USAID
FROM THE AMERICAN PEOPLE

Addressing Education in Northeast Nigeria

Non-Formal Learning Center Cohort 1: Endline Assessment

Approved Date:

Agreement Number: 72062018CA00001
Activity Start Date and End Date: September 28, 2018 to September 27, 2021
Total Award Amount: $14,876,124
AOR Name: Wale Samuel

Submitted by: Dalhatu Sulaiman Darazo, Chief of Party
FHI 360
No. 4, Hamman Maiduguri Street, Ali Kotoko Close, off Lagos Street, behind MTN office, old GRA, Maiduguri, Borno, Nigeria
Tel: (234) 9-4615555
Email: DSulaiman@FHI 360.org

AENN NFLC Cohort 1 Endline Assessment

March 2020

Contracted under 72062018CA00001

AENN

ACRONYMS

ACES	Assessment of Children's Emotional Skills
AENN	Addressing Education in Northeast Nigeria
AOR	Agreement Officer Representative
CBO	Community-Based Organization
CWPM	Correct words per minute
EGMA	Early Grade Mathematics Assessment
EGRA	Early Grade Reading Assessment
HLE	Home Literacy Environment
IDP	Internally Displaced Person
LGAs	Local Government Areas
LSPM	Letter sounds per minute
MOE	Ministry of Education
NFLC	Non-Formal Learning Center
ORF	Oral reading fluency
PCA	Principal Components Analysis
SEL	Socio-emotional Learning
SES	Socioeconomic Status
UNICEF	United Nations Children's Fund
USAID	United States Agency for International Development

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This report summarizes changes in academic and social emotional learning among a representative sample of learners who participated in the USAID Addressing Education in Northeast Nigeria (AENN) Activity's first cohort of basic level non-formal learning centers (NFLCs). The first cohort was implemented across Borno and Yobe states from July 2019 through January 2020, targeting out-of-school children and youth aged 6-15. Within the NFLCs, AENN uses a condensed curriculum that is aligned to the formal curriculum and is divided into two levels (Basic Literacy, which is equivalent to primary 1-3, and Post-Literacy, which his equivalent to primary 4-6, which will be offered in Year 2), according to the national non-formal education policy. Students attending the NFLC Basic Literacy program in Cohort 1 received classes four times per week for two hours and 15 minutes per day for approximately seven months. Each day, learners participated in 45-minute lessons for Hausa literacy, mathematics, and social emotional learning (SEL).

Data Collection Strategy

As part of program implementation, the AENN team administered baseline and endline assessments with a sample of learners across AENN implementation sites. Assessment tools included the Early Grade Reading Assessment (EGRA), Early Grade Mathematics Assessment (EGMA), two assessments of social-emotional learning (Assessment of Children's Emotional Skills and Children's Stories), and a student-level Safer Learning Environments survey. Respondents included a cross-section of four randomly selected learners from every AENN NFLC across Yobe and Borno states. Baseline data for Cohort 1 NFLCs was collected between mid-July and mid-August 2019, and endline data was collected in mid-January 2020. This analysis is drawn from a sample of 1,700 learners assessed at baseline and endline from all NFLCs that had at least one record at baseline and endline.

Learner Demographics

The majority of learners sampled in both states at baseline and endline are female (between 53% and 63% across the two states and two assessments), which is consistent with the AENN target of reaching more females than males. In both states, the average age of learners sampled is approximately 11 years old, which is consistent with the AENN target group of children aged 6-15. The majority of learners sampled in Borno are IDPs (72%), while the majority of learners sampled in Yobe are host community members (~55%), which is consistent with the much higher proportion of Internally Displaced Persons (IDPs) in Borno state than in Yobe. More learners sampled at baseline and endline in Yobe live with both parents (81%, 76%) compared to learners sampled in Borno (58%, 60%), which is consistent with the much higher level of conflict-related disruption experienced in Borno state. Approximately two-thirds of the learners sampled in Yobe at baseline and endline speak Hausa at home, while in Borno, just over one-third of learners sampled speak Hausa and almost half speak Kanuri at home, which is consistent with the high prevalence of Kanuri speakers in Borno state, particularly among displaced populations. Across the sample, less than 10% of learners sampled report having physical or cognitive disabilities, with slightly more learners reporting cognitive disabilities.

Reading and Numeracy Outcomes

Overall, there are strong, statistically significant improvements across every reading and numeracy subtask. The mean oral reading fluency (ORF) score improved from 6.73 correct words per minute (CWPM) to 19.24 CWPM from baseline to endline, an increase of 0.52 standard deviations. At baseline, the mean letter sound score was 6.15 letter sounds per minute (LSPM), and at endline, the mean letter sound score

improved by 16.51 LSPM, an increase of 0.76 standard deviations. From baseline to endline, the mean syllable scores saw comparable improvements, increasing by 15.70 SPM or 0.65 standard deviations. Additionally, mean reading comprehension scores increased by a statistically significant margin of 12.6 percentage points from baseline to endline. In terms of numeracy performance, learners had statistically significant improvements across all three subtasks: mean number identification scores improved by 21.3 percentage points; mean addition scores increased by 17.1 percentage points; and mean subtraction scores increased by 16.4 percentage points.

Results by State and Gender

The EGRA results by state and gender demonstrate that girls are starting at a lower baseline than boys but demonstrating a higher growth rate than boys, decreasing gender disparities in learning and improving equity. For example:

- In Borno, average oral reading fluency (ORF) scores increased by 10 correct words per minute (CWPM) for males and almost 14 CWPM for females from baseline to endline. In Yobe, average ORF scores increased by 11 CWPM for males and 14 CWPM for females.
- While the average Borno male could read 11 words per minute and the average Borno female could read just over 6 words per minute at baseline, by endline both males and females in Borno could read approximately 20 words per minute. Similar patterns are seen for average letter sound scores and average syllable scores.
- Boys in Yobe had a mean score of approximately 40% on the number identification subtask at baseline and 62% at endline, an increase of 22 percentage points.
- At baseline, girls in Yobe had a slightly lower mean score of 32% on the number identification subtask. However, at endline, the mean score for girls in Yobe increased by 30 percentage points to 62%, the same mean scores as boys in Yobe.

Results by Age Group

The analysis also examined mean differences in EGRA/MA subtask scores for learners ages 6-9 and learners ages 10-15 from baseline to endline and determined, as expected, that older learners demonstrate higher gains in learning. For example:

- Learners who are 6-9 years old had a mean ORF score of 4.3 words per minute at baseline and saw a significant increase to 10.3 words per minute by endline, while learners who are 10-15 years old had a mean ORF score of 7.6 words per minute at baseline and saw a significant increase to 20.9 words per minute by endline.
- Learners in the older age bracket had higher mean math scores than younger learners at baseline and endline, as well as slightly larger gains in mean math scores.

Results by Displacement Status

There are no major differences between IDPs and host communities in terms of learning outcomes, although it appears that host communities learn at a slightly faster rate than IDP children. This could be explained by the relative stability of host community children's lives in comparison with IDP children and also the fact that more IDPs are Kanuri speakers, while more host community members are Hausa speakers (see below).

Results by Language

In terms of reading performance, Hausa speakers have the highest mean scores at endline and have the largest margins of improvement from baseline to endline. This is not surprising given that the materials

were taught in Hausa (Kanuri materials have been developed and will be used for Cohort 2). However, the results for Hausa speakers are only slightly higher than Kanuri speakers, who have similar mean reading scores at baseline. Hausa and Kanuri speakers have slightly larger improvements in mean math scores compared to Fulani speakers and speakers of other languages.

Social-emotional Learning Outcomes

The results for Socio-emotional Learning (SEL) are less dramatic than for reading, but we still see improvements across all measures from baseline to endline. Across both states and genders, learners have slight improvements in mean Assessment of Children's Emotional Skills (ACES) scores from baseline to endline, with statistically significant improvements for males in both states. ACES scores were already relatively high at baseline. Mean ACES scores and the percentage of learners with perfect ACES scores are slightly higher for IDPs as compared to host community members and returnees. Additionally, ACES outcomes improved by a statistically significant margin for IDPs from baseline to endline.

In Borno, more learners display hostile attribution bias at both baseline and endline compared to learners in Yobe. However, the percentage of learners who display hostile attribution bias decreases by a statistically significant margin from baseline to endline for all groups except for Borno males. Conflict resolution strategies stay relatively stagnant from baseline to endline across all groups. Problem solving is the most used technique for all groups at both baseline and endline. Notably, Yobe males have a statistically significant increase in the use of problem solving and a statistically significant decrease in the use of aggression from baseline to endline. IDPs have the highest percentage of learners who display hostile attribution bias at both baseline and endline, however, this bias sees a statistically significant decrease from 42% to 33% from baseline to endline. The percentage of host community members who display hostile attribution bias also decreases by a statistically significant margin from 38% to 24% from baseline to endline.

School Safety Outcomes

Across all groups, there are increases in school climate and safety scores from baseline to endline, with all but one change being statistically significant. These scores are generally high at baseline and endline, with safety scores all over 94% for each group at endline.

Conclusions

Overall, although this is not an impact evaluation, the results from this pre-post assessment are quite clear: NFLC participants are learning, and at a high enough rate that it is reasonable to assume that some participants could feasibly mainstream into Grade 4 in the formal system or move on to Post-Basic NFLCs, which are primarily taught in English. While the academic results appear high in comparison to early grade reading programs from other parts of Northern Nigeria, it is important to remember that, on average, AENN targets an older age group, and older participants should learn at a faster rate than younger children. The SEL and school safety results are less dramatic, but still positive, and all results point in the direction of improved social-emotional skills and perceptions of school safety over the course of the seven month intervention. When correlating social-emotional learning indicators with reading performance, there does not appear to be a relationship. All of these results are very positive and paint a picture of an educational intervention that is working in the way that it was designed.

1. INTRODUCTION

Purpose

This report summarizes changes in academic and social emotional learning among a representative sample of learners who participated in AENN's first cohort of Basic level non-formal learning centers (NFLCs), from July 2019 through January 2020.

Activity Description

The AENN Activity targets the immediate educational needs of children and youth ages 6-15 who are internally displaced, and their host communities, in eight local government areas (LGAs) across Borno and Yobe, while building long-term resilience across the system. In Borno state, in Year 1, AENN was implemented in Maiduguri Metropolitan Council, Jere, Monguno, Hawul, and Dikwa. In Yobe, the LGAs include Damaturu, Potiskum, and Bade. The program was rolled out in 150 communities in Year 1. It is important to note that AENN ceased programming in Hawul midway through Year 1 due to low numbers of out-of-school children and the evident risk of drawing children out of the formal system. Therefore, children from Hawul were not included in the endline assessment and are therefore not included in this report.

To meet the overwhelming educational needs of out-of-school children and youth, AENN provides quality basic education for 6-15 year olds via non-formal learning centers (NFLCs). These classes are primarily delivered in formal school buildings in the afternoon (after formal school lesson is over) or in temporary learning spaces that have been established by United Nations Children's Fund (UNICEF) or other education actors. Community-Based Organizations (CBOs) that oversee the implementation of non-formal activities provide seating mats, WASH materials, and other support the process of making the classrooms safer and more comfortable for learners. Additionally, CBOs provide learners with snacks and female learners with hygiene kits to promote regular attendance. The size of classes ranges between 50-60 learners per class. Classes are delivered by trained local learning facilitators. In Year 1, all NFLCs were delivered in Hausa language (Kanuri language was added in Year 2). All non-formal education activities are accompanied by safety interventions and community-level activities that support enrollment and awareness around the importance of education.

Within the NFLCs, AENN uses a condensed curriculum that is aligned to the formal curriculum and is divided into two levels (Basic Literacy, which is equivalent to primary 1-3, and Post-Literacy, which his equivalent to primary 4-6) according to the national non-formal education policy. This assessment only examines students participating in Basic Literacy (NFLC Level 1). Students attending the NFLC Basic Literacy program receive classes 4 times per week for 2 hours and 15 minutes per day for approximately nine months (although due to time constraints the first cohort was completed in seven months). Each day, learners participate in 45-minute lessons for literacy, math, and social emotional learning (SEL). After completing the Basic Literacy program, learners can either mainstream into formal schools at grade 4 or continue into the two-year Post-Literacy program. After completing the Post-Literacy program, learners can mainstream into formal schools at grade 7 (junior secondary school).

NFLC Cohort 1 included 600 NFLCs across eight LGAs in Borno (MMC, Hawul, Jere, Dikwa, and Monguno) and Yobe (Damaturu, Potiskum, and Bade). The program began in July 2019 and concluded in January 2020. The baseline and endline assessments were carried out in July and August 2019, and January 2020 respectively.

2. DATA COLLECTION STRATEGY

As part of program implementation, the AENN team administered the EGRA/EGMA/SEL assessment and student survey to four randomly selected learners from a random sample of NFLCs across Yobe and Borno states. Baseline data for Cohort 1 NFLCs was collected from mid-July to mid-August, 2019. At baseline, data was collected for 1,721 learners from 396 NFLCs. Endline data collection occurred in January 2020. At endline, data was collected for 982 learners from 240 NFLCs. Of the 240 NFLCs with valid observations at endline, 197 matched with NFLCs with valid observations at baseline. Thus, our analysis draws from learners sampled from these NFLCs that were visited at baseline and endline.

Because there was significant attrition in NFLCs from baseline to endline, we test for non-random attrition to determine potential implications on external validity and broader generalizations that can be made using the results of the following analysis.[1] Overall, we find that the children's baseline characteristics and outcomes are stable between the NFLC sample that remained and the sample that dropped out. However, we note a few exceptions where the sample that dropped out had slightly higher ORF and ACES scores, which may potentially lead to a downward bias in the pre-post learning growth measures. More notable, however, is that NFLCs from Borno state were more likely to be among the NFLCs that were not visited at endline. As such, we caution that the findings of this report may not be generalizable to the entire AENN population due to the disproportionate geographic distribution of the dropped sample.

