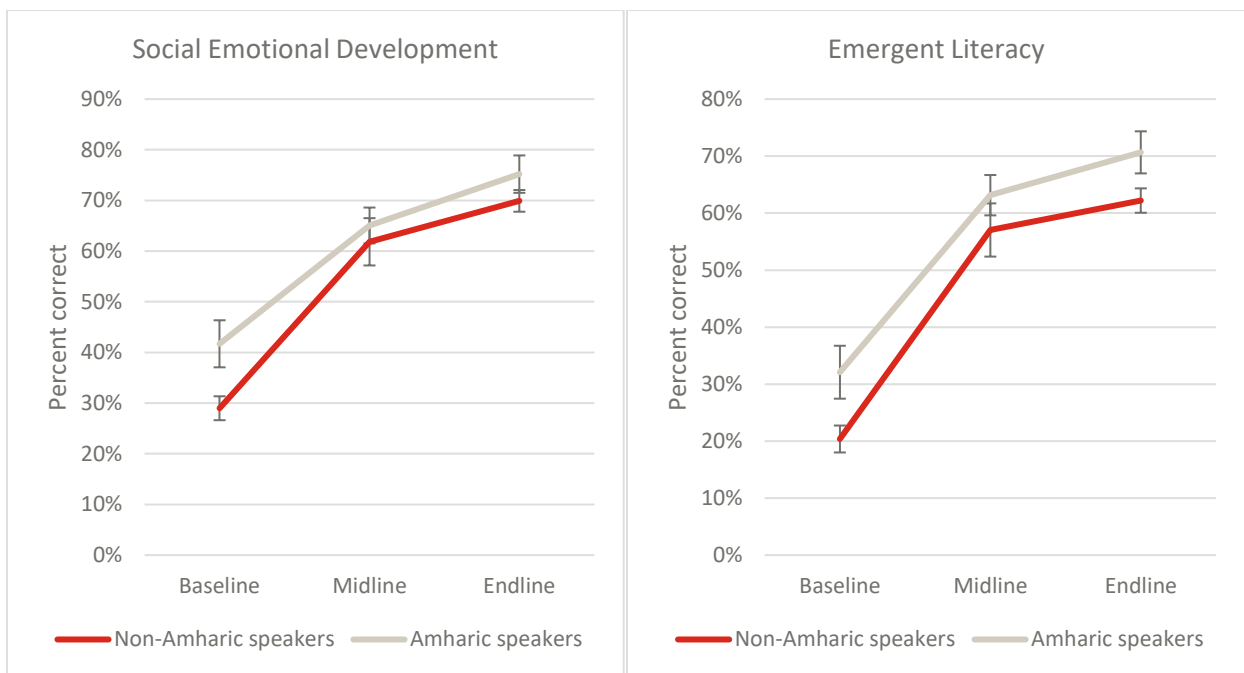
### Background characteristics

Children who speak primarily Amharic at home started off with significantly higher baseline scores than children whose home language was different than Amharic. As the midline report noted, this gap narrowed in Afar at midline. Unfortunately, this gap increased at endline again. As **Figure 10** shows, at midline, Total IDELA scores were statistically indistinguishable between the two regions, but then the gap again widened at endline. In the Social-Emotional domain, however, the result was different, and **Figure 11** presents a more positive story. Non-Amharic children partially closed the gap in Social-Emotional development at midline, a trend that continued at endline.