This study employs a pre-post analysis from baseline to endline across a variety of academic and social-emotional learning outcomes. We measure literacy outcomes using the Early Grade Reading Assessment (EGRA) in Hausa, including letter sounds, syllables, oral reading fluency (ORF), and reading comprehension. We use three Early Grade Mathematics (EGMA) subtasks to measure numeracy outcomes, including number identification, addition, and subtraction. We deploy several social-emotional learning (SEL) and safe learning environment modules, including the Assessment of Children's Emotional Skills (ACES) module to measure children's ability to correctly identify emotions of others. The Children's Stories module measures learners' hostile attribution bias, which is their intent to attribute hostile intent to a third party, as well as conflict resolution strategies, which identifies whether a learner uses aggressive or problem-solving strategies to resolve a conflict, or disengages from the conflict. Drawing from USAID's Safer Learning Environment Toolkit, we use a series of school climate questions to measure learners' perception of the teacher and classroom environment. We also ask a series of questions to measure learners' perceptions of safety in the classroom and community. We measure differences in all of these outcomes from baseline to endline using simple t-tests.

3. RESULTS: DESCRIPTIVE ANALYSIS

Learner Sample

This analysis is drawn from a sample of 1,700 learners assessed at baseline and endline in Cohort 1 NFLCs that had at least one record at baseline and endline. In total, 933 learners were sampled in Borno, with 505 sampled at baseline and 428 sampled at endline. In Yobe, a total of 767 learners were sampled, with 391 sampled at baseline and 376 sampled at endline. Table 1 shows that more learners were sampled in Borno than Yobe at both baseline and endline, which is consistent with the higher number of learners in Borno overall. Additionally, both states have a slight decline in the number of observations from baseline

[1] Refer to the Appendix for further detail regarding the attritition analysis.

to endline. The learner sample in Borno is comprised of approximately two-thirds of learners from MMC and approximately one-fourth of learners from Jere. While there were a small number of observations in Dikwa and Hawul at baseline, there were no observations from these LGAs at endline, so they were excluded from this analysis. In Yobe, approximately 40% of the sample is in Bade, 32% of the sample is in Damaturu, and approximately 28% of the sample is in Potiskum at both baseline and endline.

Table 1. Learner sample by state and LGA at baseline and endline

	Borno			Yobe	
LGA	Baseline	Endline	LGA	Baseline	Endline
Jere	0.25	0.28	Bade	0.40	0.41
MMC	0.66	0.63	Damaturu	0.32	0.32
Monguno	0.08	0.10	Potiskum	0.28	0.27
Observations	505	428	Observations	391	376

Learner Demographics

Table 2 displays learner demographic characteristics by state at baseline and endline. As seen in the table, the majority of learners sampled in both states at baseline and endline are female. The percentage of females sampled in Borno slightly increases from baseline to endline in Borno and stays relatively similar in Yobe. In both states, the average age of learners sampled at baseline is approximately 11 years old. As expected, average ages slightly increase in both states by endline. At both baseline and endline, the majority of learners sampled in Borno are IDPs (72%), while the majority of learners sampled in Yobe are host community members (55%, 56%). Approximately 70% of learners sampled in Borno at baseline and endline did not attend school last year. In Yobe, 85% of learners sampled at baseline did not attend school and 76% of learners sampled at endline did not attend school. More learners sampled at baseline and endline in Yobe live with both parents (81%, 76%) compared to learners sampled in Borno (58%, 60%). When comparing language spoken by learners at home, the table shows that approximately two-thirds of the learners sampled in Yobe at baseline and endline speak Hausa at home, while in Borno, just over one-third of learners sampled speak Hausa. Almost half of the learners sampled in Borno speak Kanuri at home. Across the sample, less than 10% of learners sampled report having physical or cognitive disabilities, with slightly more learners reporting cognitive disabilities. Learners with physical disabilities report having a lot of difficulty seeing, hearing, walking, and doing basic self-care, such as washing or dressing. Learners with cognitive disabilities report having a lot of difficulty remembering or communicating.

Table 2. Learner demographic characteristics by state at baseline and endline

	Baseline		Endline	
	Borno	Yobe	Borno	Yobe
Female	0.53	0.63	0.60	0.59
Age	11.12	11.04	12.03	11.61
Ate before school	0.85	0.81	0.88	0.89
Displacement status:				
IDP	0.72	0.42	0.72	0.40
Host community	0.25	0.55	0.24	0.56
Returnee	0.03	0.03	0.04	0.03
Attended formal school last year	0.23	0.14	0.24	0.22
Grade last year:				
KG2	0.01	0.00	0.00	0.00
Primary 1	0.10	0.03	0.11	0.14
Primary 2	0.09	0.04	0.05	0.02
Primary 3	0.05	0.02	0.04	0.01
Primary 4	0.02	0.03	0.04	0.03
Primary 5	0.03	0.01	0.04	0.03

Primary 6	0.02	0.03	0.05	0.01
Not in school last year	0.69	0.85	0.67	0.76
Lives with:				
Both parents	0.58	0.81	0.60	0.76
Single parent	0.29	0.12	0.31	0.17
Other family member	0.13	0.07	0.08	0.07
Other adult	0.01	0.00	0.01	0.00
Language spoken at home:				
Hausa	0.36	0.67	0.42	0.62
Fulani	0.01	0.12	0.00	0.14
English	0.00	0.00	0.00	0.00
Kanuri	0.47	0.13	0.44	0.16
Other	0.16	0.08	0.13	0.09
Disability:				
Has difficulty seeing	0.00	0.01	0.00	0.01
Has difficulty hearing	0.01	0.01	0.00	0.02
Has difficulty walking	0.01	0.02	0.02	0.02
Has difficulty dressing	0.01	0.02	0.00	0.01
Has difficulty remembering	0.03	0.05	0.01	0.02
Has difficulty communicating	0.04	0.06	0.05	0.05
Has a physical disability	0.03	0.06	0.02	0.05
Has a cognitive disability	0.07	0.09	0.06	0.07
Observations	896		804	

Table 3 shows that slightly more learners sampled in Yobe work outside the home or in the fields compared to learners sampled in Borno at both baseline and endline. This percentage slightly decreases across both states from baseline to endline. The measures for type of work only show responses from learners who responded that they do work at home or in the fields. We see that of the learners sampled who work, the majority in both states work in a household for another family. In terms of attendance, we see that 68% of learners sampled in Borno at baseline and endline had zero absences last week. This percentage is slightly higher for learners sampled in Yobe at both baseline (74%) and endline (81%). In Borno, approximately 90% of learners sampled spend time with friends outside of school at least one day a week at both baseline and endline. In Yobe, this percentage is slightly less, at approximately 80%.

Table 3. Learner demographic characteristics by state at baseline and endline (cont.)

	Baseline		**Endline**	
	Borno	Yobe	Borno	Yobe
Works outside home or in fields	0.28	0.36	0.19	0.29
Type of work:				
Work in household for another family	0.54	0.39	0.63	0.41
Work in market	0.17	0.11	0.16	0.15
Work in a store	0.07	0.01	0.01	0.03
Work in the fields	0.08	0.15	0.04	0.27
Other work	0.13	0.34	0.16	0.14
Days absent last week:				
No absences	0.68	0.74	0.68	0.81
One day	0.12	0.10	0.17	0.05
Two days	0.10	0.05	0.08	0.05
Three days	0.05	0.02	0.03	0.04
Four days	0.02	0.07	0.02	0.03
Five days	0.01	0.01	0.00	0.01
NFLC was not open last week	0.02	0.01	0.01	0.00
Time spent with friends outside school:				
Never	0.09	0.22	0.10	0.20

One day per week	0.08	0.09	0.06	0.06
Two days per week	0.14	0.13	0.10	0.16
Three days per week	0.18	0.12	0.27	0.24
Four days per week	0.07	0.05	0.07	0.09
Five days per week	0.06	0.04	0.05	0.02
Six days per week	0.04	0.02	0.03	0.02
Seven days per week	0.35	0.32	0.31	0.20
Observations	896		804	

Table 4 shows the socioeconomic status (SES) and home literacy environment (HLE) of learners sampled in Borno and Yobe at baseline and endline. To determine household SES, we ask learners if they have a series of seven relevant household assets. We then employ Principal Components Analysis (PCA) to construct an SES (wealth) index. The index is created as a factor that reduces the dimensionality of the household ownership across the seven household assets. As such, the SES index is measured in terms of standard deviations from the mean, where values less than zero denote lower SES and values above zero denote relatively higher SES. As seen in the table below, a majority of learners sampled in both states at baseline and endline have low SES. Slightly more learners sampled in Borno at endline have low SES (76%) compared to Borno learners sampled at baseline (70%). When comparing learners sampled in Yobe to learners sampled in Borno, we see slightly fewer learners have low SES at both baseline (63%) and endline (56%).

To determine learners' HLE, we employ similar methods that were used to determine SES. To construct the HLE index, we use three relevant items: 1) if the student reports being read to at home, 2) if the student reports receiving help with homework at home, and 3) if the student has materials to read at home. We then employ PCA to construct an HLE index, which reduces the dimensionality across the three items. As such, the HLE index is measured in terms of standard deviations from the mean, where values less than zero denote lower HLE and values above zero denote relatively higher HLE. Table 4 shows that more learners sampled in Yobe have low HLE at baseline and endline compared to learners in Borno. The percentage of learners sampled with low HLE slightly increases from baseline to endline in both states.

Table 4. Learner SES and HLE by state at baseline and endline

	Baseline		**Endline**	
	Borno	Yobe	Borno	Yobe
SES:				
Low SES	0.70	0.63	0.76	0.56
High SES	0.30	0.37	0.24	0.44
Wealth Index	-0.12	0.01	-0.09	0.05
Home Learning Environment:				
Read to at home	0.45	0.31	0.49	0.35
Receives help with homework at home	0.42	0.29	0.46	0.29
Has materials to read at home	0.29	0.21	0.35	0.16
Low HLE	0.46	0.65	0.50	0.69
High HLE	0.54	0.35	0.50	0.31
HLE Index	0.11	-0.15	0.14	-0.19
Observations	896		804	

Table 5 shows incidences of victimization experienced by learners sampled in the last two weeks. Two items comprise being victimized by another child: 1) being criticized by other children, and 2) being hit by other children. Similarly, two items comprise being victimized by an adult: 1) being pinched by an adult, and 2) being hit with an object by an adult. For each individual item, we see decreases from baseline to endline across both states. In Borno, we see slightly higher levels of victimization compared to Yobe. The

percentage of learners victimized in the last two weeks is the percentage of learners who experienced at least one type of victimization in the last two weeks. As seen in the table, 49% of learners sampled in Borno experienced victimization in the last two weeks at baseline, compared to 40% at endline. In Yobe, 26% of learners sampled experienced victimization in the last two weeks at baseline, compared to 18% at endline.

Table 5. Learner victimization by state at baseline and endline

	Baseline		Endline	
	Borno	Yobe	Borno	Yobe
Victimization:				
Criticized by other children	0.31	0.15	0.22	0.12
Hit by other children	0.34	0.16	0.22	0.11
Pinched by adult	0.21	0.12	0.18	0.09
Hit with object by adult	0.19	0.11	0.15	0.08
Victimized by other children	0.42	0.21	0.31	0.15
Victimized by adults	0.29	0.15	0.25	0.11
Victimized in last two weeks	0.49	0.26	0.40	0.18
Observations	896		804	

Outcomes: EGRA/MA

In this section, we calculate mean EGRA/MA subtask scores by state and gender and test for statistical significance of mean differences from baseline to endline. The specific reading subtask measures include letter sounds identified correctly per minute, syllables read correctly per minute, words (from a passage) read correctly per minute, and reading comprehension questions answered correctly. We also measure the percentage of learners who receive zero scores on each of these subtasks. For math, we assess learners on their ability to identify numbers correctly, complete addition exercises, and complete subtraction exercises. Each of the three math subtasks are scored by adding the number of questions that learners answered correctly and dividing by 20, which is the total number of questions for each math subtask.

Overall EGRA/MA Outcomes

Table 6 displays the difference in mean EGRA/MA outcomes from baseline to endline for the entire sample. The table shows that there are statistically significant improvements across each reading and numeracy subtask. At baseline, the mean letter sound score was 6.149 letter sounds per minute (LSPM). At endline, the mean letter sound score improved by 16.511 LSPM, an increase of 0.76 standard deviations. From baseline to endline, the mean syllable scores saw comparable improvements, increasing by 15.702 SPM or 0.65 standard deviations. The mean ORF score improved from 6.734 correct words per minute (CWPM) to 19.241 CWPM from baseline to endline, an increase of 0.52 standard deviations. Additionally, mean reading comprehension scores increased by a statistically significant margin of 12.6 percentage points from baseline to endline. For reading subtasks, the percentage of learners who received zero scores on all four subtasks decreased by a statistically significant margin from baseline to endline. For example, the percentage of learners who received zero scores of the ORF subtask decreased from 53.2% at baseline to 25.8% at endline, a change of 27.4 percentage points.

In terms of numeracy performance, learners had statistically significant improvements across all three subtasks, with learners performing better on number identification compared to addition, and better on addition compared to subtraction. Mean number identification scores improved from 43.3% at baseline to 64.7% at endline. Mean addition scores increased by 17.1 percentage points from baseline to endline, while mean subtraction scores increased by 16.4 percentage points. The percentage of learners who received

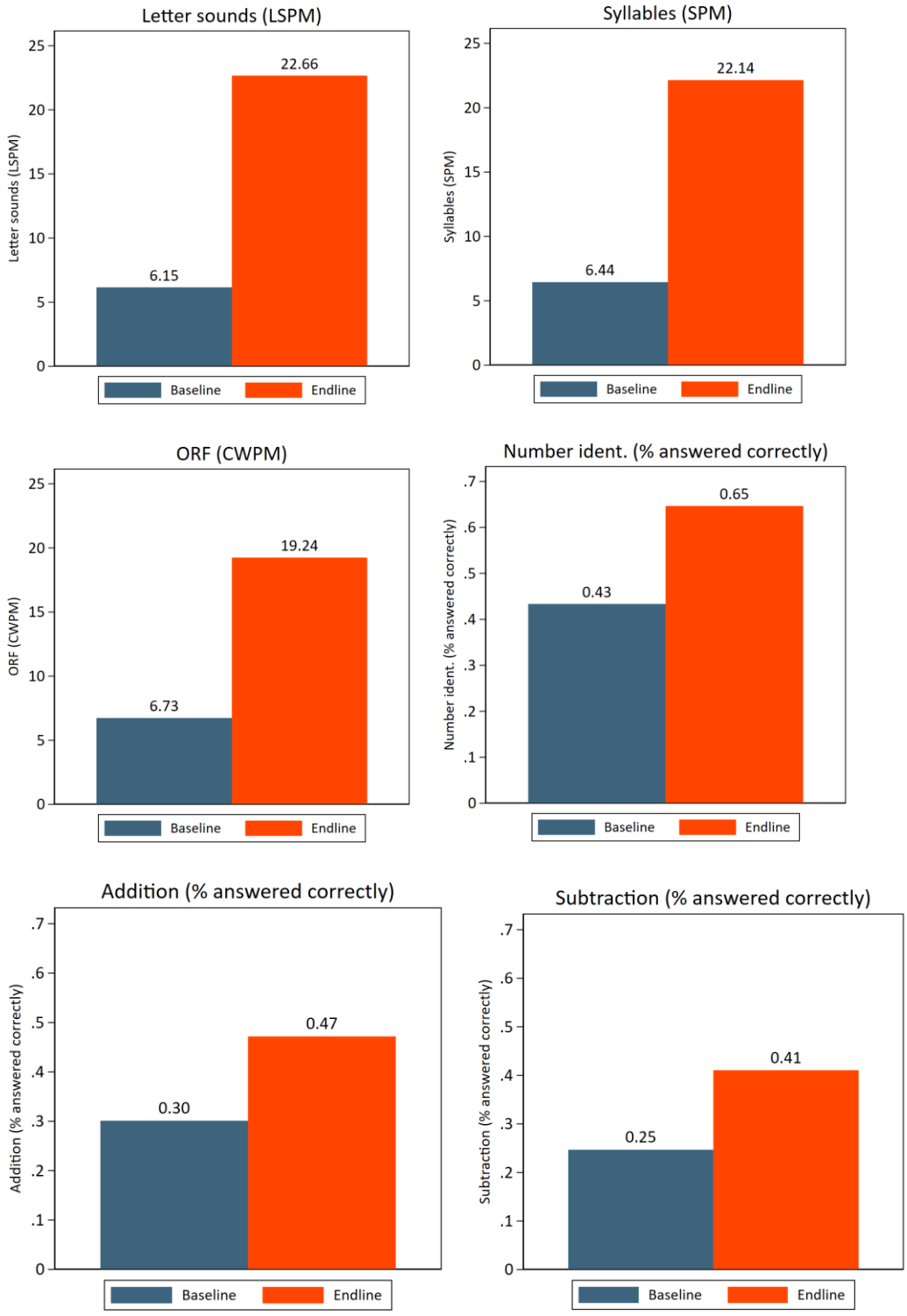
zero scores on all three math subtasks decreased from baseline to endline. The percentage of learners with zero scores on number identification decreased by 6 percentage points and the percentage of learners with zero scores on addition decreased by 11.4 percentage points. The decrease in learners with zero scores on math subtasks was greatest for subtraction, which had 29.5% of learners receiving zero scores at baseline and only 14% of learners receiving zero scores on subtraction at endline.

Table 6. Overall EGRA/EGMA outcomes at baseline and endline

	Baseline	Endline	Diff.
Reading Performance:			
Letter sounds (LSPM)	6.149	22.660	16.511***
Syllables (SPM)	6.441	22.142	15.702***
ORF (CWPM)	6.734	19.241	12.507***
Reading comp. (% answered correctly)	0.254	0.380	0.126***
Letter sounds (% with zero scores)	0.393	0.133	-0.261***
Syllables (% with zero scores)	0.488	0.224	-0.264***
ORF (% with zero scores)	0.532	0.258	-0.274***
Reading comp. (% with zero scores)	0.342	0.179	-0.163***
Numeracy Performance:			
Number ident. (% answered correctly)	0.433	0.647	0.213***
Addition (% answered correctly)	0.301	0.472	0.171***
Subtraction (% answered correctly)	0.247	0.410	0.164***
Number ident. (% with zero scores)	0.142	0.075	-0.067***
Addition (% with zero scores)	0.221	0.107	-0.114***
Subtraction (% with zero scores)	0.295	0.140	-0.155***
Observations	896	804	

Figure 1 displays the mean scores at baseline and endline as presented in the table above. These visuals show the substantial increases in mean EGRA/MA scores from baseline to endline for the entire sample. At baseline, mean EGRA subtask scores are approximately 6 and increase to approximately 20 at endline. For each EGMA subtask, mean scores increase by approximately 16-21 percentage points from baseline to endline.

Figure 1. Mean EGRA/MA subtask scores at baseline and endline



EGRA/MA *Outcomes by State and Gender*

Table 6 shows that there are statistically significant improvements in all reading subtasks from baseline to endline for girls in both states. Additionally, we see that boys in Yobe have statistically significant

improvements across all reading subtasks from baseline to endline. Boys in Borno have statistically significant improvements across all reading subtasks, except for the percentage of learners who receive zero scores on the reading comprehension subtask. Learners in Borno slightly outperform learners in Yobe on most reading subtasks. Additionally, girls have higher gains than boys from baseline to endline across almost all reading subtasks, decreasing the disparity in reading outcomes by gender. For example, in Borno, average ORF scores increase by 10 CWPM for males and almost 14 CWPM for females from baseline to endline. In Yobe, average ORF scores increase by 11 CWPM for males and 14 CWPM for females. As such, while the average Borno male could read 11 words per minute and the average Borno female could read just over 6 words per minute at baseline, both males and females in Borno could read approximately 20 words per minute by endline. Similar patterns are seen for average letter sound scores and average syllable scores.

In terms of the percentage of learners who receive zero scores on reading subtasks, we see statistically significant decreases in almost all categories, with girls having larger decreases than boys across all reading subtasks. For example, approximately 35% of boys in Borno could not identify a single letter sound at baseline. This percentage decreased by a statistically significant margin of 20 percentage points from baseline to endline, meaning that at endline, only 14% of Borno male learners could not identify a single letter sound. When comparing this percentage to Borno female learners, we see that 47% of female learners in Borno could not identify a single letter sound at baseline. Yet, the percentage of female learners in Borno who could not identify a single letter sound at endline was only 13%, a drop of almost 34 percentage points from baseline to endline. Again, we see similar patterns for zero scores across each reading subtask in both states.

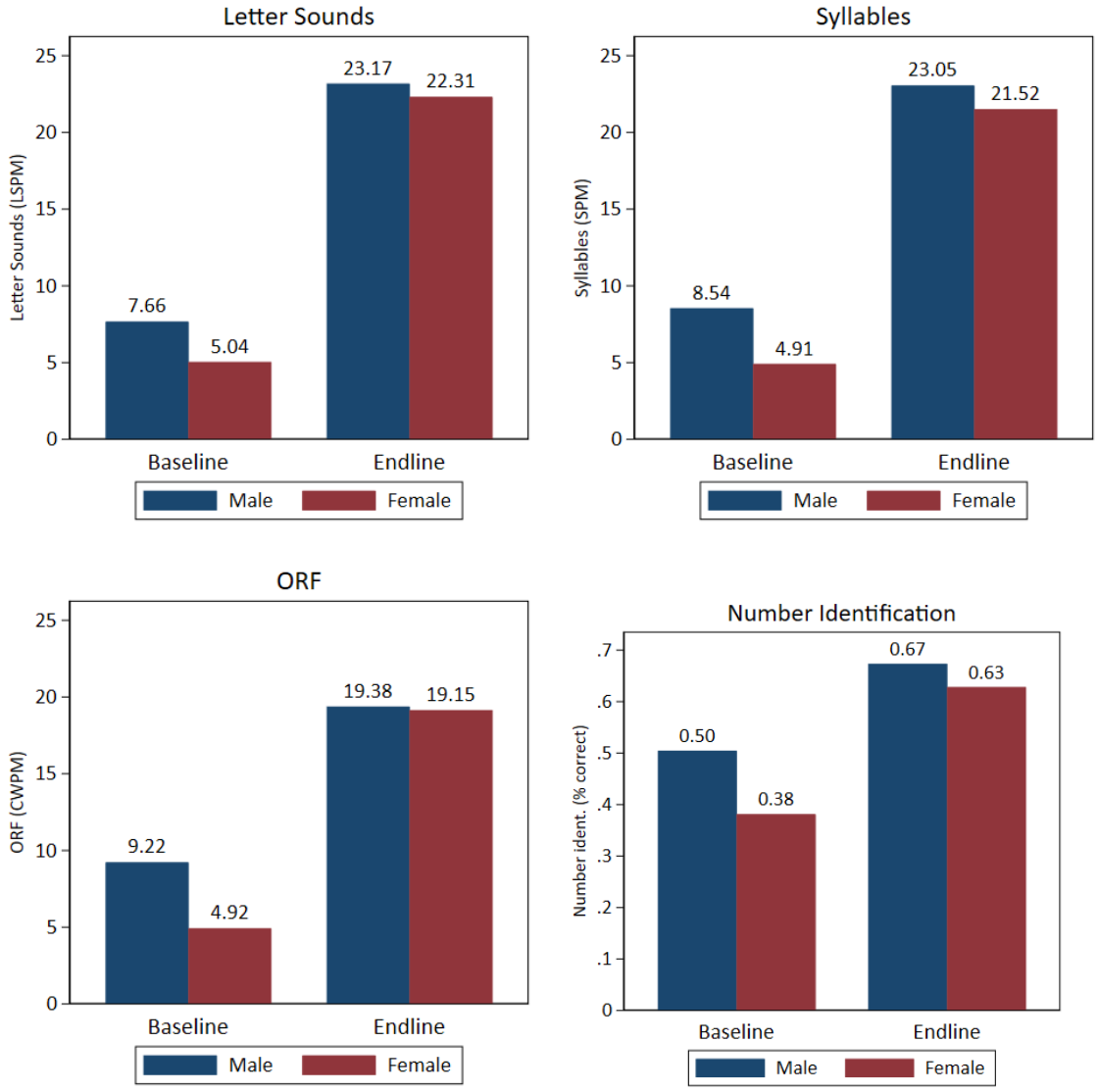
Across each of the three math subtasks, we see statistically significant improvements in average scores from baseline to endline for boys and girls in each state. Across both states and genders, learners perform best in number identification, followed by addition, then subtraction. Again, learners in Borno slightly outperform learners in Yobe across most math subtasks. When comparing average math scores by gender, we see that girls in both states have larger gains from baseline to endline compared to boys, decreasing the disparity in math outcomes by gender. For example, boys in Yobe had a mean score of approximately 40% on the number identification subtask at baseline and 62% at endline, an increase of 22 percentage points. At baseline, girls in Yobe had a slightly lower mean score of 32% on the number identification subtask. However, at endline, the mean score for girls in Yobe increased by 30 percentage points to 62%, the same mean scores as boys in Yobe at endline. In terms of zero scores on the math subtasks, almost all decreases in the percentage of learners receiving zero scores are statistically significant. We see similar patterns of Borno learners slightly outperforming Yobe learners and girls having slightly larger decreases than boys.