**Figure 10. Total IDELA for Amharic and non-Amharic children**



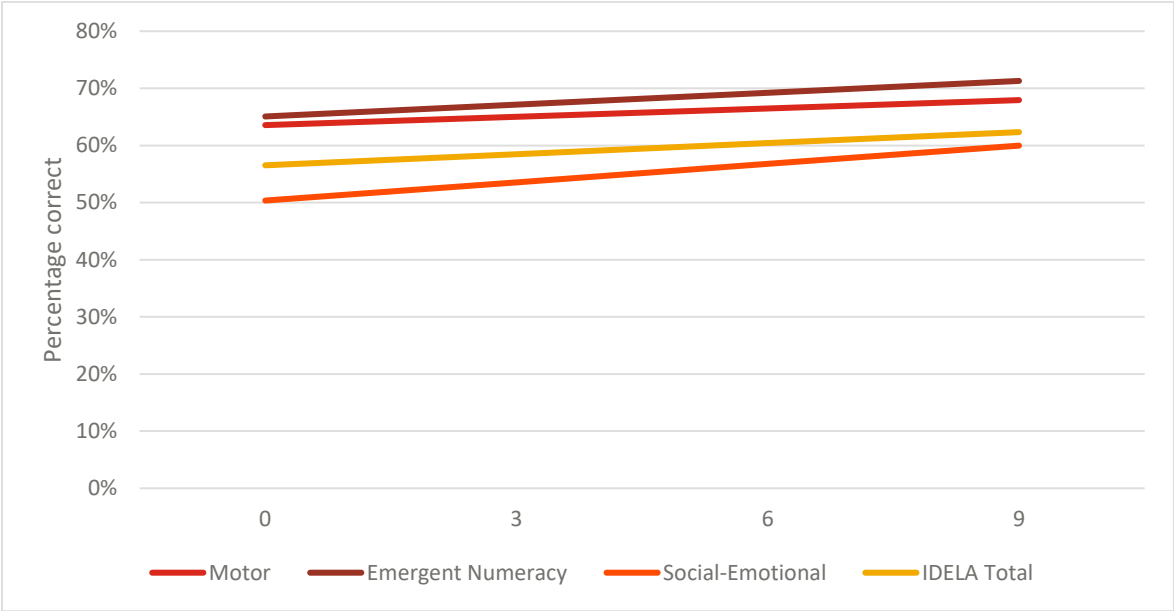
**Figure 11. Social Emotional and Emergent Literacy domains for Amharic and non-Amharic children**



It is also unsurprising to learn that, as **Figure 11** shows, non-Amharic speaking children scored consistently lower in the Emergent Literacy domain—a domain that relies highly on language skills. These results emphasize the importance of ensuring that learning, and assessment, happens in a language that children understand.

The final area where we find significant differences is according to children’s socio-economic status. As **Figure 12** shows, an children’s socio-economic status (as measured by the number of types of common household items the family possesses) has a significant relationship with overall IDELA, and the Motor, Emergent Numeracy, and Social-Emotional Domains.

**Figure 12. IDELA domains by number of types of home possessions**



### Conclusion

Our research design prohibits us from making strong causal conclusions about the impact of the program. Nevertheless, the results are good and suggest a large and positive effect of the program. Though still very, very large, our estimates of impact at endline are slightly smaller than at midline, suggesting a slight fade out of impact. However, we cannot stress enough the caution with which these results should be interpreted. The fact that the comparison group was only added at midline means that we have no baseline scores to analyze differential gains that students may be making.

The caregiver survey revealed mixed signs regarding changes in caregiver practices. We saw positive improvement within the intervention area in Afar on both the home learning environment and home learning activities: caregivers reported significantly richer learning environments at midline and endline than at baseline. In South Omo, there was no difference. Finally, we found persistently higher negative discipline in Afar than South Omo, with no significant change from baseline to endline.

Regarding equity, the results were mixed. Gender appeared not to be a significant factor for IDELA outcomes, meaning that boys and girls were doing equally well at baseline, midline, and endline. Ethnicity and language proved a more interesting dimension. Overall, we found that after the gap between Amharic and non-Amharic speaking children shrunk between baseline and midline, it re-emerged at endline.

The most interesting question as a result of this analysis is the substantial deviation in results between Afar and South Omo. Children in Afar experienced consistently larger positive gains both between baseline and midline and midline and endline. Investigating the differences in programming between these regions may help identify key factors that made the program successful.

Overall, the analysis paints a positive picture of the ELM programming in Afar and South Omo. Children undeniably demonstrated improved learning and development at both midline and endline than and reported positive growth in their environment, attitudes, and actions.

### **Limitations**

As discussed, the research design of this report prohibits any strong causal conclusions or estimate of the impact of ELM. Attrition did not appear to be differential on measured characteristics. However, there may be other unobserved differences between children who attrited and those who didn't that bias the results. More importantly, there appears to be substantial selection bias. The large and significant differences in background characteristics between children receiving ELM and those who didn't was indicative that the differences in learning and development between them cannot be fully attributed to ELM.