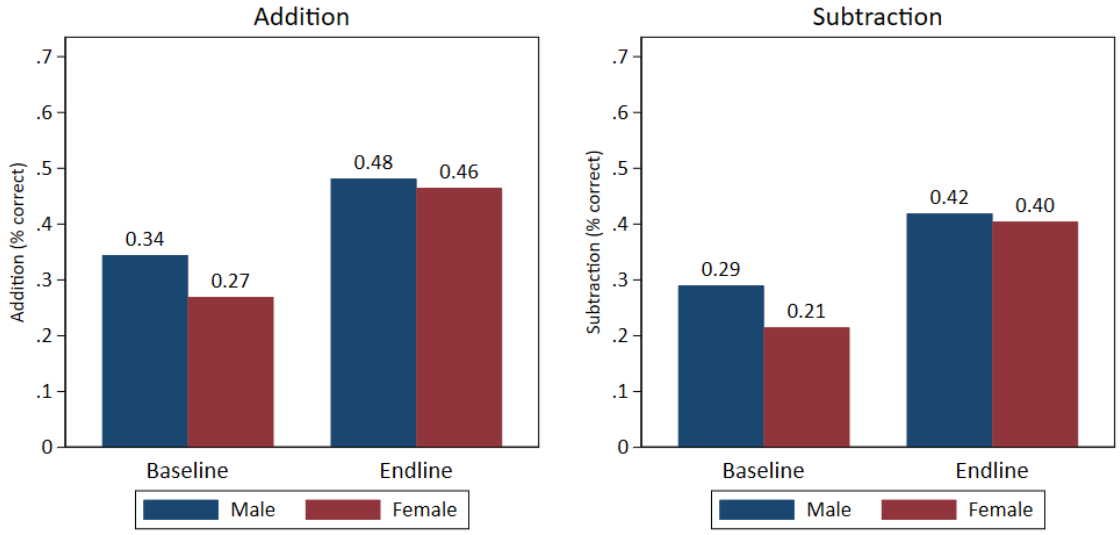
Table 7. EGRA/MA outcomes by state and gender at baseline and endline

	Borno						Yobe					
	Male			Female			Male			Female		
	Baseline	Endline	Diff.	Baseline	Endline	Diff.	Baseline	Endline	Diff.	Baseline	Endline	Diff.
Reading Performance:												
Letter sounds (LSPM)	9.036	23.885	14.849***	5.265	23.049	17.784***	5.417	22.386	16.968***	4.794	21.456	16.662***
Syllables (SPM)	10.292	24.675	14.383***	6.282	22.682	16.400***	5.658	21.263	15.605***	3.428	20.159	16.732***
ORF (CWPM)	11.128	21.208	10.080***	6.520	20.470	13.950***	6.084	17.379	11.295***	3.192	17.617	14.425***
Reading comp. (% answered correctly)	0.291	0.356	0.065*	0.256	0.402	0.146***	0.240	0.398	0.158***	0.165	0.363	0.198***
Letter sounds (% with zero scores)	0.346	0.142	-0.204***	0.472	0.133	-0.339***	0.308	0.124	-0.184***	0.402	0.132	-0.271***
Syllables (% with zero scores)	0.382	0.201	-0.181***	0.525	0.216	-0.309***	0.444	0.260	-0.184***	0.575	0.226	-0.349***
ORF (% with zero scores)	0.424	0.193	-0.231***	0.561	0.244	-0.317***	0.507	0.303	-0.205***	0.616	0.292	-0.325***
Reading comp. (% with zero scores)	0.257	0.203	-0.054	0.346	0.190	-0.156***	0.350	0.128	-0.222***	0.550	0.176	-0.374***
Numeracy Performance:												
Number ident. (% answered correctly)	0.570	0.725	0.155***	0.439	0.634	0.195***	0.398	0.617	0.220***	0.319	0.621	0.302***
Addition (% answered correctly)	0.378	0.515	0.138***	0.289	0.476	0.186***	0.291	0.445	0.153***	0.247	0.453	0.205***
Subtraction (% answered correctly)	0.337	0.430	0.093***	0.235	0.404	0.169***	0.215	0.407	0.192***	0.194	0.405	0.211***
Number ident. (% with zero scores)	0.083	0.030	-0.053**	0.129	0.028	-0.101***	0.177	0.145	-0.033	0.193	0.116	-0.077**
Addition (% with zero scores)	0.133	0.042	-0.091***	0.226	0.058	-0.168***	0.221	0.191	-0.031	0.298	0.153	-0.144***
Subtraction (% with zero scores)	0.169	0.079	-0.090***	0.297	0.115	-0.182***	0.307	0.199	-0.108**	0.402	0.172	-0.230***
Observations	235	170		270	258		144	155		247	221	

Figure 2 displays mean EGRA/MA subtask scores by gender at baseline and endline. This figure shows the significant increases in mean scores for both males and females across reading and math subtasks. It also shows the decreased disparity in mean score differences between girls and boys from baseline to endline. For example, boys read an average of 9.22 CWPM on the ORF subtask at baseline, while girls read a slightly lower 4.92 CWPM. However, at endline, boys and girls were both reading just over 19 CWPM, on average.

Figure ***2. Mean EGRA/MA subtask scores by gender at baseline and endline***





EGRA/MA Outcomes by Displacement Status

Table 8 shows differences in mean EGRA/MA scores from baseline to endline by displacement status. This table also shows differences in the percentage of learners with zero scores from baseline to endline. The results show statistically significant improvements across all subtasks for learners in each of the three categories- IDP, host community, and returnee. Additionally, returnees have the highest gains across all reading and math subtasks, however it is important to note the small sample size. When comparing the reading performance of IDPs and host community members, we see that host community members outperform IDPs in all subtasks, except for reading comprehension. Although IDPs and host community members all begin with similar scores on letter sounds, syllables, and ORF, host community members have slightly higher scores across these three subtasks at endline. Host community members also have larger decreases in the percentage of learners with zero scores across these three subtasks from baseline to endline. In terms of numeracy performance, host community members have slightly higher gains in number identification compared to IDPs, and both groups perform similarly in addition and subtraction.

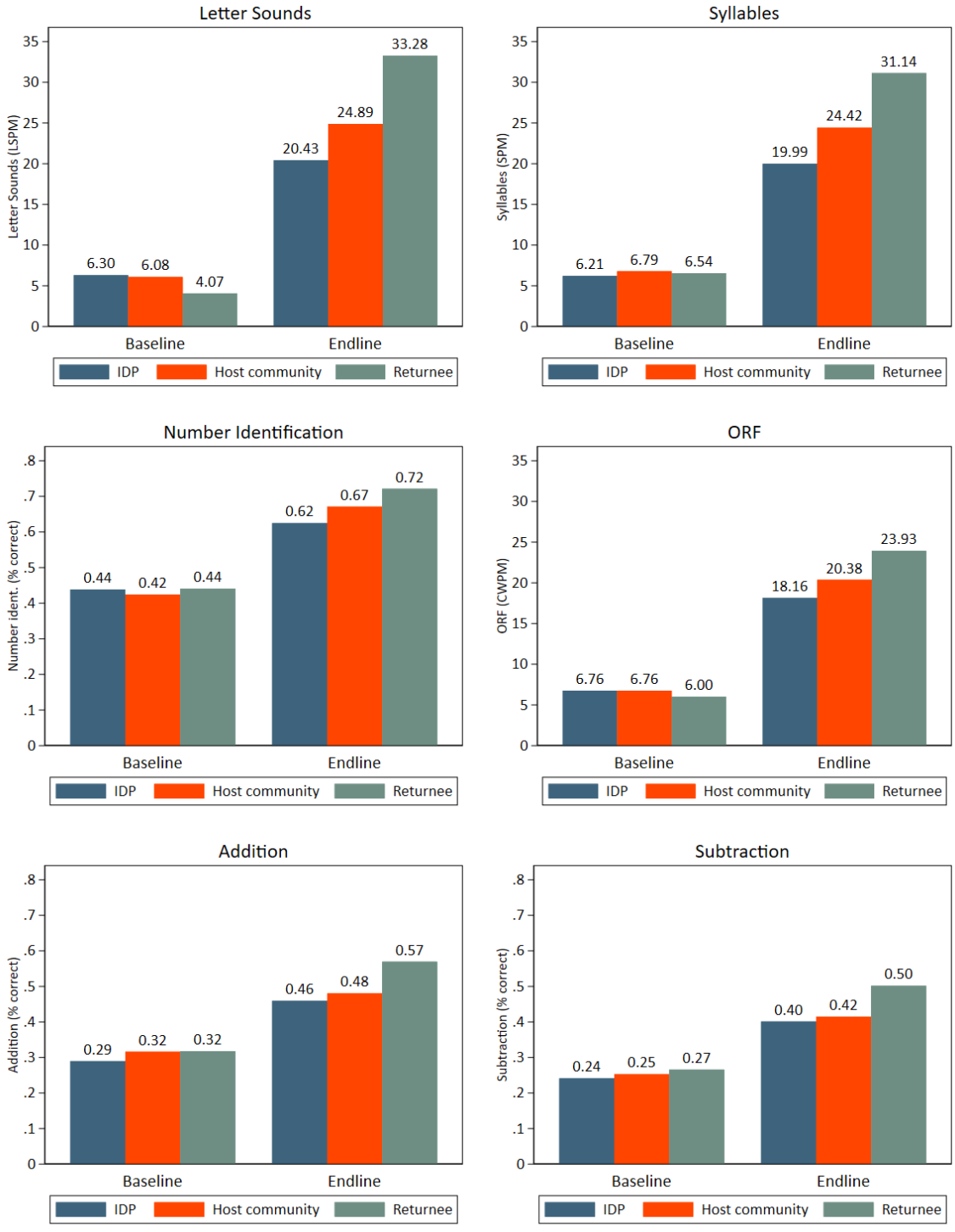
Table 8. EGRA/MA outcomes by displacement status at baseline and endline

	IDP			Host Community			Returnee		
	Baseline	Endline	Diff.	Baseline	Endline	Diff.	Baseline	Endline	Diff.
Reading Performance:									
Letter sounds (LSPM)	6.301	20.434	14.133***	6.083	24.890	18.807***	4.071	33.278	29.207***
Syllables (SPM)	6.208	19.993	13.784***	6.793	24.420	17.627***	6.536	31.138	24.602***
ORF (CWPM)	6.758	18.163	11.405***	6.757	20.376	13.619***	6.000	23.927	17.927***
Reading comp. (% answered correctly)	0.240	0.387	0.148***	0.287	0.372	0.085**	0.160	0.373	0.213*
Letter sounds (% with zero scores)	0.402	0.167	-0.234***	0.372	0.089	-0.283***	0.500	0.069	-0.431***
Syllables (% with zero scores)	0.489	0.254	-0.235***	0.481	0.188	-0.292***	0.571	0.138	-0.433***
ORF (% with	0.532	0.275	-0.257***	0.529	0.236	-0.292***	0.571	0.214	-0.357***

zero scores)									
Reading comp. (% with zero scores)	0.321	0.153	-0.168***	0.340	0.208	-0.132**	0.700	0.227	-0.473***
Numeracy Performance:									
Number ident. (% answered correctly)	0.439	0.625	0.186***	0.424	0.671	0.247***	0.441	0.721	0.280***
Addition (% answered correctly)	0.290	0.459	0.170***	0.316	0.480	0.164***	0.318	0.570	0.252***
Subtraction (% answered correctly)	0.241	0.401	0.160***	0.253	0.415	0.162***	0.266	0.502	0.236***
Number ident. (% with zero scores)	0.144	0.085	-0.059***	0.143	0.068	-0.076***	0.107	0.000	-0.107*
Addition (% with zero scores)	0.233	0.111	-0.122***	0.204	0.107	-0.097***	0.214	0.036	-0.179**
Subtraction (% with zero scores)	0.289	0.143	-0.146***	0.298	0.144	-0.154***	0.357	0.036	-0.321***
Observations	529	459		339	316		28	29	

Figure 3 displays mean EGRA/MA subtask scores by displacement status from baseline to endline. This figure shows the significant increases by each group across each subtask from baseline to endline. It also shows the relatively similar performance of each group in all subtasks at baseline and the differences in outcomes at endline, in which returnees perform the best, followed by host community members, then IDPs.

Figure 3. Mean EGRA/MA subtask scores by displacement status at baseline and endline



EGRA/MA Outcomes by *Age Group*

Table 9 shows mean differences in EGRA/MA subtask scores for learners ages 6-9 and learners ages 10-15 from baseline to endline. The table shows that the older group has statistically significant improvements across all subtasks and the younger group has statistically significant improvements across

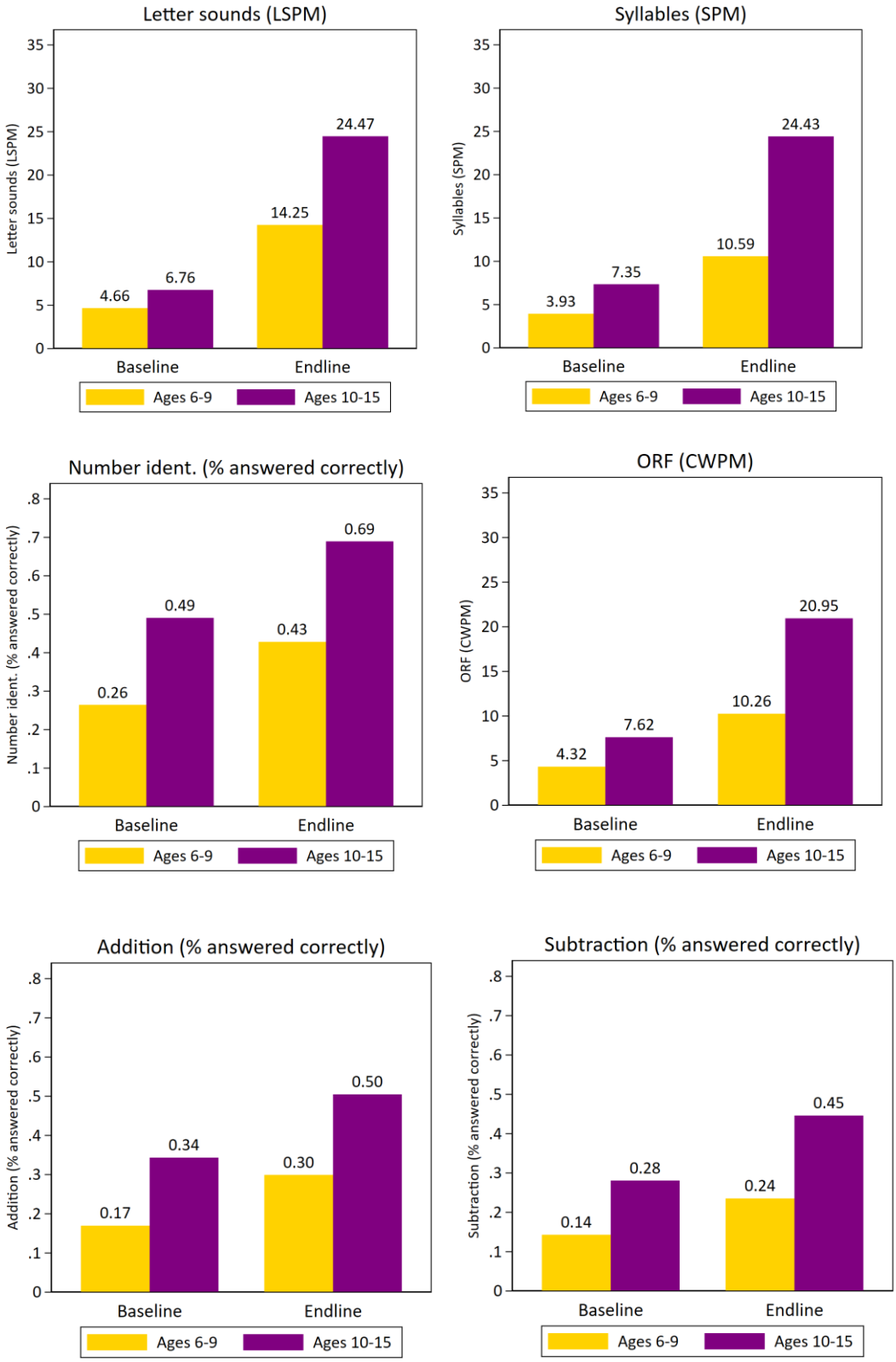
all subtasks from baseline to endline except reading comprehension. As expected, learners in the older age bracket have higher mean reading scores at baseline and endline. Older learners also have larger gains in mean reading scores from baseline to endline. For example, learners who are 6-9 years old had a mean ORF score of 4.3 words per minute at baseline and saw a significant increase to 10.3 words per minute by endline. On the other hand, learners who are 10-15 years old had a mean ORF score of 7.6 words per minute at baseline and saw a significant increase to 20.9 words per minute by endline. Additionally, a higher percentage of learners in the younger age bracket received zero scores on the four reading subtasks at both baseline and endline. The decrease in the percentage of learners who received zero scores on the letter sounds subtask was slightly larger than that of older learners. However, learners in the older age bracket saw larger decreases in the percentage of learners who received zero scores on syllables, ORF, and reading comprehension compared to younger learners. In terms of numeracy performance, learners in the older age bracket have higher mean math scores than younger learners at baseline and endline. Additionally, older learners have slightly larger gains in mean math scores from baseline to endline compared to younger learners. However, the decreases in the percentage of learners who receive zero scores on the math subtasks is larger than the decreases for older learners. It is important to note that the younger learners were beginning at higher percentages of learners with zero scores on the math subtasks.

Table 9. EGRA/MA outcomes by age group at baseline and endline

	Ages 6-9			Ages 10-15		
	Baseline	Endline	Diff.	Baseline	Endline	Diff.
Reading Performance:						
Letter sounds (LSPM)	4.664	14.246	9.582***	6.759	24.475	17.715***
Syllables (SPM)	3.934	10.588	6.655***	7.348	24.427	17.078***
ORF (CWPM)	4.315	10.256	5.941***	7.622	20.946	13.323***
Reading comp. (% answered correctly)	0.218	0.257	0.039	0.257	0.397	0.140***
Letter sounds (% with zero scores)	0.513	0.252	-0.261***	0.358	0.099	-0.259***
Syllables (% with zero scores)	0.607	0.425	-0.183***	0.464	0.181	-0.282***
ORF (% with zero scores)	0.639	0.467	-0.171***	0.505	0.217	-0.288***
Reading comp. (% with zero scores)	0.412	0.357	-0.055	0.336	0.161	-0.175***
Numeracy Performance:						
Number ident. (% answered correctly)	0.264	0.428	0.164***	0.490	0.689	0.199***
Addition (% answered correctly)	0.169	0.299	0.130***	0.343	0.505	0.161***
Subtraction (% answered correctly)	0.143	0.235	0.092***	0.281	0.446	0.165***
Number ident. (% with zero scores)	0.285	0.110	-0.175***	0.098	0.069	-0.030*
Addition (% with zero scores)	0.415	0.222	-0.193***	0.167	0.084	-0.083***
Subtraction (% with zero scores)	0.443	0.287	-0.156***	0.255	0.106	-0.149***
Observations	309			1303		

Figure 4 displays mean subtask scores by age group at baseline and endline, as seen in the table above. The figure shows that for each subtask, younger learners (ages 6-9) have lower mean scores than older learners (ages 10-15) at baseline and endline. We also see that for each subtask, both groups have substantial increases from baseline to endline. For example, mean number identification scores increase by approximately 17 percentage points from 26% to 43% for younger learners, and approximately 20 percentage points from 49% to 69% for older learners.

Figure 4. Mean EGRA/MA subtask score by age group at baseline and endline



EGRA/MA Outcomes by *Language*

Table 9 shows differences in mean EGRA/MA subtask scores by learner language spoken at home from baseline to endline. The table also shows the percentage of learners with zero scores for each subtask by language at baseline and endline. It is important to note that the majority of the sample is comprised of Hausa speakers, followed by Kanuri speakers, then speakers of other languages, and finally, Fulani speakers. As seen in the table, Hausa and Kanuri speakers have statistically significant improvements across all EGRA/MA outcomes. Fulani speakers have statisitically significant improvements across all but three EGRA/MA outcomes from baseline to endline. Speakers of other languages improve across all EGRA/MA outcomes from baseline to endline, and have statistically significant improvements for letter sounds, syllables, and math subtasks. In terms of reading performance, Hausa speakers have the highest mean scores at endline and have the largest margins of improvement from baseline to endline. However, these are only slightly higher than Kanuri speakers, who have similar mean reading scores at baseline. Additionally, the decrease in the percentae of Hausa speakers who receive zero scores on reading scores is lower than the decrease in the percentage of Kanuri speakers who receive zero scores on all reading subtasks, except syllables. However, it is important to note that at baseline, there were slightly higher percentages of Kanuri speakers who were receiving zero scores on the reading subtasks. Fulani speakers have slightly lower mean reading scores at baseline compared to Hausa and Kanuri speakers and have slightly smaller increases in mean reading scores from baseline to endline. Interestingly, speakers of other languages have comparable mean reading scores at baseline, but see substantially less improvement from baseline to endline compared to the other groups, particularly in ORF and reading comprehension scores. Additionally, the percentage of learners who speak other languages who receive zero scores on reading subtask does not decrease by as much as the percentages did for Hausa, Kanuri, and Fulani speakers.

In terms of numeracy performance, mean scores are relatively similar across all subtasks and groups at baseline. Hausa and Kanuri speakers have slightly larger improvements in mean math scores compared to Fulani speakers and speakers of other languages from baseline to endline. Thus, at endline, Hausa and Kanuri speakers have slightly higher mean math scores. At endline, Kanuri speakers have the lowest percentage of learners receiving zero scores across all three math subtasks.

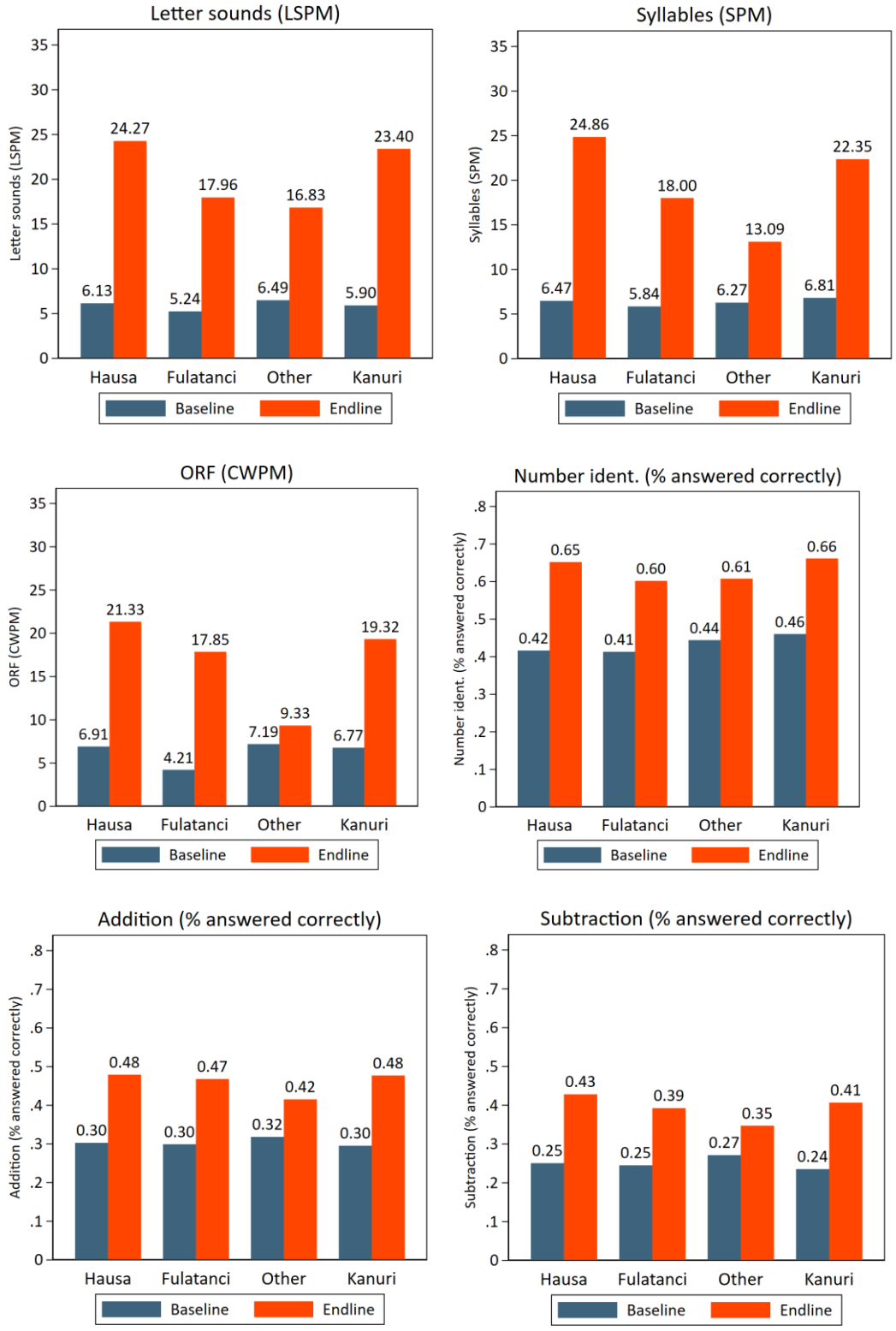
Table 10. EGRA/MA outcomes by language at baseline and endline

	Hausa			Kanuri			Fulani		
	Baseline	Endline	Diff.	Baseline	Endline	Diff.	Baseline	Endline	Diff.
Reading Performance:									
Letter sounds (LSPM)	6.135	24.266	18.131***	5.900	23.395	17.495***	5.240	17.960	12.720***
Syllables (SPM)	6.473	24.856	18.384***	6.815	22.350	15.536***	5.840	18.002	12.162***
ORF (CWPM)	6.907	21.329	14.422***	6.773	19.324	12.551***	4.210	17.853	13.643***
Reading comp. (% answered correctly)	0.278	0.398	0.120***	0.235	0.389	0.155***	0.200	0.320	0.120
Letter sounds (% with zero scores)	0.382	0.118	-0.264***	0.406	0.120	-0.286***	0.400	0.140	-0.260***
Syllables (% with zero scores)	0.475	0.172	-0.304***	0.509	0.239	-0.270***	0.480	0.288	-0.192**
ORF (% with zero scores)	0.499	0.208	-0.291***	0.576	0.252	-0.324***	0.620	0.358	-0.262***
Reading comp. (% with zero scores)	0.333	0.184	-0.149***	0.358	0.164	-0.194***	0.417	0.167	-0.250*

Numeracy Performance:									
Number ident. (% answered correctly)	0.416	0.652	0.235***	0.460	0.661	0.201***	0.413	0.602	0.189***
Addition (% answered correctly)	0.303	0.479	0.176***	0.295	0.477	0.181***	0.299	0.468	0.169***
Subtraction (% answered correctly)	0.251	0.428	0.177***	0.235	0.407	0.171***	0.245	0.392	0.147***
Number ident. (% with zero scores)	0.173	0.102	-0.071***	0.107	0.034	-0.073***	0.140	0.057	-0.083
Addition (% with zero scores)	0.225	0.133	-0.092***	0.228	0.064	-0.164***	0.220	0.113	-0.107
Subtraction (% with zero scores)	0.288	0.148	-0.140***	0.313	0.098	-0.215***	0.340	0.189	-0.151*
Observations	819			507			103		

Figure 5 displays mean EGRA/MA subtask scores from the table above. The figure shows that across the reading subtasks, speakers of all languages performed similarly at baseline. However, at endline, Hausa speakers had the highest mean scores, followed closely by Kanuri speakers. Speakers of other languages had the lowest mean reading scores at endline. In terms of numeracy performance, the figure shows that mean math scores were similar across all groups at baseline and endline, with speakers of other languages performing slightly worse on addition and subtraction at endline.

Figure 5. Mean ERGA/MA subtask scores by language at baseline and endline



Social-emotional Learning Outcomes

In addition to assessing children's reading and numeracy skills using EGRA/MA, we assessed learner Socio-emotional Learning (SEL) outcomes using two modules adapted for the northeast Nigeria context. We deployed the Assessment of Children's Emotional Skills (ACES), which measures the ability of children to recognize the emotions of other people accurately. In this module, learners are presented with a series of scenarios in which a child in a story may feel anger, sadness, happiness, or fear, and the respondent is asked to identify the emotion that the child in the story should feel. Second, we deployed the Children's Stories module, which includes an ambiguous scenario in which one child does something to another child in a story. The respondent is asked to interpret whether the child in the story intended to hurt the other child or if it was an accident. This module helps to measure hostile attribution bias, which assess the tendency of children to attribute hostile intent to a third party in response to a provocation. Additionally, the Children's Stories module measures conflict resolution strategies employed by respondents by asking how the respondent would react in the scenario presented. Conflict resolution strategies include responses that are categorized as aggression, disengagement, or problem solving.