## Appendix A1: Background characteristics & SES in intervention and comparison areas

		Midline			Endline		
		ELM	Comparison	p-value	ELM	Comparison	p-value
Sample background characteristics	Father is literate	59%	27%	0.00	60%	27%	0.00
	Mother is literate	46%	12%	0.00	30%	12%	0.04
	Percent female	44%	45%	0.91	44%	45%	0.97
	Child's age	5.3	5.0	0.00	6.0	6.5	0.00
	Mother's age	29.5	29.9	0.50	31.0	30.8	0.76
	Father's age	35.8	36.4	0.37	37.1	37.5	0.52
	Number of children in household	3.2	3.4	0.32	3.4	3.7	0.03
	Average hours child is with mother per day	2.9	2.6	0.01	3.1	2.6	0.00
	Average hours child is with father per day	2.2	1.9	0.02	2.3	1.9	0.00
	Average time child spends along per day	1.4	0.9	0.00	1.3	1.0	0.01
Does your family have ___ in the home?	Bedroom	39%	22%	0.06	40%	21%	0.04
	Kitchen	36%	22%	0.13	39%	25%	0.13
	Living room	39%	23%	0.06	63%	35%	0.00
	Washroom	8%	6%	0.77	11%	9%	0.89
	Indoor toilet	26%	28%	0.84	30%	34%	0.61
	Radio	41%	19%	0.01	35%	20%	0.09
	Television	22%	12%	0.27	24%	11%	0.14
	Refrigerator	10%	4%	0.49	16%	6%	0.27
	Bicycle	9%	2%	0.44	10%	0%	0.31
	Motor bike	13%	2%	0.23	11%	3%	0.34
	Mobile Telephone	69%	53%	0.09	75%	57%	0.05
	Electricity	43%	16%	0.00	41%	15%	0.00
	Land	62%	87%	0.00	66%	89%	0.01
Livestock	80%	88%	0.33	84%	91%	0.45	

## Appendix A2: Background characteristics & SES in intervention and comparison areas

		Midline			Endline		
		ELM	Comparison	p-value	ELM	Comparison	p-value
Does your family have _____ in the home?	Storybooks	30%	19%	0.26	31%	17%	0.13
	Textbooks	23%	25%	0.83	39%	35%	0.67
	Magazines	19%	16%	0.77	14%	12%	0.83
	Religious books	49%	36%	0.17	55%	47%	0.43
	Coloring books	17%	7%	0.27	13%	6%	0.52
	Comics	15%	5%	0.29	10%	5%	0.60
	<b>Total number of types of reading materials</b>	1.47	0.91	<b>0.00</b>	1.46	1.03	<b>0.00</b>
	Homemade toys	43%	33%	0.28	32%	26%	0.49
	Shop-made toys	36%	19%	0.06	34%	19%	0.10
	Household objects as toys	78%	75%	0.72	81%	76%	0.64
	Outside toys	72%	76%	0.66	81%	72%	0.32
	Drawing toys	39%	22%	<b>0.05</b>	50%	29%	<b>0.02</b>
	Puzzle toys	22%	8%	0.14	15%	6%	0.35
	Hand-eye coordination toys	17%	9%	0.42	10%	8%	0.81
	Color/shape toys	17%	7%	0.30	9%	6%	0.75
	Number toys	81%	77%	0.68	10%	7%	0.75
<b>Total number of types of toys</b>	3.40	2.53	<b>0.00</b>	3.23	2.42	<b>0.00</b>	
In the past week, have you _____?	Read to child	50%	24%	<b>0.01</b>	57%	28%	<b>0.00</b>
	Tell child stories	57%	41%	0.10	64%	53%	0.25
	Sing to child	77%	59%	0.08	86%	63%	<b>0.02</b>
	Go out with child	76%	56%	<b>0.03</b>	79%	59%	<b>0.04</b>
	Play with child	74%	56%	0.07	80%	69%	0.26
	Draw with child	60%	40%	<b>0.04</b>	68%	44%	<b>0.01</b>
	Teach child something new	51%	27%	<b>0.01</b>	59%	36%	<b>0.01</b>
	Teach letters	53%	24%	<b>0.00</b>	61%	30%	<b>0.00</b>
	Teach numbers	52%	27%	<b>0.01</b>	60%	34%	<b>0.01</b>
	Tell the child they are loved	82%	74%	0.48	82%	76%	0.59
	<b>Total number of types of home learning activities</b>	4.62	2.85	<b>0.00</b>	5.05	3.41	<b>0.00</b>
	Spank the child	71%	65%	0.52	69%	69%	0.99
	Hit the child	53%	50%	0.74	45%	41%	0.71
	Yell or shout at the child	52%	40%	0.22	43%	28%	0.12
<b>Total number of types of negative discipline</b>	1.54	1.23	<b>0.01</b>	1.32	1.11	0.06	