Table 11 shows differences in mean SEL outcomes from baseline to endline for the entire sample. From baseline to endline, both mean ACES scores and the percentage of pupils with perfect ACES scores increase by statistically significant margins. Additionally, the percentage of pupils who display hostile attribution bias decreases by 11.4 percentage points from baseline to endline, a statistically significant margin. From baseline to endline, the percentage of pupils who report feeling angry after listening to a story decreases by a statistically significant margin of 6.9 percentage points. In terms of conflict resolution strategies employed by pupils, the percentageof pupils who use aggressive conflict resolution strategies decreases by a statistically significant margin from baseline to endline, while the percentage of pupils who disengage or use problem solving techniques increases, but by margins that are not statistically significant.

Table 11. Overall SEL summary scores at baseline and endline

	Baseline	Endline	Diff.
ACES:			
ACES score (% answered correctly)	0.822	0.848	0.026***
Perfect ACES score (%)	0.297	0.359	0.062***
Displays hostile attribution bias	0.403	0.289	-0.114***
Feeling reported by student after story:			
Calm	0.288	0.318	0.030
Surprised	0.135	0.189	0.054***
Sad	0.312	0.297	-0.015
Angry	0.265	0.196	-0.069***
Conflict resolution strategies:			
Aggression	0.147	0.104	-0.043**
Disengagement	0.338	0.352	0.014
Problem Solving	0.515	0.544	0.029
Observations	1630		

Table 11 shows differences in SEL summary scores by state and gender from baseline to endline. Across both states and genders, learners have slight improvements in mean ACES scores from baseline to endline, with statistically significant improvements for males in both states. It is important to note that ACES scores were already relatively high at baseline. The percentage of learners who received perfect ACES scores increased by a statistically significant margin for boys in both states. ACES outcomes are slightly better for learners in Borno compared to learners in Yobe. In Borno, more learners display

hostile attribution bias at both baseline and endline compared to learners in Yobe. However, the percentage of learners who display hostile attribution bias decreases by a statistically significant margin from baseline to endline for all groups except for Borno males. Conflict resolution strategies stay relatively stagnant from baseline to endline across all groups. Problem solving is the most used technique for all groups at both baseline and endline. Notably, Yobe males have a statistically significant increase in the use of problem solving and a statistically significant decrease in the use of aggression from baseline to endline.

Table 12. SEL summary scores by state and gender at baseline and endline

	Borno					
	Male			Female		
	Baseline	Endline	Diff.	Baseline	Endline	Diff.
ACES:						
ACES score (% answered correctly)	0.828	0.864	0.037**	0.843	0.860	0.016
Perfect ACES score (%)	0.281	0.451	0.169***	0.363	0.434	0.071
Displays hostile attribution bias	0.496	0.426	-0.070	0.477	0.349	-0.128***
Feeling reported by student after story:						
Calm	0.339	0.364	0.025	0.302	0.286	-0.016
Surprised	0.094	0.123	0.030	0.109	0.104	-0.005
Sad	0.268	0.253	-0.015	0.341	0.349	0.007
Angry	0.299	0.259	-0.040	0.248	0.261	0.013
Conflict resolution strategies:						
Aggression	0.155	0.154	-0.001	0.165	0.120	-0.045
Disengagement	0.329	0.385	0.056	0.313	0.365	0.052
Problem Solving	0.516	0.462	-0.055	0.522	0.515	-0.007
Observations	386			501		
	Yobe					
	Male			Female		
	Baseline	Endline	Diff.	Baseline	Endline	Diff.
ACES:						
ACES score (% answered correctly)	0.783	0.828	0.044**	0.816	0.838	0.021
Perfect ACES score (%)	0.179	0.267	0.088*	0.310	0.271	-0.039
Displays hostile attribution bias	0.286	0.187	-0.099**	0.305	0.192	-0.114***
Feeling reported by student after story:						
Calm	0.200	0.313	0.113**	0.276	0.322	0.046
Surprised	0.157	0.273	0.116**	0.188	0.276	0.087**
Sad	0.350	0.273	-0.077	0.301	0.290	-0.012
Angry	0.293	0.140	-0.153***	0.234	0.112	-0.122***
Conflict resolution strategies:						
Aggression	0.138	0.047	-0.090***	0.126	0.089	-0.037
Disengagement	0.377	0.324	-0.052	0.348	0.333	-0.014
Problem Solving	0.486	0.628	0.143**	0.526	0.577	0.051
Observations	290			453		

Table 12 shows the same outcomes as Table11, but disaggregated by displacement status. It is important to again note the small sample size for returnees. Mean ACES scores and the percentage of learners with perfect ACES scores are slightly higher for IDPs compared to host community members and

returnees. Additionally, ACES outcomes improve by a statistically significant margin for IDPs from baseline to endline. IDPs have the highest percentage of learners who display hostile attribution bias at both baseline and endline, however, they see a statistically significant decrease from 42% to 33% from baseline to endline. The percentage of host community members who display hostile attribution bias also decreases by a statistically significant margin from 38% to 24% from baseline to endline. Conflict resolution strategies used remain relatively similar across all groups from baseline to endline, with problem solving being the most used technique for all three groups at baseline and endline. Host community members have a statistically significant decrease in the use of aggressive strategies from baseline to endline as only 7% of these learners use aggressive techniques at endline.

Table 13. SEL summary scores by displacement status at baseline and endline

	IDP			Host Community			Returnee		
	Baseline	Endline	Diff.	Baseline	Endline	Diff.	Base	End	Diff.
ACES:									
ACES score (% answered correctly)	0.816	0.857	0.041***	0.832	0.836	0.004	0.808	0.848	0.040
Perfect ACES score (%)	0.300	0.404	0.103***	0.302	0.296	-0.006	0.179	0.357	0.179
Displays hostile attribution bias	0.418	0.331	-0.087***	0.384	0.237	-0.147***	0.357	0.214	-0.143
Feeling reported by student after story:									
Calm	0.301	0.294	-0.007	0.278	0.345	0.067*	0.179	0.393	0.214*
Surprised	0.124	0.170	0.047**	0.151	0.217	0.066**	0.143	0.179	0.036
Sad	0.287	0.324	0.037	0.335	0.263	-0.072**	0.500	0.250	-0.250*
Angry	0.289	0.211	-0.077***	0.236	0.174	-0.061*	0.179	0.179	0.000
Conflict resolution strategies:									
Aggression	0.161	0.125	-0.037	0.125	0.074	-0.051**	0.148	0.107	-0.041
Disengagement	0.344	0.369	0.025	0.342	0.330	-0.012	0.185	0.321	0.136
Problem Solving	0.495	0.506	0.011	0.533	0.596	0.063	0.667	0.571	-0.095
Observations	939			635			56		

Assessment of Children's Emotional Skills

Table 13 shows differences in the percentage of learners who answered each of the eight ACES stores correctly by state and gender from baseline to endline. The table also shows the same summary scores from Table 10 – mean overall ACES scores and the percentage of learners with perfect ACES scores. Overall ACES scores are calculated by adding the number of items that a learner answered correctly and dividing by the total number of ACES items (8). Perfect ACES scores are calculated by counting the number of learners who responded correctly to all eight ACES stories and dividing by the total number of learners.

Table 14. ACES components and scores by state and gender at baseline and endline

	Borno					
	Male			Female		
	Baseline	Endline	Diff.	Baseline	Endline	Diff.
ACES 1	0.826	0.840	0.014	0.869	0.884	0.016
ACES 2	0.763	0.759	-0.004	0.768	0.798	0.029
ACES 3	0.857	0.963	0.106***	0.822	0.926	0.103***
ACES 4	0.786	0.790	0.004	0.811	0.831	0.020
ACES 5	0.911	0.883	-0.028	0.911	0.884	-0.027
ACES 6	0.804	0.889	0.085**	0.830	0.831	0.000
ACES 7	0.777	0.864	0.087**	0.819	0.851	0.033

ACES 8	0.897	0.926	0.029	0.915	0.872	-0.043
ACES score (% answered correctly)	0.828	0.864	0.037**	0.843	0.860	0.016
Perfect ACES score (%)	0.281	0.451	0.169***	0.363	0.434	0.071
Observations	386			501		

Yobe

	Male			Female		
	Baseline	Endline	Diff.	Baseline	Endline	Diff.
ACES 1	0.921	0.853	-0.068*	0.870	0.874	0.004
ACES 2	0.629	0.713	0.085	0.674	0.715	0.041
ACES 3	0.829	0.927	0.098**	0.833	0.902	0.069**
ACES 4	0.664	0.867	0.202***	0.770	0.855	0.085**
ACES 5	0.943	0.893	-0.050	0.937	0.897	-0.040
ACES 6	0.679	0.693	0.015	0.812	0.762	-0.050
ACES 7	0.714	0.773	0.059	0.749	0.794	0.045
ACES 8	0.886	0.900	0.014	0.887	0.902	0.015
ACES score (% answered correctly)	0.783	0.828	0.044**	0.816	0.838	0.021
Perfect ACES score (%)	0.179	0.267	0.088*	0.310	0.271	-0.039
Observations	290			453		

Figure 6 displays mean ACES scores by gender and displacement status at baseline and endline. From this figure, we can see that ACES scores slightly increase for boys and girls in this sample from baseline to endline. We can also see that mean ACES scores slightly increased for IDPs, host community members, and returnees from baseline to endline.

Figure 6. Mean ACES scores by gender and displacement status at baseline and endline

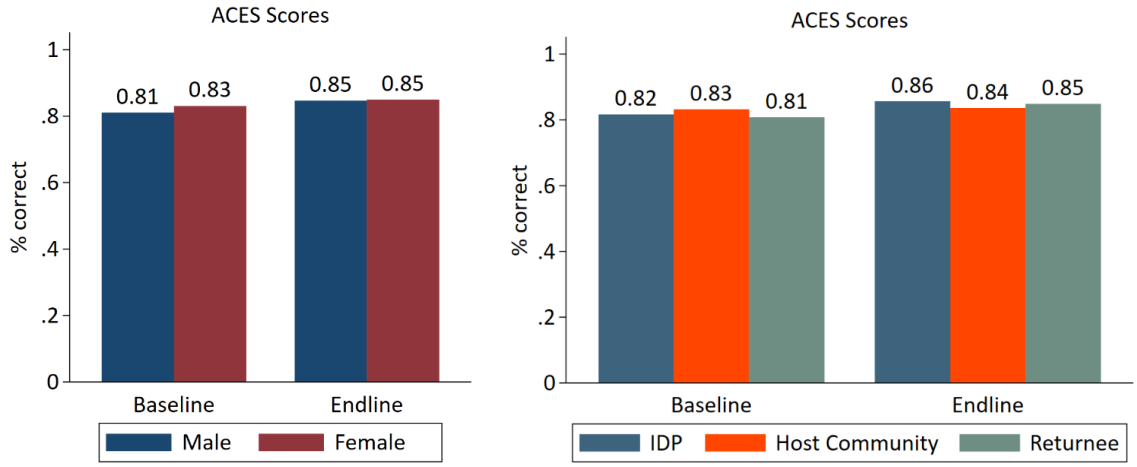
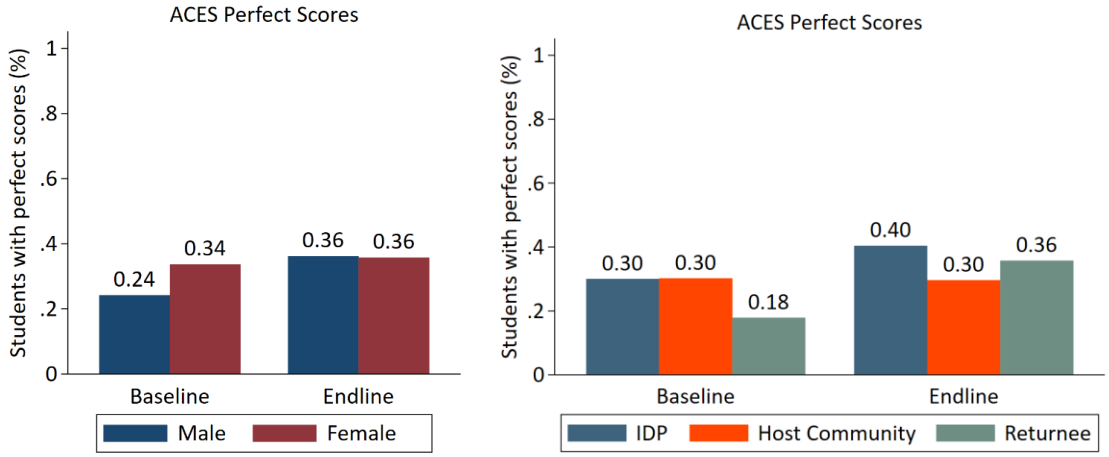


Figure 7 displays the percentage of learners with perfect ACES scores by gender and displacement status at baseline and endline. This figure shows that at baseline, 24% of males and 34% of females across the sample received perfect ACES scores. At endline, 36% of males and females received perfect ACES scores. This figure also shows that the percentage of IDPs who received perfect ACES scores increases by 10% from baseline to endline, while there was no change in the percentage of host community members who received perfect ACES scores from baseline to endline.

Figure 7. Percentage of learners with perfect ACES scores by gender and displacement status at baseline and endline



Children's Stories: Hostile Attribution Bias and Conflict Resolution

The Children's Stories module allows us to measure hostile attribution bias and conflict resolution strategies employed by learners. Learners display hostile attribution bias if they respond to a story by saying that a child did an act on purpose, as opposed to saying that the event was an accident. After being asked how they would react the given scenario, learners' responses are categorized into conflict resolution strategies, including aggression, disengagement, and problem solving. Table 14 shows the difference in the percentage of learners who display hostile attribution bias by state and gender from baseline to endline. The table also shows the percentage of learners who use each type of conflict resolution strategy by state and gender at baseline and endline. These results were discussed in summary of SEL results.

Table 15. Hostile attribution bias and conflict resolution by state and gender at baseline and endline

	Borno					
	Male			Female		
	Baseline	Endline	Diff.	Baseline	Endline	Diff.
Displays hostile attribution bias	0.496	0.426	-0.070	0.477	0.349	-0.128***
Feeling reported by student after story:						
Calm	0.339	0.364	0.025	0.302	0.286	-0.016
Surprised	0.094	0.123	0.030	0.109	0.104	-0.005
Sad	0.268	0.253	-0.015	0.341	0.349	0.007
Angry	0.299	0.259	-0.040	0.248	0.261	0.013
Conflict resolution strategies:						
Aggression	0.155	0.154	-0.001	0.165	0.120	-0.045
Disengagement	0.329	0.385	0.056	0.313	0.365	0.052
Problem Solving	0.516	0.462	-0.055	0.522	0.515	-0.007
Observations	386			501		
	Yobe					
	Male			Female		
	Baseline	Endline	Diff.	Baseline	Endline	Diff.
Displays hostile attribution bias	0.286	0.187	-0.099**	0.305	0.192	-0.114***
Feeling reported by student after story:						
Calm	0.200	0.313	0.113**	0.276	0.322	0.046
Surprised	0.157	0.273	0.116**	0.188	0.276	0.087**
Sad	0.350	0.273	-0.077	0.301	0.290	-0.012
Angry	0.293	0.140	-0.153***	0.234	0.112	-0.122***

Conflict resolution strategies:						
Aggression	0.138	0.047	-0.090***	0.126	0.089	-0.037
Disengagement	0.377	0.324	-0.052	0.348	0.333	-0.014
Problem Solving	0.486	0.628	0.143**	0.526	0.577	0.051
Observations	290			453		

Figure 8 displays the percentage of learners who display hostile attribution bias by gender and displacement status at baseline and endline. The figure shows that both the percentages of boys and girls who display hostile attribution bias decreases from baseline to endline, with girls being slightly less likely to display hostile attribution bias. Additionally, we also see that a slightly higher percentage of IDPs display hostile attribution bias compared to host community members. However, both groups see substantial decreases in the percentage of learners who display hostile attribution bias from baseline to endline.

Figure 8. Percentage of learners who display hostile attribution bias by gender and displacement status at baseline and endline

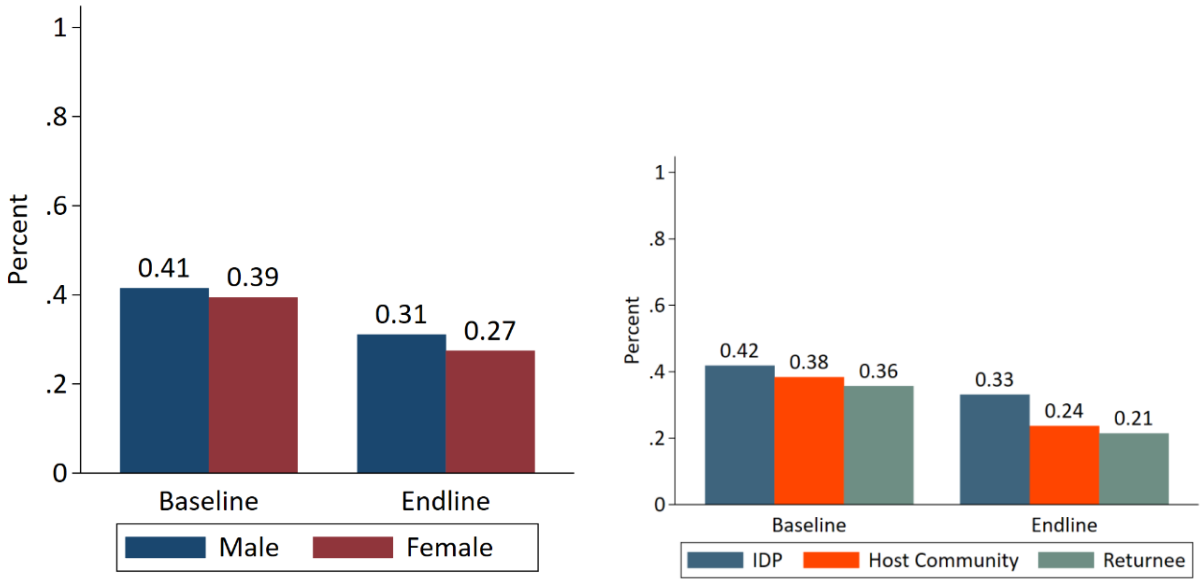
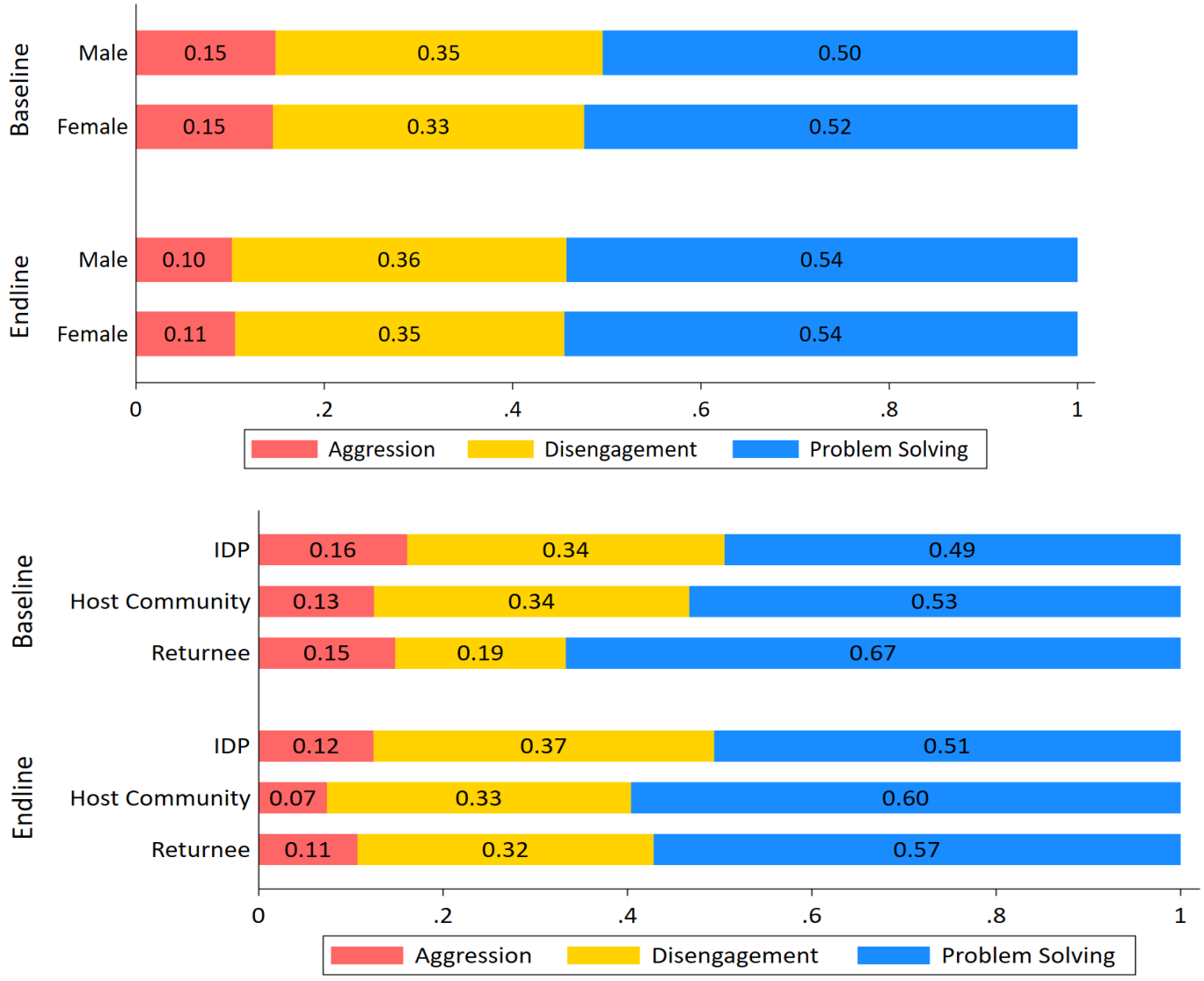


Figure 9 shows the percentage of learners who use each type of conflict resolution strategy by gender and displacement status at baseline and endline. As seen in the figure, problem solving is the most used strategy across all groups at both baseline and endline. Disengagement is the second most used technique, followed by aggressive strategies. Across all groups, we see slight decreases in the use of aggressive strategies from baseline to endline. We also see slight increases in the use of problem solving techniques from baseline to endline in all groups, except for returnees.

Figure 2. Conflict resolution strategies by gender and displacement status at baseline and endline



School Safety Outcomes

To measure pupils' perceptions of the NLFC learning environment, we use the school climate questionnaire from USAID's Safer Learning Environment Toolkit, which assesses if children feel that their teachers create a safe learning environment and whether children are treated fairly and with respect. We also measure safety of learners by asking four questions related to community and school safety.

Table 16 shows mean school climate and safety scores from baseline to endline for the entire sample. Both scores increase by a statistically significant margin from baseline to endline, with mean school climate scores increasing from 80.2% at baseline to 86.1% at endline, and mean safety scores increasing from 90.6% at baseline to 95.5% at endline.

Table 16. Overall school climate and safety mean scores at baseline and endline

	Baseline	Endline	Diff.
School climate score (%)	0.802	0.861	0.060***
Safety score (%)	0.906	0.955	0.049***
Observations	1630		

Table 17 shows differences in mean school climate and safety scores by state and gender at baseline and endline. Across all groups, there are increases in school climate and safety scores from baseline to endline, with all but one change being statistically significant. These scores are generally high at baseline and endline, with safety scores all over 94% for each group at endline.

Table 17. Mean school climate and safety scores by state and gender at baseline and endline

	Borno					
	Male			Female		
	Baseline	Endline	Diff.	Baseline	Endline	Diff.
School climate score (%)	0.784	0.817	0.034**	0.770	0.829	0.060***
Safety score (%)	0.922	0.946	0.024	0.907	0.955	0.048***
Observations	386			501		
	Yobe					
	Male			Female		
	Baseline	Endline	Diff.	Baseline	Endline	Diff.
School climate score (%)	0.820	0.902	0.082***	0.843	0.901	0.059***
Safety score (%)	0.889	0.967	0.077***	0.902	0.954	0.053***
Observations	290			453		

Table 18 shows differences in mean school climate and safety scores by displacement status from baseline to endline. Again, we see statistically significant increases across all groups, besides returnee mean school climate scores. Mean school climate scores increase by 5-6 percentage points for all three groups, with increases for IDPs and host community members being statistically significant. Mean safety scores increase by 3.7-5.6 percentage points from baseline to endline and are statistically significant increases for all three groups.

Table 18. Mean school climate and safety scores by displacement status at baseline and endline

	IDP			Host Community			Returnee		
	Baseline	Endline	Diff.	Baseline	Endline	Diff.	Baseline	Endline	Diff.
School climate score (%)	0.794	0.854	0.060***	0.816	0.875	0.059***	0.774	0.825	0.052
Safety score (%)	0.898	0.955	0.056***	0.917	0.954	0.037***	0.929	0.982	0.054*
Observations	939			635			56		

School Climate

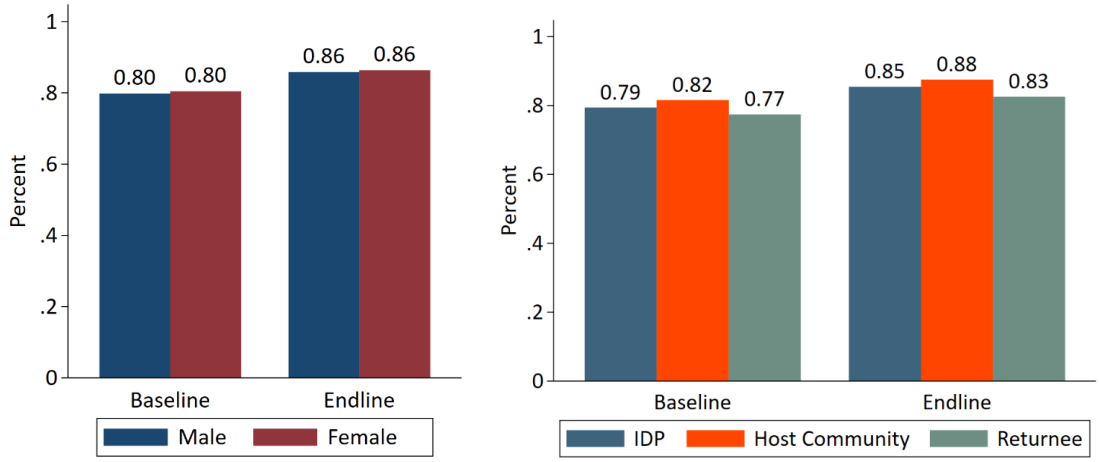
This section examines results from the school climate module. In total, learners are asked about the extent to which they agree with nine school climate statements. Table 15 shows the percentage of learners who agree with each individual statement by state and gender at baseline and endline. The table also shows mean school climate scores, which are calculated by adding the number of school climate questions that students agree with and dividing by the total number of school climate statements (9).