## Appendix B1: Simplified multivariate model building process predicting Total IDELA (standard errors clustered at community level)

	(1)	(2)	(3)	(4)	(5)
Variables	Total IDELA	Total IDELA	Total IDELA	Total IDELA	Total IDELA
Midline	0.451***	0.441***	0.436***	0.451***	0.442***
Endline	0.595***	0.582***	0.576***	0.595***	0.584***
South Omo	0.109***	0.116***	0.104***	0.095***	0.093*
South Omo x Midline	-0.204***	-0.203***	-0.191***	-0.204***	-0.196**
South Omo x Endline	-0.263***	-0.263***	-0.249***	-0.266***	-0.256***
Female	-0.003				
#types of possessions		0.018***			
#types of reading materials			0.016***		0.009*
Amharic speaking				0.090***	0.082**
Constant	0.287***	0.220***	0.270***	0.259***	0.251***
R-squared	0.643	0.659	0.652	0.676	0.680
Number of observations	751	756	756	756	756

\*\*\* p<0.001, \*\* p<0.01, \* p<0.05



## Appendix B2: Multivariate models predicting IDELA Domains (complex model, standard errors clustered at the community level)

	Motor	Emergent Literacy	Emergent Numeracy	Social-Emotional	IDELA Total	Approaches to Learning
<b>Midline</b>	0.462***	0.464***	0.480***	0.439***	0.464***	0.606***
<b>Endline</b>	0.588***	0.623***	0.597***	0.632***	0.612***	0.652***
<b>South Omo</b>	0.236***	0.090*	0.078~	0.044	0.093*	0.336***
<b>South Omo x Midline</b>	-0.307***	-0.155*	-0.195***	-0.158*	-0.184**	-0.417***
<b>South Omo x Endline</b>	-0.455***	-0.278***	-0.076~	-0.300***	-0.258***	-0.440***
<b>#types of possessions</b>	0.007	0.003	0.010~	0.016~	0.010~	-0.004
<b>#types of reading materials</b>	0.011	0.014*	0.005	0.003	0.013~	0.037*
<b>#types of reading materials x Midline</b>	-0.003	0.004	0.003	-0.003	-0.005	-0.033
<b>#types of reading materials x Endline</b>	-0.006	-0.012	-0.006	-0.014	-0.014~	-0.040*
<b>Amharic speaking</b>	0.093	0.117**	0.064	0.127*	0.108*	0.052
<b>Amharic speaking X Midline</b>	-0.061	-0.056	-0.053~	-0.095	-0.074	0.010
<b>Amharic speaking X Endline</b>	0.048	-0.033	-0.021	-0.074*	-0.028	0.035
<b>Constant</b>	0.250***	0.111*	0.279***	0.196***	0.205***	0.341***
<b>R-squared</b>	0.483	0.628	0.715	0.546	0.690	0.410
<b>Number of observations</b>	756	756	756	755	756	744

\*\*\* p<0.001, \*\* p<0.01, \* p<0.05, ~ p<0.1

**Appendix B3: Multivariate models predicting IDELA Domains  
(parsimonious model, standard errors clustered at the community level)**

	<b>Motor</b>	<b>Emergent Literacy</b>	<b>Emergent Numeracy</b>	<b>Social- Emotional</b>	<b>IDELA Total</b>	<b>Approaches to Learning</b>
<b>Midline</b>	0.443***	0.458***	0.474***	0.415***	0.442***	0.567***
<b>Endline</b>	0.596***	0.589***	0.583***	0.591***	0.584***	0.604***
<b>South Omo</b>	0.233***	0.093*	0.075	0.044	0.093*	0.343***
<b>South Omo x Midline</b>	-0.316***	-0.167**	-0.203***	-0.173*	-0.196**	-0.419***
<b>South Omo x Endline</b>	-0.445***	-0.274***	-0.074	-0.305***	-0.256***	-0.430***
<b>#types of reading materials</b>	0.010~	0.013*	0.007*	0.002	0.009*	0.010
<b>Amharic speaking</b>	0.095***	0.090***	0.049	0.085~	0.082**	0.064
<b>Constant</b>	0.277***	0.132***	0.320***	0.269***	0.251***	0.351***
<b>R-squared</b>	0.473	0.624	0.710	0.529	0.680	0.402
<b>Number of observations</b>	756	756	756	755	756	744
*** p<0.001, ** p<0.01, * p<0.05						

## Appendix C1: Estimates of impact on IDELA using multivariate regression at midline

	Motor	Emergent Literacy	Emergent Numeracy	Social-Emotional	Total IDELA
<b>ELM</b>	0.268***	0.285***	0.300***	0.329***	0.296***
<b>South Omo</b>	-0.004	0.040	-0.030	-0.052	-0.012
<b>Father is literate</b>	-0.134*	-0.170**	-0.077~	-0.086~	-0.117**
<b>Mother is literate</b>	0.062	0.091*	0.032	0.065	0.063
<b>Age</b>	0.018	0.013	0.032*	0.008	0.018
<b>Number of children in family</b>	-0.007	-0.007	0.004	-0.012	-0.005
<b>Father's age</b>	0.002	0.001	0.001	0.002	0.001
<b>Mother's age</b>	0.001	0.004*	0.001	0.004*	0.002
<b>Mother went to primary school</b>	0.001	-0.021	0.014	-0.008	-0.004
<b>Father went to primary school</b>	0.023	0.040~	0.013	0.018	0.023
<b>Number of home learning activities</b>	0.007	0.009	0.003	0.008	0.007
<b>Number of negative discipline</b>	0.001	-0.007	-0.003	-0.013	-0.006
<b>Number of reading materials</b>	0.012	0.024*	0.001	0.011	0.012
<b>Number of types of toys</b>	-0.005	-0.005	0.003	-0.003	-0.002
<b>Constant</b>	0.279*	0.038	0.238	0.175	0.183
<b>R-sq</b>	0.417	0.478	0.492	0.526	0.565
<b>N</b>	380	380	380	380	380

	<b>Motor</b>	<b>Emergent Literacy</b>	<b>Emergent Numeracy</b>	<b>Social-Emotional</b>	<b>Total IDELA</b>
<b>ELM</b>	0.187***	0.224***	0.266***	0.244***	0.230***
<b>South Omo</b>	-0.196***	-0.166***	-0.013	-0.267***	-0.160***
<b>Father is literate</b>	-0.006	-0.003	-0.000	0.052	0.011
<b>Mother is literate</b>	-0.032	0.035	0.001	-0.007	-0.001
<b>Age</b>	0.029	0.027*	0.018*	0.021~	0.023*
<b>Number of children in family</b>	-0.019*	-0.004	-0.008~	-0.015*	-0.012*
<b>Father's age</b>	0.002	0.002~	0.000	0.003~	0.002
<b>Mother's age</b>	0.001	-0.002	0.000	0.001	-0.000
<b>Mother went to primary school</b>	0.031~	0.008	0.005	0.022	0.017
<b>Father went to primary school</b>	0.015	0.013	0.010	-0.003	0.009
<b>Number of home learning activities</b>	-0.003	-0.002	0.000	-0.002	-0.002
<b>Number of negative discipline</b>	0.019	0.004	0.002	-0.007	0.005
<b>Number of reading materials</b>	-0.002	-0.002	0.001	-0.013~	-0.004
<b>Number of types of toys</b>	0.003	-0.001	-0.004	0.005	0.001
<b>Constant</b>	0.681***	0.542***	0.590***	0.722***	0.634***
<b>R-sq</b>	0.512	0.622	0.690	0.624	0.704
<b>N</b>	380	380	380	380	380