Table 19. School climate components and scores by state and gender at baseline and endline

	Borno						Yobe					
	Male			Female			Male			Female		
	Baseline	Endline	Diff.	Baseline	Endline	Diff.	Baseline	Endline	Diff.	Baseline	Endline	Diff.
Students treat each other with respect	0.920	0.993	0.073***	0.923	0.987	0.064***	0.970	0.986	0.017	0.991	0.986	-0.005
Teachers listen to students	0.907	0.961	0.054**	0.912	0.968	0.056**	0.955	0.986	0.031	0.969	0.990	0.021
Boys and girls are not nice to each other	0.514	0.447	-0.067	0.481	0.454	-0.027	0.363	0.229	-0.134**	0.378	0.205	-0.174***
Students often treat disabled students unkindly	0.325	0.139	-0.186***	0.319	0.176	-0.143***	0.336	0.121	-0.215***	0.291	0.089	-0.203***
Teachers treat girls and boys equally	0.953	0.966	0.013	0.883	0.974	0.090***	0.962	0.979	0.018	0.978	0.990	0.013
Teachers treat students of all races the same	0.938	0.958	0.020	0.908	0.972	0.064***	0.947	0.980	0.032	0.978	0.986	0.008
Students asked to help decide what is best for class	0.650	0.664	0.015	0.715	0.731	0.016	0.695	0.847	0.152***	0.667	0.810	0.143***
Students rewarded when they do well on classwork	0.950	0.961	0.011	0.915	0.992	0.076***	0.992	0.993	0.001	0.987	0.995	0.008
Students know who to report to when they experience or witness violence	0.962	1.000	0.038**	0.967	0.974	0.007	0.969	0.987	0.017	0.982	0.986	0.003
School climate score (%)	0.784	0.817	0.034**	0.770	0.829	0.060***	0.820	0.902	0.082***	0.843	0.901	0.059***
Observations	381			496			288			447		

Figure 9 shows mean school climate scores by gender and displacement status at baseline and endline. The figure shows that boys and girls in the sample had mean school climate scores of 80% at baseline, and these scores increased to 86% at endline. We can also see that scores increased for IDPs, host community members, and returnees from baseline to endline.

Figure 3. Mean school climate scores by gender and displacement status at baseline and endline



Perceptions of School (NFLC) Safety

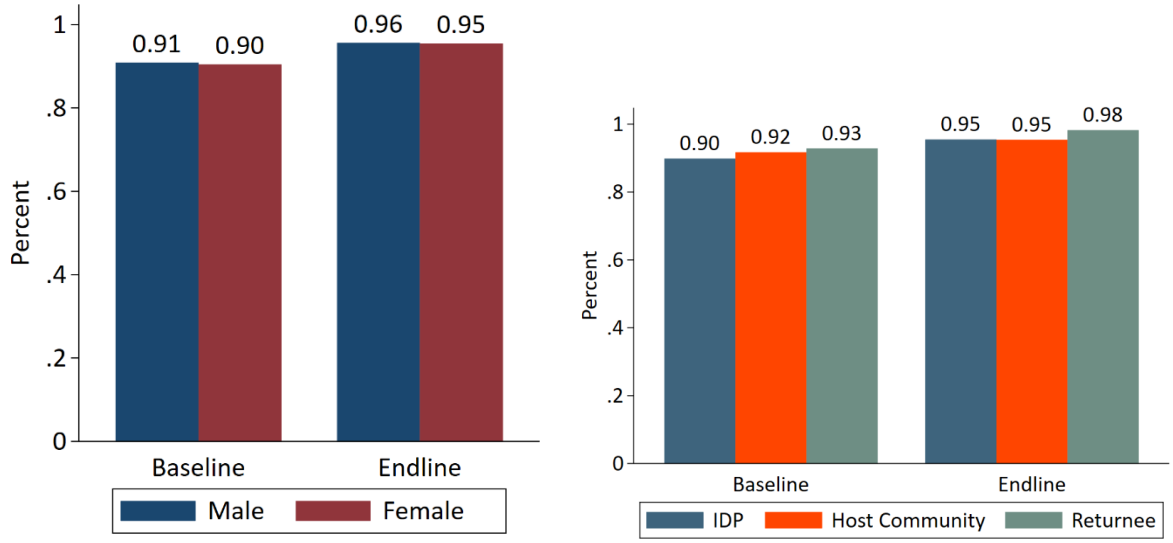
Table 16 shows each individual safety item and overall mean safety scores by state and gender at baseline and endline. For each of the four safety items, the mean shows the percentage of students who agree with the statement. Overall safety scores are calculated by adding the number of safety items that a learner agrees with and dividing by the total number of safety items (4). As seen in the table, mean safety scores are generally high at baseline and increase to above 94% for all groups from baseline to endline.

Table 20. Safety components and scores by state and gender at baseline and endline

	Borno					
	Male			Female		
	Baseline	Endline	Diff.	Baseline	Endline	Diff.
Feels safe at school	0.973	0.943	-0.029	0.957	0.979	0.022
Has safe places to play in school	0.955	0.975	0.020	0.950	0.967	0.017
Feels safe in community	0.927	0.963	0.035	0.906	0.979	0.073***
Has safe places to play in community	0.861	0.926	0.065**	0.837	0.929	0.092***
Safety score (%)	0.922	0.946	0.024	0.907	0.955	0.048***
Observations	385			499		
	Yobe					
	Male			Female		
	Baseline	Endline	Diff.	Baseline	Endline	Diff.
Feels safe at school	0.935	0.993	0.058***	0.966	0.991	0.025*
Has safe places to play in school	0.886	0.973	0.088***	0.912	0.986	0.074***
Feels safe in community	0.913	0.993	0.080***	0.944	0.967	0.023
Has safe places to play in community	0.843	0.913	0.070*	0.837	0.883	0.046
Safety score (%)	0.889	0.967	0.077***	0.902	0.954	0.053***
Observations	290			453		

Figure 10 shows mean safety scores by gender and displacement status at baseline and endline. As seen in the figure, mean safety scores for boys and girls are both approximately 90% at baseline and increase to approximately 95% at endline. Mean safety scores also increase for IDPs, host community members, and returnees from baseline to endline, with IDPs and host community members having mean safety scores of 95% at endline.

Figure 4. Mean safety scores by gender and displacement status at baseline and endline



Relationship between EGRA/EGMA and SEL Results

When correlating social-emotional learning indicators with reading performance, there does not appear to be a relationship. Most correlation coefficients are below 0.1. For example, when comparing ACES scores to letter sound scores, the correlation coefficient is 0.10. For ACES and syllable scores, the correlation coefficient is 0.14, and for ACES and ORF scores, the correlation coefficient is 0.11. The correlation coefficient for ACES scores and reading comprehension scores is 0.08.

4. CONCLUSION

Overall, although this is not an impact evaluation, the results from this pre-post assessment are quite clear: NFLC participants are learning, and at a high enough rate that it is reasonable to assume that some participants could feasibly mainstream into Grade 4 in the formal system or move on to Post-Basic NFLCs, which are primarily taught in English. While the academic results appear high in comparison to early grade reading programs from other parts of Northern Nigeria, it is important to remember that, on average, AENN targets an older age group, and older participants should learn at a faster rate than younger children. Improving educational equity for over-age out-of-school children and girls are both goals of the AENN program, and the EGRA and EGMA results show that the groups benefiting most from the intervention include older children (ages 10-15) and girls. Both IDPs and host community children and youth have access to the program, as designed, and both groups appear to benefit similarly from their participation. All of these are results are positive and paint a picture of an educational intervention that is working in the way that it was designed.

The SEL and school safety results are less dramatic, but still positive. These contructs are much more difficult to measure objectively than EGRA/EGMA, and often suffer from high baselines, which limits opportunity for growth. Still, all results point in the direction of improved social-emotional skills and perceptions of school safety over the course of the seven month intervention. When correlating social-emotional learning indicators with reading performance, there does not appear to be a relationship.

The primary recommendation of this report is that the NFLCs are working as designed and should be continued. The second cohort of NFLCs is already underway, having begun in February 2020, with several new elements: the program is now being offerent at the Basic level in both Hausa and Kanuri, and Post-Basic is available. It will be useful to consider these literacy results in light of the reading benchmarking and policy linking conversations that are taking place in Nigeria this year with USAID's support and participation.

Another recommendation relates to fine-tuning the metrics for social-emotional learning (SEL) and improved child well-being, which is one of the AENN Activity indicators. Since reliable metrics for SEL are still under development globally, the AENN team used the best tools available at the time when the baseline was carried out in July 2019. Based on the results from this study, the team has determined that the ACES modules, which were designed for use with younger children, do not appear to be particularly reliable or valid with this population. On the other hand, the Children's Stories modules are more appropriate and valid given the age group and context of the intervention. Therefore, the team recommends to drop ACES from the Cohort 2 assessments and replace it with several additional Children's Stories (the current tool has only one story; the revised version would increase it up to five stories) and also add a very brief module on participants' mental health. We believe this will provide improved insight into the effects of the NFLCs on SEL and learner well-being.

5. APPENDIX

Sample Attrition Analysis

As discussed earlier, the original sampling plan involved sampling 4 children from each of about 394 NFLCs twice, once at baseline and again at endline. Due to budgetary and programmatic issues, 198 out of 394 NFLCs from baseline were not visited at endline to complete the second round of EGRA/MA and SEL survey data collection. However, at endline we collected data from 43 new NFLCs to replace some of the attrited centers from the baseline. As is the case with any instances of attrition, there are some concerns regarding the internal and external validity of the findings if the composition of the sampled NFLCs significantly changed between baseline and endline. As such, we complete simple t-tests of the differences in baseline child characteristics, learning outcomes, and SEL outcomes between the group of schools who dropped out of the sample and the group that remained.

Table A1, below displays the baseline differences in observed child characteristics from the NFLCs who remained and dropped out of the sample. Overall, we find that the average characteristics of children were mostly stable between the schools who dropped out and those who did not with the exception of language spoken at home, whether the child lives with both parents, cognitive disability, and the state in which the NFLC is located. We find that the NFLCs that dropped out of the sample have a higher percentage of children who speak Kanuri (smaller percentage of Hausa speakers), while children from the remaining sample were more likely to live with both parents (as opposed to living with only one or neither). This exercise also shows that the NFLCs that dropped out were more likely to be located in Borno state.

Table A1. Attrition test – Baseline characteristics

	Remained	Dropped	Diff
Female	0.58	0.55	0.03
Age	11.08	11.20	-0.12
Hausa	0.50	0.39	0.11***
Kanuri	0.32	0.38	-0.07***
Attended formal school last year	0.19	0.18	0.01
Both parents	0.68	0.58	0.11***
Works outside home or in fields	0.31	0.32	-0.01
Time spent with friends outside school	3.87	3.92	-0.06
Days learner was absent last week	0.61	0.53	0.09
Low HLE	0.54	0.52	0.02
Low SES	0.66	0.63	0.04
Has a physical disability	0.04	0.03	0.01
Has a cognitive disability	0.08	0.03	0.05***
IDP	0.59	0.61	-0.02
Host community	0.38	0.34	0.03
Yobe	0.43	0.26	0.18***

Next, we test for non-random attrition in baseline outcomes and present the results in Table A2. Similar to the results from Table A1, we find that samples are somewhat balanced along the majority of the outcomes with the exception of scores on the reading assessment, hostile attribution bias, and ACES. Moreover, the children in the NFLCs that dropped out of the sample had slightly higher baseline letter

sound, syllable, and oral reading fluency scores by 2.5, 2.4, and 1.5 points per minute, respectively. Although, it is important to mention that although the averages may have been slightly different between the two groups, the percentage of zero scores across all EGRA and EGMA subtasks were not statistically significant. In terms of SEL outcomes, we find that the children from the dropped sample had higher hostile attribution bias and a higher ACES score, by about 7 and 8 percentage points, respectively.

Table A2. Attrition test – Baseline outcomes

	Remained	Dropped	Diff
Letter sounds (LSPM)	6.13	8.63	-2.50***
Syllables (SPM)	6.44	8.82	-2.38***
ORF (CWPM)	6.73	8.18	-1.45**
Reading comp. (% answered correctly)	0.25	0.29	-0.04
Number ident. (% answered correctly)	0.43	0.45	-0.02
Addition (% answered correctly)	0.30	0.31	-0.01
Subtraction (% answered correctly)	0.25	0.27	-0.02
Letter sounds (% with zero scores)	0.40	0.43	-0.04
Syllables (% with zero scores)	0.49	0.51	-0.02
ORF (% with zero scores)	0.53	0.54	-0.01
Reading comp. (% with zero scores)	0.34	0.27	0.07
Number ident. (% with zero scores)	0.14	0.17	-0.02
Addition (% with zero scores)	0.22	0.24	-0.02
Subtraction (% with zero scores)	0.29	0.30	0.00
Displays hostile attribution bias	0.30	0.37	-0.07***
Conflict solving - problem solving	0.40	0.42	-0.02
ACES score (% answered correctly)	0.51	0.59	-0.08***
Victimized	0.40	0.42	-0.02

Across both sets of attrition tests, we maintain a fairly stable sample of children, with potential bias in favor of the NFLCs that dropped out of the sample especially in terms of baseline outcomes. The major determinant of attrition, however, seems to be the state in which the NFLC is located which could explain some of the differences in reading averages. This means that potential bias may be introduced to the analysis due to the non-random nature of the attrition, especially when the NFLCs that were not visited a second time were of a slightly higher reading level and may have consequences to the generalizability to the entire project population. This also means, however, that the direction of any consequent pre-post analysis using only the remaining sample of NFLCs may, overall, be biased downward. Nevertheless, because of the relatively small magnitude of the statistically significant differences, we would not expect any potential bias to be similarly small. We apply caution when interpreting the results of such analyses and maintain that findings reported in this document are not causal.